WISS 2013 について

五十嵐 健夫*

本論文集は，日本ソフトウェア科学会インタラクティブシステムとソフトウェア (ISS) 研究会が主催し，2013 年 12 月 4 日 (水) から 6 日 (金) にかけて，ザ クラウンパレス新阪急高知 (高知県高知市本町) で開催される「第 21 回インタラクティブシステムとソフトウェアに関するワークショップ」(Workshop on Interactive Systems and Software: WISS 2013) での発表をまとめたものです．

WISS は，2 泊 3 日の泊り込み形式で，インタラクティブシステムにおける未来を切り拓くような新しいアイディア・技術を議論するワークショップです．1993 年にスタートした WISS は今年で 21 回目を迎えました．この分野において国内でもっともアクティブな学術会議のひとつであり，近年では，170 名以上の参加者が，朝から深夜まで活発で意義深い情報交換をおこなっています．

本年は，登壇発表に対して，66 件のフルペーパーの投稿がありました．それぞれに対してプログラム委員から 3 名の査読者を割り当てて査読をおこないました．各論文に 1 名のプライマリ査読者を割り当てたほか，各論文を担当した 3 名の査読者が査読システム上で議論を行うための掲示板を設けました．また，必要な場合には，外部査読者の追加も行いました．そして，プログラム委員会で詳細に検討した結果，投稿数の 28.8%の 19 件の論文を採択しました．

カメラレディの提出後に，プログラム委員による投票を行い，論文賞を選出しました．その結果，以下の論文が WISS2013 論文賞に選ばれました．おめでとうございます．

SmartVoice：言語の壁を越えたプレゼンテーションサポーティングシステム 李 翔 (東大) 暦本 純一 (東大／ Sony CSL)

登壇発表論文に加えて，本論文集にはデモ／ポスター発表の論文が 46 件掲載されています．これらは，査読をせずに発表申し込みのあった全件を受け付けました．この他にも，当日は論文集には掲載されないデモ／ポスターが展示される予定です．また，海外の学会等で発表されていた研究を中心に，招待デモの展示も行います．実際に動くシステムを見ながらインタラクティブに議論をすることで，発表者は参加者から多くのフィードバックを得る事ができ，また参加者はその研究の本質をより深く理解することが可能となります．

第 12 回までは，論文集に論文 PDF およびデモビデオなどを収録した資料 CD-ROM を添付しておりましたが，ネットワークインフラの充実によりその使命を終えたと判断し，添付しておりません．論文 PDF およびデモビデオ等は，本研究会ホームページで公開する予定です．なお，今年初めての試みとして，採録論文については査読コメントをホームページ上で論文とともに公開予定です．

さらに今年も，昨年に続き，講演の模様をリアルタイムに動画中継したり (ニコニコ生放送と Ustream を予定)，Twitter の特定ハッシュタグの内容をテキストチャットへ中継したりすることを計画しています．それに，WISS Challenge で公募した議論を活発にする様々な工夫が加わります．また，参加者間の議論を活発にするために，名札を大きくして自己紹介・研究紹介等を記載する「名札改革」を継続し，名札がポジションペーパー的な役割を果たせるようにします．

最後に，登壇論文を投稿いただいた皆様，デモ／ポスターを発表いただく皆様，議論に参加し会議の成功に貢献してくださる参加者の皆様，スポンサー企業の皆様，協賛学会関係者の皆様，査読者の皆様，そして，WISS 2013 プログラム委員／運営委員の皆様に深く感謝いたします．

WISS2013 プログラム委員長

五十嵐 健夫 (東大)

WISS2013 運営委員長

平井 重行（京産大）

WISS2013 運営委員

会計	伊藤 貴之 (お茶大)
出版	渡邊 恵太 (明治大)
電子投稿	三浦 元喜 (九州工業大)
広報	
・学会リエゾン	中野 倫靖 (産総研)
WISS Challenge	綾塚 祐二 (ISID)
	野嶋　琢也 (電通大)
チャット	西田 健志 (神戸大)
動画生中継	坂本 大介 (東大)
	馬場　哲晃 (首都大学東京)
デモ・ポスター	竹川 佳成 (はこだて未来大)
	福本 雅朗（MSRA）
	松村 耕平（はこだて未来大）
ネットワーク	中村 嘉志 (国士舘大)
	瀬川典久（岩手県立大）
Web	濱崎 雅弘 (産総研)
投票	

* Takeo Igarashi <takeo[at]acm.org>, 東京大学

・ローカルサーバ	濱崎 雅弘 (産総研)
懇親会	岡部 誠 (電通大)
	橋本 直 (明治大)
表彰	宮下 芳明 (明治大)
企業スポンサー	神場 知成 (NEC ビッグローブ)
デザインコンペ	渡邊 恵太 (明治大)
学生ボランティア	佐藤 俊樹 (電通大)
参加マニュアル	
・名札	脇田 建 (東工大)

ISS 研究会連絡先

日本ソフトウェア科学会
「インタラクティブシステムとソフトウェア研究会」
主査: 五十嵐 健夫
〒 113-0033
東京都文京区本郷 7-3-1
東京大学大学院情報理工学系研究科
Email: takeo[at]acm.org
本研究会に関する情報は http://www.wiss.org/
からアクセス可能です.

WISS2013 プログラム委員

綾塚 祐二（ISID）
伊藤 貴之（お茶大）
伊藤 雄一（阪大）
五十嵐 健夫（東大）
江渡 浩一郎（産総研）
岡部 誠（電通大）
筧 康明（慶應大）
梶本 裕之（電通大）
神場 知成（NEC ビッグローブ）
岸野 泰恵（NTT）
佐藤 俊樹（電通大）
坂本 大介（東大）
瀬川 典久（岩手県立大）
高田 哲司（電通大）
竹川 佳成（はこだて未来大）
塚本 昌彦（神戸大）
豊田 正史（東京大）
中野 倫靖（産総研）
中村 嘉志（国士舘大）
西田 健志（神戸大）
野嶋 琢也（電通大）
橋本 直（明治大）
馬場 哲晃（首都大学東京）
濱崎 雅弘 (産総研)
原田 康徳（NTT）
平井 重行（京産大）
土方 嘉徳（大阪大）
福本 雅朗（MSRA）
増井 俊之（慶應大）
松村 耕平（はこだて未来大）
三浦 元喜（九州工業大）
宮下 芳明（明治大）
森島 繁生（早稲田大）
山本 岳洋（京大）
脇田 建（東工大）
渡邊 恵太（明治大）

表紙デザイン

中村 裕美 (明治大学)

もくじ

まえがき: WISS 2013 について
五十嵐 健夫 ..... *i*

セッション1 対話と共有

ネイティブ英語発話の日本人風の発音への変換による国際的な意識の促進
西田 健志 *1*
SmartVoice：言語の壁を越えたプレゼンテーションサポーティングシステム
李 翔, 暦本 純一 *7*
視覚障害者のための写真撮影・閲覧・共有インタフェース
浅川 智恵子, 原田 丞, 佐藤 大介, 石原 辰也, 髙木 啓伸 *13*
個人的な小さな幸せを実現するブラウザ上での動画編集・共有手法
中村 聡史, 石川 直樹, 渡邊 恵太 *19*

セッション2 読み書き

HirakuReader：行間を拡張する電子書籍
中村 将達, 西田 健志 *25*
SuprIME: IME によるテキスト編集機能の統合
中園 翔, 増井 俊之 *31*
CapStudio: プログラムの実行画面の録画映像を利用した統合開発環境
深堀 孔明, 坂本 大介, 加藤 淳, 五十嵐 健夫 *37*
順解析と逆解析を相互に利用する打楽器のインタラクティブなデザインインタフェース
山本 和彦, 五十嵐 健夫 *43*

セッション3 実世界

SensingFork：フォーク型センサを用いた食行動改善手法の研究
門村 亜珠沙, 李 爭原 , 陳 嚴章, 塚田 浩二, 朱 浩華, 椎尾 一郎 *49*
鍵盤上への演奏補助情報投影機能をもつピアノ学習支援システムにおける熟達化プロセスに関する調査
竹川 佳成, 椿本 弥生, 田柳 恵美子, 平田 圭二 *55*
LumoSpheres: 複数浮遊物体の実時間追跡にもとづく映像投影
山口 裕明, 小池 英樹 *61*

セッション4 動画

画像照合に基づく多メディア間の話題追跡システム
豊田 正史, 伊藤 正彦, Cai-Zhi Zhu, 佐藤 真一, 喜連川 優 *67*
静止画内物体への変形指示による動画検索
川手 裕太, 岡部 誠, 尾内 理紀夫, 平野 廣美 *73*
料理画像をアニメーションすることによる魅力的な料理動画生成システム
崎山 翔平, 岡部 誠, 尾内 理紀夫, 平野 廣美 *79*

セッション5 I/O

SyncBlink: 同期刺激によるナビゲーション方式の提案と評価
志摩 駿介, 綾塚 祐二, 寺田 努, 塚本 昌彦 *85*
Traxion:仮想力覚提示デバイス
暦本 純一 *91*
HandyScope：引き出しジェスチャを用いたテーブルトップ用遠隔地操作手法
栗原 拓郎, 三田 裕策, 大西 主紗, 志築 文太郎, 田中 二郎 *97*

セッション6 記憶

ライフログによる記憶拡張のための探索手法とその実践
中村 聡史 *103*
EpisoPass: エピソード記憶にもとづくパスワード管理
増井 俊之 *109*

デモ・ポスター発表

研究分野と参照関係からたどる文献可視化インタフェース
中澤 里奈, 伊藤 貴之, 斎藤 隆文 *115*
偶然の再会を演出する消極的な人向け人間関係継続支援システム
芥川 洋平, 山本 景子, 倉本 到, 辻野 嘉宏 *117*
ThermoGuide：表面温度分布を利用した加熱管理支援手法
斉藤 清隆, 高橋 伸, 田中 二郎 *119*
Harry: ペーパクラフトからの 3 次元モデル生成
三浦 元喜, 山本 将史 *121*
話者による正確なポインティングのできるプレゼンテーションシステム
若松 翔, 梅村 恭司, 岡部 正幸, 三輪 多恵子 *123*
「へやくる！」だらだらと賃貸物件情報をみることについて
大坪 五郎 *125*
ポータブルビジュアルプレイリスト GRAPE の主観評価
魚田 知美, 伊藤 貴之 *127*
肌マイクロスコープを用いた肌微細構造パラメータの診断
番場 文枝, 猪股 真美, 伊藤 貴之, 豊田 成人, 大高 瞳, 笹本 裕美 *129*
多様性と類似性を重視したアパレル商品閲覧システム
小池 恵里子, 伊藤 貴之 *131*
Touch-Shake:高齢者のコミュニケーションを促すインタラクティブシステムの開発と評価
山口 陽平, 柳 英克, 竹川 佳成 *133*
倍音分析によるいい声作りの支援アプリに向けて
菅原 衣織, 伊藤 貴之 *135*
Unisoner: 同一楽曲を歌った異なる歌声を重ね合わせる合唱制作支援インタフェース
都築 圭太, 中野 倫靖, 後藤 真孝, 山田 武志, 牧野 昭二 *137*
自由な筆記にロバストな手書き表認識システム
中洲 俊信, 柴田 智行, 井本 和範, 三原 功雄 *139*
部品へのプロジェクションマッピングによる通電しない電子工作
秋山 耀, 宮下 芳明 *141*

VisualMimic: スクリーンショットを利用した GUI 操作自動化のためのビジュアル開発環境
深堀 孔明, 坂本 大介, 五十嵐 健夫 *143*
災害対策本部における記録・文書管理システムの提案
仙道 航, 瀬川 典久, 澤本 潤, 杉野 栄二, 秋冨 慎司 *145*
誰でも神プレイできるシューティングゲーム。
簗瀬 洋平 *147*
叩打音を利用した操作インタフェースと浴槽への適用
伊藤 大毅, 平井 重行 *149*
住宅内の機器・サービスを統合するためのビジュアルプログラミング環境
橋岡 良, 平井 重行 *151*
1 つのセンサを用いたマルチタッチジェスチャセンサ
真鍋 宏幸 *153*
ライフログに基づく音楽推薦システム MALL とその検証実験
宇野 愛, 伊藤 貴之 *155*
Autonomous Interactive Desk Environment の構築
藤田 真康, 諏訪 博彦, 栗原 聡 *157*
BrightFace: プロジェクションによる表情拡張システム
辻田 眸, 暦本 純一 *159*
ボケる対話型システムの検討
藤本 浩介, 岡部 誠, 尾内 理紀夫 *161*
紙窓: そこに置くだけで操作可能なカードインタフェース
加藤 邦拓, 宮下 芳明 *163*
擬音語アニメーションによる動画音響の可視化手法
王 方舟, 柏野 邦夫, 永野 秀尚, 五十嵐 健夫 *165*
ダンスパフォーマンスにおける 2 軸動作可能な電動カーテンを用いた隊形練習支援
土田 修平, 寺田 努, 塚本 昌彦 *167*
モチベーションの維持を考慮したピアノ学習支援システム
福家 悠人, 竹川 佳成, 柳 英克 *169*
Bezel Check: ベゼル上におけるダブルクロッシングジェスチャ
黒澤 敏文, 志築 文太郎, 田中 二郎 *171*
スマートフォンと MAD-SS を用いた長距離センサーノードの開発
内田 泰広, 瀬川 典久, 澤本 潤, 杉野 栄二, 矢澤 正人, 後閑 政昭 *173*
暗記学習のための文節とモーラ数に基づいた替え歌自動生成システム
伊藤 悠真, 寺田 努, 塚本 昌彦 *175*
軟らかくてよく弾むセンサ内蔵ボール
佐藤 俊樹 *177*
Overview Scrollbar: ページ全体を概観として表示するスクロールバー
溝口 晃, 坂本 大介, 五十嵐 健夫 *179*
行動ログの入力を支援するユーザインタフェースの設計と実装
鈴木 孝宏, 美馬 義亮 *181*
ペイントアプリケーションにおけるタッチアクションを用いたエフェクトの提案
武藤 寛, 井川 裕貴, 瀬川 典久, 杉野 栄二, 澤本 潤 *183*
デジタルコンテンツ生成のためのクラウドソーシングによるパラメタ最適化
小山 裕己, 坂本 大介, 五十嵐 健夫 *185*
UnReC：文字入力に基づいた Undo/Redo
大江 龍人, 志築 文太郎, 田中 二郎 *187*

どこでも「いいね！」ボタン:位置評価情報の提示によるユーザ興味の変化に関する研究
佐々木 佳祐, 寺田 実 *189*
SmarTrackBall：スマートフォンのアウトカメラを用いたトラックボール
深津 佳智, 志築 文太郎, 田中 二郎 *191*
人間の表情を活用するライフログ閲覧
高井 啓, 志築 文太郎, 田中 二郎 *193*
音楽の構造情報を反映したミュージックビデオを簡単に制作・共有できる Web サービス
中嶋 誠, 五十嵐 健夫 *195*
実世界オブジェクトへの機能拡張によるインタフェースの提案
松嶌 信貴, 赤池 英夫, 角田 博保 *197*
手本データを自動抽出する運動指導システム
黒田 修平, 入江 英嗣, 吉永 努 *199*
EnchantMoon による漢字・部首認識を活用した描画補助ツールの提案
若林 花見, 井川 裕貴, 武藤 寛, 瀬川 典久, 杉野 栄二, 澤本 潤 *201*
講義中の板書を支援する人物半透明化カメラの実装
山根 恵和, 岡部 正幸, 梅村 恭司 *203*
浮遊ノートシステムにおける道案内機能の実装
小野澤 清人, 入江 英嗣, 吉永 努 *205*

著者索引

Cai-Zhi Zhu 67

赤池 英夫 197
秋冨 慎司 145
秋山 耀 141
芥川 洋平 117
浅川 智恵子 13
綾塚 祐二 85

五十嵐 健夫 37, 43, 143, 165, 179, 185, 195
井川 裕貴 183, 201
石川 直樹 19
石原 辰也 13
伊藤 貴之 115, 127, 129, 131, 135, 155
伊藤 正彦 67
伊藤 大毅 149
伊藤 悠真 175
猪股 真美 129
井本 和範 139
入江 英嗣 199, 205

魚田 知美 127
内田 泰広 173
宇野 愛 155
梅村 恭司 123, 203

王 方舟 165
大江 龍人 187
大高 瞳 129
大坪 五郎 125
大西 主紗 97
岡部 正幸 123, 203
岡部 誠 73, 79, 161
尾内 理紀夫 73, 79, 161
小野澤 清人 205

角田 博保 197
柏野 邦夫 165
加藤 淳 37
加藤 邦拓 163
門村 亜珠沙 49
川手 裕太 73

喜連川 優 67

倉本 到 117
栗原 聡 157
栗原 拓郎 97
黒澤 敏文 171
黒田 修平 199

小池 英樹 61
小池 恵里子 131
後閑 政昭 173
後藤 真孝 137
小山 裕己 185

斎藤 隆文 115
斉藤 清隆 119
坂本 大介 37, 143, 179, 185
崎山 翔平 79
佐々木 佳祐 189
笹本 裕美 129
佐藤 俊樹 177
佐藤 真一 67
佐藤 大介 13
澤本 潤 145, 173, 183, 201

椎尾 一郎 49
志築 文太郎 97, 171, 187, 191, 193
柴田 智行 139
志摩 駿介 85

菅原 衣織 135
杉野 栄二 145, 173, 183, 201
鈴木 孝宏 181
諏訪 博彦 157

瀬川 典久 145, 173, 183, 201
仙道 航 145

高井 啓 193
髙木 啓伸 13
高橋 伸 119
竹川 佳成 55, 133, 169
田中 二郎 97, 119, 171, 187, 191, 193
田柳 恵美子 55

朱 浩華 49
陳 嚴章 49

塚田 浩二 49
塚本 昌彦 85, 167, 175
辻田 眸 159
辻野 嘉宏 117
土田 修平 167
都築 圭太 137
椿本 弥生 55

寺田 実 189
寺田 努 85, 167, 175

豊田 成人 129
豊田 正史 67

中澤 里奈 115
中嶋 誠 195
中洲 俊信 139
中園 翔 31
永野 秀尚 165
中野 倫靖 137
中村 将達 25
中村 聡史 19, 103

西田 健志 1, 25

橋岡 良 151
原田 丞 13
番場 文枝 129

平井 重行 149, 151
平田 圭二 55
平野 廣美 73, 79

深津 佳智 191
深堀 孔明 37, 143
福家 悠人 169
藤田 真康 157
藤本 浩介 161

牧野 昭二 137
増井 俊之 31, 109
松嶌 信貴 197
真鍋 宏幸 153

三浦 元喜 121
溝口 晃 179
三田 裕策 97
三原 功雄 139
美馬 義亮 181
宮下 芳明 141, 163
三輪 多恵子 123

武藤 寛 183, 201

矢澤 正人 173
柳 英克 133, 169
簗瀬 洋平 147
山口 裕明 61
山口 陽平 133
山田 武志 137
山根 恵和 203
山本 景子 117
山本 将史 121
山本 和彦 43

吉永 努 199, 205

李 爭原 49
李 翔 7

暦本 純一 7, 91, 159

若林 花見 201
若松 翔 123
渡邊 恵太 19

ネイティブ英語発話の日本人風の発音への変換による国際的な意識の促進

西田 健志*

概要. 日本人は英語での読み書き能力の高さに比べて英語での会話を苦手にする傾向があると言われているが，その大きな障壁の一つにはリスニングやスピーキングにおける発音の問題が挙げられる．相手が話していることが聞き取れないことが会話を練習する大きな障害になるのはもちろんのこと，後天的にネイティブと同等の発音が身に付くことが稀であるにも関わらずネイティブの発音を模倣する練習を繰り返させることが自分の発音への自信を失わせ，英語で話すことを恥ずかしいと感じさせてしまっている面がある．一方で，英語が国際的な場でのコミュニケーションのための言語として位置づけられ，世界中で様々な訛りを持つ英語を話す人が増える中で，自分たちと同じような発音しか受け入れず，訛った英語を話す人を低く見がちなネイティブの態度にも改められるべきところがあると考える．そこで我々は，ネイティブ英語発話を日本人風の親しみやすい発音に変換することで，日本人の英語に対する心理的な障壁を緩和するとともに，ネイティブに対しては自分の発話よりも訛った発話の方が日本人には通じやすいという体験を通じて意識の改革を促す手法を提案する．本稿では，国際コミュニケーションを支援する様々な他の方法との関係性を議論するとともに，音声認識と TTS を組み合わせて発音の変換を実装した実験用アプリケーションとそれを利用してもらったときの日・米・中の出身者の感想を報告する．

1 はじめに

高校・大学入試での合格を最優先の目標とする英語教育の結果，文法理解や読解・英作文といった読み書きの能力に比べて，日本人は英会話を苦手にしていると言われている．その反省を踏まえ，ネイティブの英語教師の採用増や小学校からの英会話授業の導入など，会話能力の獲得を中心とした英語教育の改善が議論されている．

英語での会話能力の獲得にあたって特に障害になりやすいのはリスニングや発話における発音の問題である．まず，語彙や文法の知識があったとしても相手が話していることを聞き取ることができなければ，自分が言いたいことを英語で話す練習をするスタートラインにつくことすら困難である．また，発音についてはネイティブの発音を模倣する練習を重ねるのが通例となっているが，後天的にネイティブと同等の発音に達するのは例外的な少数の者だけであり，そのような練習を当然のように課すのは自信を失わせて学習意欲を削ぐばかりなので改められるべきだという声もある [2].

英語をあくまで国際的なコミュニケーションの道具として捉える「国際語としての英語 (English as an International Language)」の動きやインド英語やシンガポール英語など世界各地でそれぞれの訛りを持った「世界英語 (World English)」を話す人口が増えている現状を鑑みるに，英米などのネイティブからも，相手が聞き取りやすいように話す，相手の訛った英語を聞き取れるように努力するといった歩み寄りの精神が求められるのは当然のことと思われる．

発音に寛容で歩み寄りを心がける人は少なくないが，ネイティブらしくない英語に対する忌避的な反応にも根深いものがある．たとえば，ネイティブ発話と同程度に聞き取れる発話であっても訛りがあると聞き取りにくいと評価される傾向 [1] や，訛った英語で話す人が様々な差別を受けていること [6] が知られている．非ネイティブも発音の悪さを認識することだけはネイティブ並みにでき，訛った英語やその話者をネイティブと同様に低く評価する傾向があるが，自分と近い発音については比較的寛容であることが知られている [9].

本研究では，日本人英語話者とネイティブ英語話者の意識ギャップを埋めるためにネイティブの英語発話を日本人風の訛った英語発話に変換するシステムを利用して，英会話の練習および実践を行うことを提案する．日本人にとっては日本訛りの英語が聞き取りやすい [2] のでリスニングが楽になる分，言いたいことを英語で言うことにより集中することができる可能性がある．また，ネイティブにとっては，訛った発話の聞き取り練習に利用できるだけでなく，自分の元々の発音よりも訛った英語の方が日本人相手に通じやすいという，先入観が覆される体験を通じて，様々な発音の世界英語に対する寛容性を高めることが期待される．

本稿ではまず，提案手法の応用可能性について議論し，続いて様々な実験ができるように実装したアプリケーションについて，利用した日本人，英語ネ

Copyright is held by the author(s).

* 神戸大学大学院国際文化学研究科

イティブ，および中国人の声とともに報告する．さらに広範にわたる関連研究を提案手法との関係性・利害得失・応用場面の違いの観点から整理し，最後に今後の課題をまとめる．

2 考えられる応用例

本稿で提案するネイティブの発話を日本人風の訛った発話に変換する手法には，日本人には聞き取りやすくなることによる効果およびネイティブには聞き取りにくくなることによる効果を活用した様々な応用が考えられる．

2.1 ネイティブと会話するための補助

まず，提案手法はリスニングがあまり得意ではない日本人が英語ネイティブと会話する際の補助として利用することができる．発音が日本人風に変換されるのを聞くことで，訛った発音で話す気恥ずかしさも軽減されるだろう．具体的には，音声を翻訳するスマートフォンアプリのようにボタンを押してから話すと話し終わった後に変換される利用形態や，補聴器のように自動的に変換を行う利用形態などが想定される．後者のような利用形態を実現するためには発話部分を精度よく自動的に切り出しリアルタイムに変換できる必要がある．

自転車の補助輪のように，いずれは補助が無くても会話ができるようになりたい人がまずはネイティブと会話できる楽しさを味わいたいという場合にも利用できるだろう．そのような利用のためには，変換の程度を段階的に減らしていくことができる機能など，補助からの卒業しやすさも考慮することが望ましいと思われる．

2.2 ネイティブに対するカルチャーショック体験

様々な訛りを持った「世界英語」を話す人口が増え，TOEIC などの英語能力試験でも様々な発音が出題されるようになる中にあっても訛った発音で話す人を差別してしまうような人に対して，「発音で人を判断するのは良くない」，「訛った英語を聞き取れるように努力すべきだ」などと言葉で意識改革を求めるのには限界がある．

海外へ旅行や留学に行った際に言葉が相手に通じないという体験をすることは国際的な意識改革を促しやすいと思われるが，英語が国際語としての地位を確立しつつあることが影響して，英語ネイティブの中にはそのようなときにも「英語も通じないのか」と責任を相手に求めてしまう人もあるだろう．

英語ネイティブが日本人と話すときに，最初は自分の発音そのままで発話し，その後それを日本人風の発音に変換した音声をお互いに聞こえるように再生することで「元々の発音よりも訛った発音の方が日本人相手には通じやすかった」という体験ができるものと思われる．この体験を通じて，ネイティブの側からも相手に歩み寄るべき余地があることが実感され，様々な発音の世界英語に対する姿勢が改められることが期待される．

もちろん，提案手法はそのような意識改革がなされた後に訛った英語の聞き取りを練習するためにも利用することができる．ネイティブが非ネイティブの発話を聞き取れるようになるという目標は，非ネイティブにネイティブの発音を期待するよりは現実味があると言えるだろう．

2.3 会議での参加格差の是正

リスニング能力が比較的高い日本人にとっても提案手法が有益と思われる状況としてはネイティブとの会議が挙げられる．たとえばオンラインの音声会議においてネイティブの発言を日本人風の発音に変換するとネイティブだけが聞き取りに苦労するようになるため，日本人が議論のペースについていきやすくなり，発言もしやすくなる可能性がある．ただし，このような応用を実現するためにはネイティブからの歩み寄りや平等な参加に対する大きな配慮が必要となるので，まずは日本人が聞く音声だけを日本人風に変換するというセットアップも考えられる．また，逆に日本人の発話をネイティブ風に変換する手法を組み合わせることで格差をさらに小さくすることも期待できる．これらの応用を実現するためには発音の変換がリアルタイムに行えることが望ましい．

3 実験用アプリケーション

本章では，提案手法の効果を検証する様々な実験を行えるように実装した実験用のアプリケーションについて述べる．

現時点では一区切り話した後に変換した音声を再生するようになっており，2 章で議論した応用例のうち，2.1 節で述べた会話中の補助および 2.2 節で述べたカルチャーショック体験など少人数での利用に対応しているが，2.3 節で述べた会議などの多人数での応用には対応していない．

3.1 発音の変換方法

本アプリケーションでは音声認識プログラムとテキスト読み上げプログラム (text-to-speech; TTS) を用いることで日本人風の発音への変換を行う二つの方法 (図 1，2) を実装した．どちらの方法を用いても普通のノート PC を用いた場合，話し終えた直後に変換後の音声が再生される．音声認識には Windows に英語の言語パックをインストールすることで利用できるもの，TTS には Microsoft Speech Platform に含まれるものをそれぞれ利用した．

一つ目の方法は，ネイティブの発話を音声認識して得られた英語テキストを日本語 TTS に読み上げさせることで日本人風の発音に変換する方法である (図 1)．この方法は実装が非常に容易であるが，上

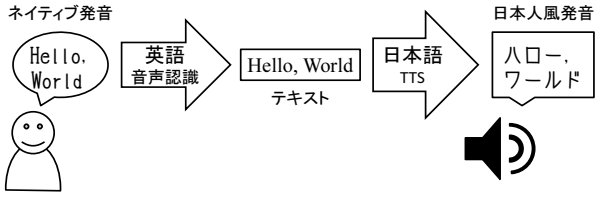


図 1. 変換方法 1: 英語音声認識の出力する英語テキストを日本語 TTS に読み上げさせる方法

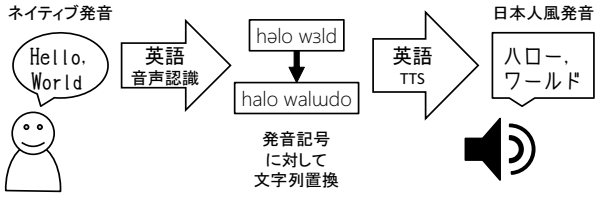


図 2. 変換方法 2: 音声認識の出力する発音記号を文字列置換で変換したものを読み上げさせる方法

述の日本語 TTS を利用することで日本人が各単語をバラバラにカタカナ読みしているような発音に変換することができる．またこの方法では，他言語の TTS を用いることで日本人風以外の様々な訛りの発音に変換することができると思われる．

しかしこの方法は英語を日本語 TTS で読み上げる段階で利用している辞書の影響か冠詞の ”a” を「エイ」，”who” を「ダブリューエイチオー」と読むなど，望ましくない変換をしてしまう場合がある．音声認識でも認識率向上のために辞書が利用されるため同様の問題が発生しうる．

そこで，二つ目の方法として，音声認識プログラムの出力する発音記号の文字列を文字列置換 (表 1) によって変換し，変換した発音記号を TTS に指定して読み上げさせる方法を用意した (図 2)．この方法では TTS の辞書に関わらず発音を変換することが可能である．また，音声認識の辞書にない単語や誤認識された語についても発音レベルで大きく外すことは少ないため，一つ目の方法よりは安定して発音を変換することができると思われる．

こちらの方法のもう 1 つのメリットとしては，文字列置換のルールを変えることで変換の種類や程度を調節できることが挙げられる．また，日本語 TTS ではなく英語など他の TTS を利用しても日本人風の発音に変換することができるので，英語ネイティブが日本人風の発音で話しているような音声などを得ることができる．その代わりに日本人風以外の発音に変換したい場合には，それぞれの訛りごとに変換ルールを用意する必要がある．

どちらの方法も TTS を利用しているので話者とは違う声になってしまう欠点がある．

表 1. 発音記号の変換ルールの例．発音記号の表記には国際発音記号 (IPA) を使用．

母音

変換前	変換後	例
[ɑʌɒ]	a	s<u>u</u>n
ɪ	iː	f<u>i</u>t
[uʊ]	ɯ	f<u>oo</u>t
ɛ	e	b<u>e</u>d
ɔ	o	d<u>o</u>g

子音

変換前	変換後	例
f	ɸ	<u>f</u>at
v	b	<u>v</u>est
θ	s	<u>th</u>ink
ð	z	<u>th</u>e
ʧ	ʦ	<u>ch</u>eep
r	l	<u>r</u>ight
ʍ	w	<u>wh</u>en

子音＋母音

変換前	変換後	例
ɪr, ɛr, ʊr	ɪa, ɛa, ʊa	<u>ear</u>
s[iɪ]	ʃi	<u>se</u>e
t[iɪ]	ʧi	<u>t</u>ip
t[uɯʊ]	ʦɯ	<u>to</u>o

子音の連続と末子音
（母音を後ろに補う）

子音	補う母音	例
dt	o	ves<u>t</u>
bfgklmpsvz	ɯ	ves<u>t</u>

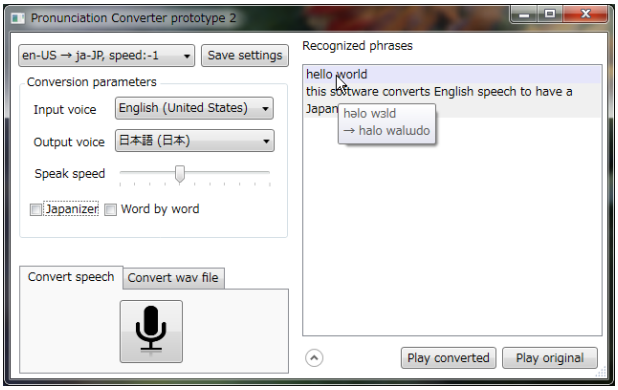


図 3. 実験用アプリケーションのスクリーンショット

3.2 実験用 UI

図 3 は，前節の変換方法を用いて様々な形態の実験ができるように作成した Windows アプリケーションのスクリーンショットである．ウィンドウには，変換パラメータの設定 UI（左上），マイク入力か音声ファイルを指定する UI（左下），音声認識結果の履歴（右）がそれぞれ表示される．

3.2.1 変換パラメータの設定

設定できる変換パラメータには，使用する音声認識と TTS のエンジン，発音変換方法，そして話速度がある．

音声認識と TTS のエンジン 使用する音声認識エンジンを変更することで英語以外の言語を日本人風に変換する実験，TTS エンジンを変更することで様々な発音に変換する実験を行うことができる．

発音変換方法 チェックボックス ”Japanizer” にチェックを入れると前章の二つの発音変換方法のうち，発音記号の文字列置換を行う方法が使用される．

話速度 話速度の影響を調べることができるよう，全体の話速度を調節するスライダーと単語間を長くするチェックボックス ”Word by word” の二種類の設定を用意した．

3.2.2 入力と出力

本アプリケーションはマイク入力によるその場での変換と録音済み wav ファイルの変換に対応している．マイク入力の場合は，マイクボタンを押してから発話が途切れるまでを 1 単位として認識し，それを変換した結果が直後に再生される．ファイル入力の場合には全体がまとめて認識，再生される．どちらの方法を利用した場合にも，変換前後の音声と認識結果をファイルに保存することができる．

3.2.3 音声認識結果の履歴

右側には音声認識結果の履歴が表示されており，カーソルを各結果の上に持って行くことで変換前後の発音記号を確認することができ，クリックすることで発音変換後の音声を再び聞くことができるようになっている．これにより，変換パラメータを変更しながら様々な発音を聞き比べる実験や，音声認識結果や発音記号を相手に見せながら会話するなど他の支援方法と比較する実験が可能である．

3.3 利用者の反応

本節では，実験用アプリケーションを日本人大学生 4 人，日本の大学で英会話の授業を担当しているネイティブ英語教員 1 人，日本の大学への中国人留学生 3 人に簡単に利用してもらった反応をまとめる．

発音変換の質 日本語 TTS をそのまま利用する一つ目の方法で変換した音声については紛れもなく日本人の話す英語らしく聞こえるが，発音記号を変換する二つ目の方法で変換した音声については，日本人らしく聞こえるものの一つ目の方法と比べるとやや聞き取りづらいことがあるとの反応で一致した．

応用可能性 ネイティブ英語教員から，自身の担当している英会話の授業で利用してみたいとの意見が得られた．自身の経験上，日本人訛りの英語が聞き取りづらいと感じるのは最初だけで，少し慣れれば聞き取りやすい部類に入ると思っているが，その割に英会話の授業などで恥ずかしがって話したがらない学生が多いので，そのようなときにこれを利用すると話しやすくなりそうだとの意見をいただいた．

他の訛りへの変換 中国語 TTS を利用して変換した英語音声を聞いてもらったところ，ネイティブ英語教員と中国人留学生のうち 1 人が，中国人の話す英語らしく聞こえると答えたが，日本人風の発音よりは微妙であったためか具体的にそのように感じたポイントを指摘した．その同じ音声について日本人大学生からは，かなりネイティブに近く感じられるとの一致した感想が得られた．

4 関連研究

外国語でのコミュニケーションを支援する方法の研究としては，機械翻訳や通訳者を利用する方法，音声の変換を利用する方法に加え，その他様々なインタラクティブシステムによる支援方法がある．本章では提案手法との関係性，利害得失，および応用場面の違いについて議論する．

4.1 機械翻訳を利用する方法

互いに使い慣れた言語を共有しない者どうしが意思疎通を図るときにもっとも手軽に利用できる方法のひとつが機械翻訳である．機械翻訳機能を搭載したテキストチャットに関する研究 [12, 11] に加え，最近では Google 翻訳 [5] や docomo はなして翻訳 [3] といった音声を入力すると翻訳後の音声を出力するスマートフォンアプリも登場している．しかし，機械翻訳の精度が向上しているといってもしばしば誤った翻訳をするという問題があり，特にチャットでの砕けた表現や話し言葉に対応することは現状でも容易ではない．

翻訳精度が高い場合にも，ある言語から別の言語へと翻訳された表現とそれを元の言語へ再翻訳した表現とが一致しないせいで意思疎通が難しくなることが指摘されている [12]．たとえば，相手の使った表現を引用して返事をしたときに元々相手が使っていた表現とは異なる表現に訳されたり，一度お互いに通じると確認しあった表現であっても後にそれを短縮した表現を使ったときに異なる表現に訳されたりするせいで，コミュニケーションの基礎となる表現が安定しないことが問題となる．

機械翻訳を用いる際の問題点を克服するために機械翻訳の使い方を工夫する試みもなされてきた．まず，再翻訳した表現 (back-translation) を表示しながらコミュニケーションを行うとコミュニケーションが円滑になることが実験により明らかにされている [11]．五十嵐は，機械翻訳が失敗しがちな長く複雑な係り受けがある文章を翻訳する際に，文節の区切りや係り受け関係を専用の UI を使って入力し，区切られた文節を別々に機械翻訳して視覚化することで伝わりやすくする手法を提案している [13]．

これらの研究に対して本研究は，翻訳を介するのではない新たな歩み寄りの形を提案するものである．機械翻訳を利用すると相手の話す言語についての知

識がほとんどなくてもコミュニケーションを取ることができるので便利であるが，いずれは補助を卒業したいと考えている人にとっては提案手法の方が向いているだろう．さらに，提案手法には，ネイティブにカルチャーショック体験を与えて意識改革を促すという機械翻訳では実現できない応用がある．

技術的には，機械翻訳に付き物の翻訳ミスに比べると，発音変換の方が想定外の動作による影響を小さくできる可能性がある．特に音声入力の結果を翻訳する場合には，音声入力での認識エラーと翻訳エラーの２種類のエラーが重なって発生することになるので両者の認識精度が極めて高い必要がある．

4.2 人間の通訳を利用する方法

機械翻訳ではなく人間の通訳を利用することで，くだけた表現やその場面の文脈などへのより柔軟な対応が可能になるが，専門の通訳を用意するには高い人的コストがかかり，その場にいる人が対応するにしても高い負担がかかってしまうという問題がある．通訳にかかる負担への対処に関する研究としては，多言語対面会議において少数派言語の参加者を他の参加者が協同で支援することで一人一人の支援負担を軽減する all-for-one 型支援システムが挙げられる [14, 15, 16]．

それに対して本研究で提案するのは，話しているネイティブ自身による歩み寄りや配慮に期待する方法である．

4.3 発音の変換を利用する方法

発話の音響的な性質を変換する研究分野で主流となっているのは，声質，つまり誰の声と認識されるかを変換するものであり，それと比べて発話の発音を変換する研究は多くはない [4]．またそのいずれも，訛りを取り除く，適切な抑揚を付加するなどして，ネイティブらしい発話に近づけることを目的としたものである [4]．

訛りを取り除く方法はネイティブに話が通じやすくなるので便利であるが，時代の趨勢に逆行して「美しい発音」への憧れをさらに強くしてしまう恐れがある．それを防ぐためにはあくまで「相手に一番通じやすい発音」への変換を共通の技術目標として捉え，ネイティブ風への変換と日本人風への変換の両方の研究を進めていくことが適切であろう．

もちろん，ネイティブと非ネイティブがお互いに逆方向の変換を使うという形で組み合わせることもでき，そうした方が非ネイティブの訛りを除去するだけの場合よりもより公平な環境になると考えられる．

技術的には，ネイティブらしい発音の英語に対する音声認識の精度がかなり高くなっているのに対して，非ネイティブの発話を認識するためには様々な訛りの認識に対応するのが難しい [8] のと同様に，発音の変換についてもネイティブ発音を認識して日本人風に変換する方が容易である可能性が高いと思われる．

4.4 その他のインタラクティブシステムによる支援

単に発話速度が遅くなるように変換したり，音声認識結果を画面に表示するだけでも，リスニングが苦手な人にとっては大きな支援になるものと考えられる．音声訂正 [7] のように音声認識結果を手作業で修正する手法を組み合わせれば，4.1 節で述べた認識エラーと翻訳エラーが重なる問題も大きく解消されうる．

しかし，これらの方法はどれも利用中の「自分たちは英会話ができない」というネガティブな印象が前面に出てしまいがちなのに対して，提案手法を利用することで「自分たちは訛っているけど英会話はできる」というポジティブさを前面に出すことができるのではないかと考えている．

多人数環境における非ネイティブの支援に着目した研究としては，少人数状況では英語で会話できる非ネイティブであっても会議では議論の速さについていくことが難しく発言が少なくなりがちであることを改善しようとした研究がある [10]．その研究では，オンライン音声会議でネイティブ同士だけ音声伝達を 0.2-0.4 秒程度遅らせることで非ネイティブに時間的な余裕を与える手法が提案・評価され，確かに非ネイティブが議論に参加しやすくなる効果は認められるものの，発話の開始が重なりやすくなって進行が妨げられるという改善すべき点も確認されている．本研究もこの研究と同様に，ネイティブの側から非ネイティブに配慮することを促すことを重視しているが，より意識されやすい形を採ることで歩み寄りへの意識改革を促すことも狙っている．

5 まとめと今後の課題

本稿では英語ネイティブの発話を日本人風の発音に変換して日本人にとって聞き取りやすくする手法を提案し，日本人のリスニング支援や英語ネイティブへのカルチャーショック体験などの応用可能性を議論した．さらに音声認識と TTS を利用して実装した実験用アプリケーションを異なる母国語を持つ人たちに利用してもらい，初期的なフィードバックを得た．

今後の課題としてはまず，日本人風変換を利用した英語発話の聞き取りおよび会話に対する効果について，より多くのネイティブと日本人が参加する実験により検証することが挙げられる．また，さらなる安定性，話者の声質保持，およびリアルタイムな変換などを実現するため，音声認識や TTS を利用するのではなく，音声信号を直接処理することによって発音を変換する手法を模索する．それらの成果を基にして最終的には，ネイティブに対するカルチャーショック体験や会議での格差是正など，他の

応用例についてのシステムを開発し，実証実験を行う計画である．

参考文献

[1] T. M. Derwing and M. J. Munro. Accent, intelligibility, and comprehensibility: evidence from four L1s. *Studies in second language acquisition*, 19(01):1–16, 1997.

[2] T. M. Derwing and M. J. Munro. Second Language Accent and Pronunciation Teaching: A Research-Based Approach. *TESOL Quarterly*, 39(3):379–397, 2005.

[3] N. docomo. はなして翻訳. `http://www.nttdocomo.co.jp/service/communication/hanashite_honyaku`.

[4] D. Felps, H. Bortfeld, and R. Gutierrez-Osuna. Foreign accent conversion in computer assisted pronunciation training. *Speech Communication*, 51(10):920–932, 2009.

[5] Google. Google Translate. `https://play.google.com/store/apps/details?id=com.google.android.apps.translate`.

[6] M. J. Munro. A Primer on Accent Discrimination in the Canadian Context. *TESL Canada Journal*, 20(2):38–51, 2003.

[7] J. Ogata and M. Goto. Speech Repair: Quick Error Correction Just by Using Selection Operation for Speech Input Interface. In *Proc. Eurospeech '05*, pp. 133–136, 2005.

[8] Y. Tsubota, T. Kawahara, and M. Dantsuji. Recognition and verification of English by Japanese students for computer-assisted language learning system. In *INTERSPEECH*, 2002.

[9] P. Winke, S. Gass, and C. Myford. The Relationship Between Raters' Prior Language Study and the Evaluation of Foreign Language Speech Samples. Technical report, Educational Testing Service, 2011.

[10] N. Yamashita, A. Echenique, T. Ishida, and A. Hautasaari. Lost in transmittance: how transmission lag enhances and deteriorates multilingual collaboration. In *Proc. CSCW '13*, pp. 923–934. ACM, 2013.

[11] N. Yamashita, R. Inaba, H. Kuzuoka, and T. Ishida. Difficulties in establishing common ground in multiparty groups using machine translation. In *Proc. CHI '09*, pp. 679–688. ACM, 2009.

[12] N. Yamashita and T. Ishida. Effects of machine translation on collaborative work. In *Proc. CSCW '06*, pp. 515–524. ACM, 2006.

[13] 五十嵐 健夫. 文の構造を明示的に指定・表示することによる異言語間コミュニケーション. WISS 2011 予稿集.

[14] 吉野 孝, 井出 美奈. All for one 型多言語会議支援システムの構築と評価. 情報処理学会論文誌, 51(1):36–44, 2010.

[15] 宮部 真衣, 吉野 孝. All for one 型対面会議支援システムのためのワークスペースアウェアネスの効果. 電子情報通信学会論文誌. D, 情報・システム, 94(1):27–36, 2011.

[16] 宮部 真衣, 吉野 孝. 多言語対面会議支援システムのための All for one 型支援の効果. 情報処理学会論文誌, 52(1):90–96, 2011.

未来ビジョン

我々日本の研究者も英語には散々苦い思いをさせられている．論文を読むにも書くにも何倍も時間をかけ，高い金を払って英文添削を受ける．国内の学会で発表された研究に類似した研究を国際学会の査読で見かけても指摘することができない．「英語で発表しなければ意味がない」と発破をかける声が日本人の中からも常に出てくるが，片方に意思疎通の努力責任をすべて押し付ける状況にこそ疑うべき点があるのではなかろうか．

奇しくもインタラクティブシステムの分野では，CHI などの国際会議で文化の相違 (cultural difference) を標榜するセッションが必ずあり，不十分・不安定なインフラや低識字率など様々な問題を抱えた非先進地域に焦点を当てた研究が目立ち始めるなど，他の文化を理解しようという風潮に注目が集まっている．

今こそ，まずは研究者コミュニティの中から，非先進地域も結構だがもっと身近な問題にも目を向けるべきだということ，そして言葉の問題に限らずコミュニケーションでは双方が互いに歩み寄るべきだということを示していくべきときである．

理解さえ得ることができればすぐにでも実現できるアイデアは多い．

- 日本の学会で発表された優れた研究を紹介するセッションを国際学会に設ける
- 採択決定後に一括で英文添削する
- 日本語で投稿と査読を行って，採択が決定してから英訳することを認める

そうしたことを実現していくためには，山下らの研究 [10] や本稿のように英語ネイティブからの歩み寄りを求める研究を積極的に発表し続けることが役に立つと期待している．

SmartVoice: 言語の壁を超えたプレゼンテーションサポーティングシステム

李 翔　　暦本 純一*

概要.　国際的イベントでは, 共通言語での講演が求められることが多い. しかし, 語学力を短期間に向上させることが難しく, 共通言語が母国語ではない講演者にとって負担になる. また, 同時通訳などの方法は, コストが高い上, 講演者本人の発表を聞いているわけではないため, プレゼンテーションの自然さを損なってしまう. 本論文では, 言語の壁を越えてプレゼンテーションを可能にするシステム「SmartVoice」を提案する. SmartVoice は, 音声データ化した原稿を講演者の口の動きにあわせてスマートに送り出し, 講演者が自分のタイミングで講演できるようにする. 音声のイントネーションや強弱も, 口の位置と形状に基づいて制御するため, まるで講演者本人が直接喋っているように見える. 評価実験を行ない, 本人のプレゼンテーションと SmartVoice によるプレゼンテーションがほとんど区別できないことがわかった. SmartVoice のアプリケーションは, 多言語を用いるプレゼンテーションに限らず, 映像に音声・セリフを入れるアフターレコーディングにも適用可能である. 本論文では, 他分野における SmartVoice の応用についても議論する.

1　はじめに

国際学会などのイベントでは, 共通言語, たとえば英語で講演することが求められる. しかし, 共通言語が母国語ではない人にとって, それは大きな障壁となる. 仮に原稿を書いて読み上げることにしたとしても, 正しい発音で流暢に聞こえるためには多くの努力を要する.

言語が異なる講演のために, 同時通訳が利用されることがある. しかし同時通訳者を雇うコストや講演者とは異なるタイミングで音声が流れる不自然さなどの問題がある. また講演者の感情や非言語情報を伝達することが困難である. 予めプレゼンテーション原稿の音読を録音し, プレゼンテーションの時にそれを再生して口パクする方法もあり得るが, 講演者は録音のペースに合わせる講演をすることになり, 講演の自由さが大きく制約されてしまう.

ボイスシンセサイザーを用いることで, 原稿のテキストファイルさえあれば, コンピュータがどの言語の単語でも正確な発音で読み上げてくれる. しかし, 文章を朗読する時のイントネーション, 抑揚が単調であり, 聞き疲れするため, 人間の声の代用とすることはできない. 例えば通常の講演では, 大事なところで間を開けて聴衆の注意をひきつける, などが行われるが, ボイスシンセサイザーではこのような演出をすることは難しい.

本論文では, これらの問題を解決する手段として, 「スマートな口パク」に基づくシステムを提案する. スマートな口パクとは, まず原稿をナレーターに読んでもらうか, ボイスシンセサイザーを利用するかして音声データにしておき, それを講演者の口の動きに合わせて送り出す方式である. 講演者は自分のタイミングでプレゼンテーションすることができ, 大事なところで間を空けるなど, 語速を自由に制御できるようになる. また, イントネーションや音の強弱などを, 口の位置や形状から制御するようにしている. この結果, 聴衆からすると, 講演者本人が直接喋っているように見え, 同時通訳のような不自然さがない. この方式での講演は, 共通言語でのプレゼンテーションに留まらず, 聴衆に合わせた言語での講演 (たとえば中国で中国語で講演する) にも利用出来る. この, スマートな口パクに基づくシステムを SmartVoice と呼ぶことにする.

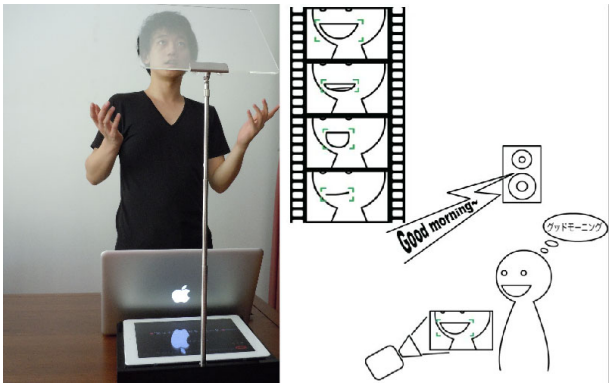


図 1. SmartVoice によるプレゼンテーション

2　SmartVoice

SmartVoice は, 再生制御およびピッチ, スピード, 音量のリアルタイム調整をプレゼンターの口の動きとその他表情にマッピングし, 音声が口パクのペース

Copyright is held by the author(s).

* Xiang Li, 東京大学大学院 学際情報学府 Jun Rekimoto, 東京大学大学院 学際情報学府, ソニーコンピュータサイエンス研究所

に合わせて自然に流すシステムである．言語の壁を越えたスピーチやプレゼンテーションのサポーティングを始めとするさまざまなアプリケーションが考えられる．SmartVoice を用いたプレゼンテーションの様子を，図 1 に示す．

2.1 SmartVoice を用いたプレゼンテーション

実際に SmartVoice を用いたプレゼンテーションの仕方について説明する．

まず，講演原稿を用意し，それに基づいた音声データを準備する．音声データは，ユーザ自身もしくは自分以外のナレーターによって原稿を読み上げた録音か，ボイスシンセサイザーによって自動生成される音声を想定している．用意した原稿テキストファイル（.txt）および音声データをシステムに与えると，SmartVoice は半自動的に原稿と音声のフレーズレベルの対応付けを行う．そして音声と正しく対応付けられる原稿中のフレーズは，ハーフミラープロンプターに表示される．

利用者は，用意した音声データをシステムに与え，顔がカメラ領域に収められ，トラッキングできる状態で，プロンプター画面を見て，原稿中のフレーズを口パクすれば，そのフレーズと対応する音声が流れてくる．一つのフレーズを読み終えて口を閉じると，再生が止まり，プロンプターには次に読むべきフレーズが強調される．再び口パクをすると，今読むべきフレーズと対応する音声が流れるが，それまでに持続時間任意の一時停止ができる．また，イントネーションも前述したように，口パクのペース，口の位置と眉毛の位置によってリアルタイムでコントロールできる．

2.2 SmartVoice におけるスマートな口パク

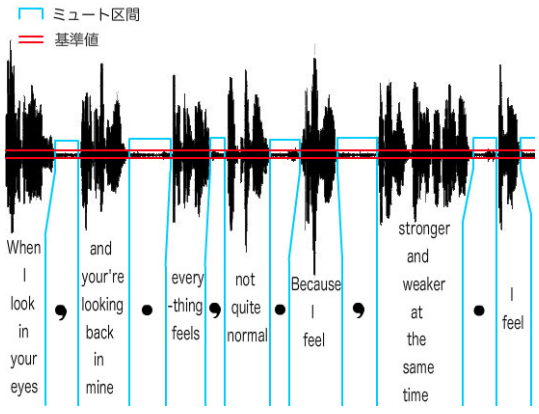


図 2. 音声と原稿からのフレーズ抽出と対応付け

2.2.1 音声の区切りおよび送り出し制御

実際に SmartVoice を使ってプレゼンテーションする時に，ユーザとインタラクションするためにプロンプターを利用している．ユーザがプレゼンター画面を見て，常に現在原稿のどこを読んでいるのかを確認することができる．そのために，音声とテキストは，正しく対応付けていることが必要である．SmartVoice は，この対応付けをシステムが半自動的に処理する．

通常の講演では，呼吸を入れることによって話のテンポを変えたり，一時停止することがある．なお，再生中の区切りに対し，違和感を感じさせないためには，最低限一つの単語の中での一時停止は避けるべきである．本システムでは，原稿テキストをカンマやピリオドなどのパンクチュエーションを区切りとしてフレーズを抽出し，音声データについては，フレーズは音声中に音響レベルが全音声データの平均値の一割以下で持続時間 0.1 秒以上の箇所（以下ミュート区間とする）によって分離された区間とする．一時停止は，ミュート区間内のみを許可する．ただし，特に録音の場合，ミュート区間中に，音響レベルが高いが，持続時間が極めて短いノイズは存在する．これらを見つけ出し，フレーズ抽出プロセスに無視されるよう，ノイズによって分断されるミュート区間が存在しないことを保証するためのアルゴリズムも用意した．(図 2)

原稿と音声の両方からフレーズが抽出されたら，本システムは後述するマッチングプロセスを用いて，原稿テキストと音声データの phrase-to-phrase の対応付けを行う．

プロンプターに表示される原稿テキストは，フレーズ単位で音声と対応し，ひとつのフレーズの音声を送り出したら，プロンプター画面には，自動的に次のフレーズのテキストに切り替わる．講演者はプロンプ他に表示されるフレーズを「口パク」で読み上げるか，概ねそのように口を動かすことで音声データが順次送り出されていく．

2.2.2 ピッチ，スピード，音量の制御

音声データは講演者の口の動きによって送り出されていくが，さらに，そのピッチ，スピード，音量も講演者の顔情報から制御できるようにした．例えば，ボイスシンセサイザー音声データを利用した場合でも，このようにして抑揚や感情表現を講演時に付与することができ，より自然な講演になる．

これらの制御は，人間が実際に喋っている時の原理を考えて，方針を決めた．具体的に，ピッチについて，頭を下げる時にピッチを下げ，逆の場合はピッチを上げるようにした (図 3 - b)．次にスピードについて，口の動きが激しければ激しいほど語速を高く設定する．最後に，音量について，目の開き具合によってコントロールしようとしたが，今回は眉毛の

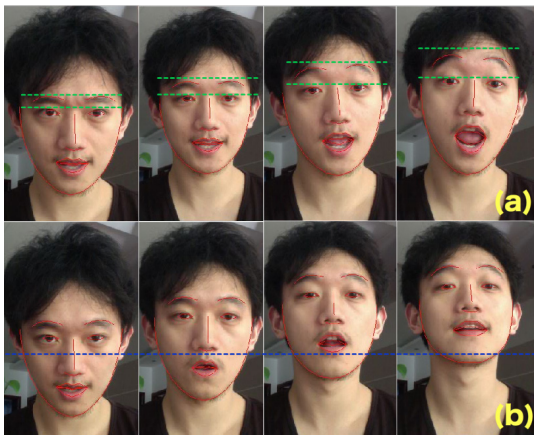


図 3. (a) 音量のコントロール (b) ピッチのコントロール

高さでコントロールすることにした（図 3 - a）. 理由は後述とする.

3 システム構成

本システムは, ホスト側とリモート側の 2 つの部分から構成されている. ホスト側は, ユーザの顔をトラッキングするためのカメラと, 音声を出すためのスピーカーが付いている諸処理を担当するノート PC からなり, リモート側は, プロンプターとして使われるタブレット PC とプロンプタースタンドからなる. なお, 両部分の間は, 無線 LAN 通信でコミュニケーションを取る.(図 4)

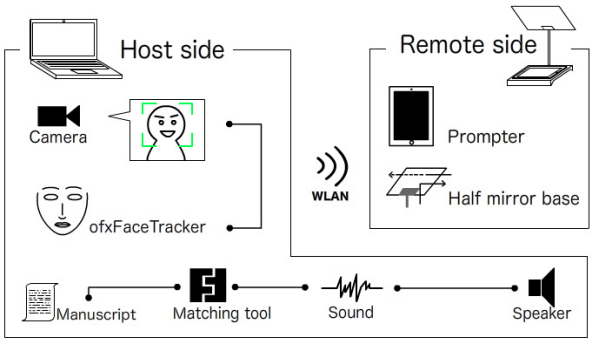


図 4. システム構成図

3.1 ホスト側

本システムは, プレゼンターに負担をかけず, より自然なパフォーマンスを実現するために, 使用中に全ての操作は, 口の動きや表情のみを使い, ハンズフリーに行う. プレゼンターは, 再生の一時停止と再開を決めることが可能のほか, 音声のスピード, ピッチそして音量をリアルタイムで変えることによって, ボイスシンセサイザーが生成した音声をより人間的にすることができる.

3.1.1 顔認識およびフェイストラッキング

SmartVoice は, 常時にプレゼンターの顔の位置, 口の動きおよび表情をトラッキングする. 本システムにおいて, openframeworks[1] 用の Add-on である ofxFaceTracker[2] を用いて, 人間の顔およびその傾きや大きさ, さらには目や鼻, 口, 眉といった顔のそれぞれのパーツの位置や大きさを立体的に検知してトラッキングする. 本システムのプロトタイプでは, Web カメラ付きノート PC を用いて, カメラから取得したリアルタイム映像からユーザの顔をトラッキングし, 画面にある顔および顔の各パーツの位置, 口の開閉パターンおよびそのペース, 眉毛の位置などを常時に記録し, 表示している.

3.1.2 Phrase-to-phrase の対応付け

前章で言及したように, 本システムにおいて, 原稿テキストと音声は, フレーズレベルで対応付けられている. この Phrase-to-phrase の対応付けの正しさを保証するために, 本システムの一部として半自動的マッチングツールを導入した.

マッチングツールはまず, 音声データ中のミュート区間を見つけ出し, フレーズを抽出し, 各フレーズの再生時間を配列にキャッシュする. 次に, 原稿テキストから抽出されたフレーズを, 順番にボイスシンセサイザーにかけて, 生成された一時的音声データの再生時間をそれぞれ取得して配列にキャッシュする.

ボイスシンセサイザーによって生成された音声の再生時間は, 語速の違いによって人間のナレーターによるものと著しく異なる場合があるため, マッチングツールは直接に前述の 2 つの再生時間の配列をマッチングするではなく, 音声データと原稿テキストのそれぞれ各フレーズの再生時間対総再生時間の比をマッチングの対象とする.

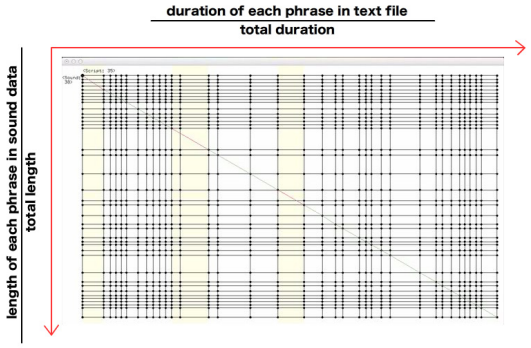


図 5. 自動マッチング結果を示すマッチングツールの画面

マッチングには DP マッチングアルゴリズムを利用している. 図 5 のように, 黄色に塗りつぶされた

部分は Miss match を表している．Miss match の原因は, 原稿と音声から抽出されたフレーズの総数が一致しないことだと考えられ, 主に 2 つのケースがある．一つは, 本来は区切りではないところで音声を止めてしまったため, ミュート区間とされた場合．もう一つは, ナレーターが原稿中の区切りを無視して続けて喋った場合である．いずれのケースに対しても, 本システムは手動修正をしやすくするための GUI をユーザに提供する．

3.1.3 再生コントロール, 感情表現付け

SmartVoice では, 一時停止の長さはユーザによって, 任意に決めることができる．ミュート区間中に一時停止するたびに, システムはそのミュート区間を丸ごと飛ばすことによって, ユーザが再び口を開くと, 音声がすぐ再生し, ディレイ無くプレゼンテーションを再開することが可能である．

前述の通り, ボイスシンセサイザーによるコンピュータボイスは, 自然なイントネーションに乏しく, 自然に喋っているように聞こえない．そこで, 本システムは, 再生時にリアルタイムで音声の音量変更, ピッチシフト, ピッチを変えずに再生スピードを変えることによって, プレゼンターの感情表現を付けることを可能にした．音声のダイナミックピッチシフトおよびストレッチは, C/C++ベースのマルチプラットフォームで動作するタイム・ピッチマニピュレーションライブラリ Dirac3 を用いて実現している．[3] ピッチシフトについては, 0.2 刻みでもとのピッチから上下 1 半音の変動幅に制限し, ストレッチについては, 0.7 倍速から 1.6 倍速の幅に制限している．音量について, 初期状態では 70 %にセットし, 再生時に 50 %から 100 %までの間で変化することができる．

3.1.4 マッピング

本システムは, プレゼンターをプレゼンテーションに専念させるために, すべての操作パラメータは, トラッキングされているプレゼンターの顔のパーツにマッピングしている．

再生の一時停止と再開は, ユーザの口の開閉にマッピングする．フレーズ再生後にプレゼンターが口を閉じ, 再び口を開くと, 次のフレーズの再生が開始される．また, 再生途中, つまりミュート区間以外の位置では, ユーザの口の開閉を問わず, 再生し続ける．これによって, 音声がスマートにプレゼンターの口の動きに合わせて流れてくるによって, まるで本人が喋っているように見えるという本システムのもっとも基本的な目的が実現される．

本システムでは, ピッチ, 音量およびスピードは, 再生時リアルタイムに変更を加えることができる．ピッチシフトについては, ユーザの顔位置にマッピングしている．具体的に, 再生時にユーザの顔の位置が初期状態顔の位置との縦方向の差で計算する．初期状態より顔を上げれば, ピッチが高くなり, 逆に, 頷くとピッチが低くなる（普通に喋る時に, 頷くと声帯が圧迫されるため）．音量の変更について, 前章で述べたように, 今回はユーザの目の開き具合にマッピングしている．しかし, 直接目の広き具合を取ると, 変化率が小さいため, 瞼と連動する眉毛の上下位置を採用した．ピッチシフトと似ており, 再生時にユーザの眉毛の位置が初期状態眉毛の位置との縦方向の差で計算し, 初期状態より眉毛を上げれば, 音量が高くなり, 逆の場合は音量が低くなる．スピード変更については, 再生時に口パクのペースにマッピングしている．具体的に, 今のフレームと直前のフレームにおいて口の面積の変化率を取って, 口の動きが激しければ, スピードを上げ, 逆にゆっくりと口を動かすと, スピードを落とす事が可能である．

3.2 プレゼンテーション時のインターフェース

3.2.1 プロンプター

本システムにおいて, プレゼンテーション時にユーザと実際にインタラクションをとるユーザインターフェースとして, 原稿を提示するためのプロンプターを設けている． なお, 本システムのプロトタイプ SmartVoice では, タブレット PC(iPad) をプ ロンプター画面とし, プロンプタースタンドとしてハーフミラーによるプロンプター台を利用している．ハーフミラー付きプロンプターを使うことによって, プレゼンターは講演会場を自然に見渡しながら原稿を読むことが可能となる．

3.2.2 ホスト側との協同作業

SmartVoice では, リモート側のプロンプターの稼働環境として iPad を採用し, リモート側とホスト側の間は, 無線 LAN 通信を用いてコミュニケーションを取っている．その理由は, なるべくプロンプターの機能をシンプルのままに保ち, プレゼンテーションできるまでの準備作業は, すべてホスト側で完結可能にしたいと考えているからである．そのために, プロンプターに表示すべき原稿は, ユーザが用意したテキストファイル (.txt) をホスト側プログラムで読み込み, 講演時に無線通信によりリモート側のプロンプター画面 (iPad) に伝送する．

実際にプレゼンテーションをする際に, ユーザとインタラクションをするのは, プロンプターであるため, ユーザは講演時にはプロンプター画面のみでシステムの挙動を確認できるようになっている．プロンプター画面は, 図 6 - c で示しているように, ホスト側でうまく顔を認識しているか否かを, ラベル「DETECTED!」でユーザが知ることができる (認識できてない場合は「???」と表示)．

音声と原稿テキストはフレーズレベルで対応付け

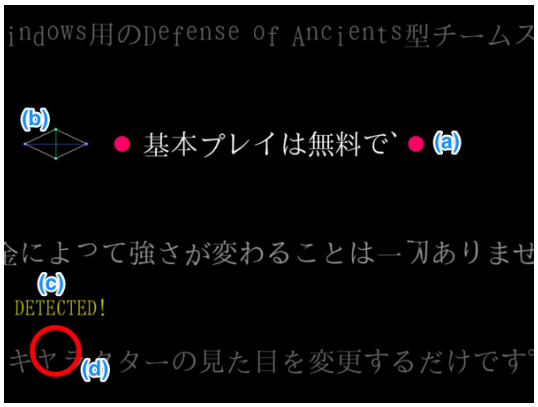


図 6. プロンプター画面

ているため, プロンプター画面にも, フレーズ単位で原稿をユーザに提示する. 具体的に, プレンゼンターが今読んでいるフレーズだけもっとも見えやすくし, 読み終わったフレーズはスライドアップするようにしている. (図 6 - a)

ピッチや音量の変化のようなすぐに分かる変化と異なり, 本システムを用いてナレーションの再生スピードを変える場合, 現在の再生速度の目安をユーザに提示する. プロンプター画面では, 口のイメージ (図 6 - d) が, 喋る速度 (ナレーションの再生速度) に合わせて, 開閉アニメーションの速度が変わる. 一時停止の時は閉じるままである. また前述のように, 本システムにおいてスピードを決めるのは口の面積の変化率であるため, そのメタファとして, 図 6 - b で示す 2 つのバーをプロンプター画面に追加した. この 2 つのバーは, ダイナミックに口の高さと広さを表し, 今取られているパラメータをユーザに提示する.

3.3 性能評価

本システムにおいて, フレーズレベルの音声と原稿テキストとの対応付けは, ホスト側が半自動的にしてくれる. ボイスシンセサイザーが生成するナレーションの場合, パンクチュエーションが遵守されるため, 対応付けは正しく行われる. ところが, 人声によるナレーションの場合, 時々パンクチュエーションを無視したり, フレーズの途中で呼吸を入れたりするナレーションがあるため, 対応付けが乱れることがある. その場合, 本システムがのマッチングツールが提供するインタフェースを用いて対応付けを手動修正することができる.

本システムに使われるノート PC に付いているカメラは, 最大解像度 720p で最大 30fps の撮影機能を備えている. 本システム稼働中, プレゼンターによるすべての操作に対し, ホスト側プログラムの FPS は 55fps を上回っており, ディレイはほとんど感じられない.

4 ユーザテスト・評価

SmartVoice の効果を確認するために, 利用者評価実験を行った. 通常のスピーチをしている動画と, SmartVoice を用いてスピーチをしている動画とをビデオ編集によって切り替えたものを実験参加者に見てもらい, どの部分が「口パク」(SmartVoice による箇所か) を判定してもらった. 判定結果はマウスボタンの押下により入力してもらった. 声色の変化による影響を避けるために, この実験では本人が読み上げた音声を, 本人が SmartVoice によって再生するということにした. 8 名 (コンピュータサイエンス関係学科の大学院生) が実験に参加した.

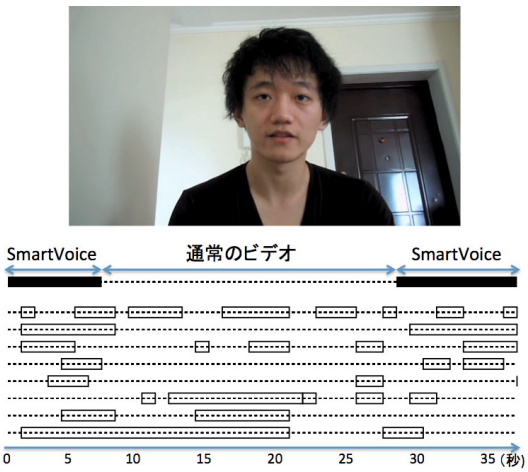


図 7. 生声原稿読み上げ v.s SmartVoice の実験結果

実験結果を図 7 に示す. それぞれの参加者が SmartVoice によるものと判定した部分を白抜きの矩型で示す. この図から明らかなように, SmartVoice ではない部分も口パクと判断した箇所がある一方で, SmartVoice による部分も本人の生のプレゼンテーションだと判断している箇所もある. この実験から, SmartVoice による音声対応は本人によるプレゼンテーションと比較しても自然であることがわかった.

5 議論

5.1 関連研究

フェイシャルアクションを使って, 何らかの制御にマッピングする前例として, Lyons らによる The Mouthesizer がある. The Mouthesizer では, ミニヘッドマウント CCD カメラを用いて, 口の中の影領域をトラッキングし, MIDI のエフェクト付きや楽器の操作などをユーザの口の動きにマッピングしている. [4] また, 口の高さ, 広さの他, 舌の動きと位置も使われている. Lyons らは, 楽器を演奏するミュージシャンに両手以外のコントローラーを与えた点では, The Mouthesizer はフットペダルと同じだが,

Mouth コントローラーはペダルより直感的で身に着けやすいことを証明し, フェイシャルアクションによるマシンインターフェイスの可能性も言及した. [5]SmartVoice はコンセプト上, The Mouthesizer のとシェアするが, 操作中に変わった表情などを取り入れていないため, プレゼンテーション中にユーザは自然な状態を保つことができる.

また, 同じくフェイストラッキングを用いるが, 音声ではなく, リアルタイムで CG キャラクターの表情を制御する研究もある. [6]

5.2 ボイスシンセサイザーによるビデオ作品の音声付け

従来, デモビデオなどにナレーションを入れたい場合, 特にナレーションが母国語ではない時に, 原稿のテキストファイルを用意し, ボイスシンセサイザーによって音声化する手法がある. しかし, 前述の通り, ボイスシンセサイザーは単語一つ一つをキレイに発音してくれるが, 抑揚に欠くため, 機械的に聞こえる難点がある. 本システムのように, 抑揚やタイミングを人間がコントロールすると, ボイスシンセサイザーによる声をビデオのナレーション作成やアニメーションの人声の作成などにも応用できると考えている.

5.3 アドリブの可能性

今まで述べてきたように, 本システムに使われる音声は, 事前に作っておかないといけない. しかし, プレゼンテーションの場合, 最初から最後まで原稿に準して喋る必要がなく, 途中で雑談を挟むことも考えられる. SmartVoice による音声と生声のスイッチによる違和感を隠蔽し, それを気づかせないようにすれば, SmartVoice と生声の間で自在に切り替え, プレゼンテーション中のアドリブも可能と考えている.

5.4 アフターレコーディングの可能性

ドラマや映画の作成は, 撮影と音声はそれぞれ単独に収録し, プロの声優のアフターレコーティングによってセリフを映像に入れるのは一般で的である. 本システムは, リアルタイムにカメラからキャプチャーした映像に限らず, 収録済みの映像中のキャラクターの顔をトラッキングして利用することもできるため, 出来上がった映像を見ながらアフターレコーディングを行わなくても, SmartVoice を用いれば, タイミング正しく映像に音声を入れることが可能である.

5.5 歌, 外国語練習への応用

繰り返しになるが, SmartVoice は「まるでその人が生で喋っているように見える」のようなシステムの実現が目的である. 広義的な「人間が喋る」という点から, 本システムが支援できるのは, スピーチやプレゼンテーションに限らないはずだ.

例えば, カラオケと類似するプロンプターを有するため, 練習しやすい上, SmartVoice はリアルタイムでピッチシフトとスピード変更もできるため, 音感・リズム感の悪い人に, ピッチ感覚とリズム感を身に着けさせ, 歌が上手くなることができるかもしれない.

また, 外国語学習者の中で, その言語をできるだけネイティブっぽく喋れることを望む人ことが多い. しかし, 発音やイントネーションを真似して, たくさん練習してそれを自分が喋る時の習慣とするのは結構大変である. 外国語学習者用に SmartVoice にネイティブ声との比較機能を付けることによって, ネイティブ風のイントネーションを身につけることができるかもしれない. さらにもっと細かな顔筋肉がトラッキングできれば, SmartVoice による発音矯正も不可能ではないと考えている.

6 結論とまとめ

本論文では, 言語の壁を超えたプレゼンテーションサポーティングシステム「SmartVoice」を提案した. 音声データを分析し, 原稿テキストとフレーズレベルの対応付けを実現することによって, ユーザの口の動きに合わせてスマートに音声を送り出すことを可能にした. また, ボイスシンセサイザーによる機械的なナレーションのイントネーションを, ユーザの口の形状・位置によってコントロールすることを実現した. 本システムの効果を検証するために, 評価実験を行い, SmartVoice による音声対応は本人によるプレゼンテーションと比較しても顕著な不自然がないことがわかった. 今後, 他の分野における SmartVoice の応用が期待できると考えている.

参考文献

[1] openframeworks. http://www.openframeworks.cc/

[2] ofxFaceTracker. http://github.com/kylemcdonald/ofxFaceTracker

[3] Dirac3L. http://dirac.dspdimension.com/

[4] L. Michael J., H. Michael, and T. Nobuji. The Mouthesizer: A Facial Gesture Musical Interface. In *Conference Abstracts, Siggraph 2001*, pp. 230, 2001

[5] L. Michael J., T. Nobuji. Facing the Music: A Facial Action Controlled Musical Interface. In *Proceedings, CHI 2001, Conference on Human Factors in Computing Systems*, pp. 309-310, 2001

[6] W. Thibaut, B. Sofien, L. Hao and P. Mark. Realtime Performance-Based Facial Animation. In *ACM SIGGRAPH 2011 Papers, SIGGRAPH ‘11*, pp. 77:1-77:10, 2011

視覚障害者のための写真撮影・閲覧・共有インタフェース

浅川 智恵子　　原田 丞　　佐藤 大介　　石原 辰也　　高木 啓伸 *

概要. 写真は視覚的なメディアではあるものの，多くの視覚障害者が記録や晴眼者とのコミュニケーション手段として興味を持っていることが明らかとなっている．そこで本研究では視覚障害者のニーズ調査の結果に基づき，写真と音声メモ・周辺音を結びつけて記録することで，非視覚的な撮影・閲覧・共有を可能にするスマートフォン・アプリケーション “Accessible Photo Album” (APA) を開発した．日常的な状況での評価のため，5 人の視覚障害者被験者が実験者とペアで外出して写真を撮影したうえで，別の実験者に写真を説明するという実験を行った．その結果，すべての被験者が音声メモと周辺音を活用して撮影・閲覧・共有を楽しむことができた．

1 はじめに

視覚障害者にとって写真は本質的にアクセス困難なメディアである．しかし，ソーシャルネットワークの普及によりコミュニケーション手段としての興味が高まってきた [6]．米国では，視覚障害者を対象とした写真教室もすでに実施されている [14]．視覚障害者が写真を撮影するためには，例えば構図の決定・露出・ブレ等の情報を非視覚的にどう伝えるか，撮影した写真を独力でどのように管理するか等の技術的課題が山積みであり，既に多くの研究が報告されている [6, 11, 1]．しかし，写真を共有するインタフェースの研究は例がなく，我々が実施したアンケート調査では，回答者の半数以上が，写真の共有に問題を感じていた [5]．

このメディアのギャップを埋めるためにはアクセス可能な情報，例えば音声を同時に記録する方法が考えられる．既にいくつかの音声録音可能なカメラやカメラ・アプリケーションが実用化されている [8, 9, 13, 16, 17, 18] が，これらは晴眼者を対象としているためアクセシビリティ上の多くの問題がある．ビデオで代替する方法も考えられるが，撮影時に対象物や人に対して一定時間カメラを向け続けなければならず，瞬間的に撮影可能な写真と比較して困難な操作が要求される．

そこで本研究では音声メモ付写真に着目し，視覚障害者による撮影・閲覧・共有を実現するためのスマートフォン・アプリケーション “Accessible Photo Album”(APA) を開発した．APA はスクリーンリーダを用いて操作することができ，撮影時に周辺音と音声メモを録音する機能を備えている．APA では閲覧時に音声に加えて日付，時刻，GPS 情報も非視覚的に参照可能なため，写真を容易に識別することが可能であり，さらなる解説を加えての共有が可能である．

本論文では関連研究に続いて APA のデザインと実装を紹介する．その後，評価実験とその結果について報告したのちに，視覚障害者が独力で情報を発信する手段としての可能性について考察する．

Copyright is held by the author(s).

* Chieko Asakawa, Susumu Harada, Daisuke Sato, Tatsuya Ishihara, Hironobu Takagi, 日本アイ・ビー・エム株式会社 東京基礎研究所

2 関連研究

視覚障害者の間で写真撮影に対する興味が高まっている [25, 10, 19, 20]．Jayant らのアンケート調査 [6] によると，アンケートに答えた 70%のユーザがきれいに写真を撮影することに興味があると答え，ほぼ同数が最近写真を撮影したことがあると答えた．写真の識別には携帯型点字メモ機を併用して写真を撮影した順序に合わせて説明文を記録する，写真のファイル名に意味を持たせる等の方法をとっている．

2.1 音声付き写真

Frohlich と Tallyn が提案した音声付き写真は，写真とそれに付加された録音音声を 1 つのメディアとしたものである [3, 4]．この音声付き写真の概念は，特に視覚障害者のために考案されたものではないが，視覚障害者が写真を撮影・閲覧・共有する上で非常に有効であると考えられる．

音声付写真を撮影するためのスマートフォン用アプリケーションは，すでにいくつか実用化されているが [8, 9, 13, 16, 17, 18]，アクセシビリティ上の多くの課題がある．例えば，スクリーンリーダ機能を用いると 1 枚の写真を撮るために，35 回のジェスチャ入力が必要となる．また音声付写真を撮影できるカメラも市販されているが，視覚障害者が操作可能な製品は筆者らの調査では見つからなかった．

2.2 視覚障害者のための写真撮影支援

iPhone の VoiceOver[15] や Android の TalkBack 等，スマートフォンのアクセシビリティ向上に伴い，

視覚障害者向けのアプリケーションが各種開発されている．写真関連では，画像認識やクラウドソーシングを用いて，視覚障害者が撮影した写真の内容を認識するサービスも利用が広がっている [2, 22, 24]．これらのサービスを利用することにより，ワイン，飲み物の銘柄，薬の種類，郵便物の内容等を遠隔からのサポートにより知ることができる．写真のブレの検出 [7]，視覚情報を音情報に変換する技術 [23]，写真のブレ，暗さ，傾きを音声で伝達する研究 [12] 等も報告されている．

3 Accessible Photo Album(APA)

視覚障害者が写真を撮影・閲覧・共有することを目的として APA を開発した．APA は以下の 3 つの方針に基づき設計した．

1. 素早く撮影し，音声メモと周辺音を録音できるように，非視覚的なジェスチャ操作を最適化する
2. 閲覧時に素早く写真を識別可能にするために音声メモの再生を効果的に取り入れる
3. 主体的に晴眼者と写真を共有可能にするために専用のモードを設けてインタフェースを最適化する

3.1 実装

APA は iPhone のアプリケーションとして実装し，読み上げには iOS 付属のスクリーンリーダである VoiceOver を用いるように設計した．APA はカメラモードとアルバムモードの二つのモードを持つ（図 1）．起動直後にカメラモードに遷移し周辺音の録音が開始され，1 本指で下スワイプすることにより写真が撮影される．1 本指でタップして押し続けること (タップ&ホールド) で音声メモの録音が開始し，指が離れるまでの間記録することができる．録音終了直後に録音した音声メモが自動で再生される．

アルバムモードでは日付別にソートされた写真のリストが表示される．リスト間の移動には 1 本指で左右スワイプ，日付間の移動には 1 本指で上下スワイプを用いる．次の写真へ移動する際に，対応する音声メモが自動的に再生される．再生と同時に写真が撮影された日時と位置情報も音声合成により読み上げられる．位置情報については，写真が保存された時点の GPS 情報を地名に変換して読み上げる．晴眼者と写真を共有するときには，リスト上の写真の位置でダブルタップすることでフルスクリーン表示することができる．フルスクリーンビューでは VoiceOver の標準ジェスチャ機能である 3 本指の左右スワイプを用いて写真間を移動することができる．新しい写真が表示されると，音声メモの再生に続いて，写真の向き (縦方向または横方向) が読み上げられるので，デバイスの方向を主体的に変更することが可能である．

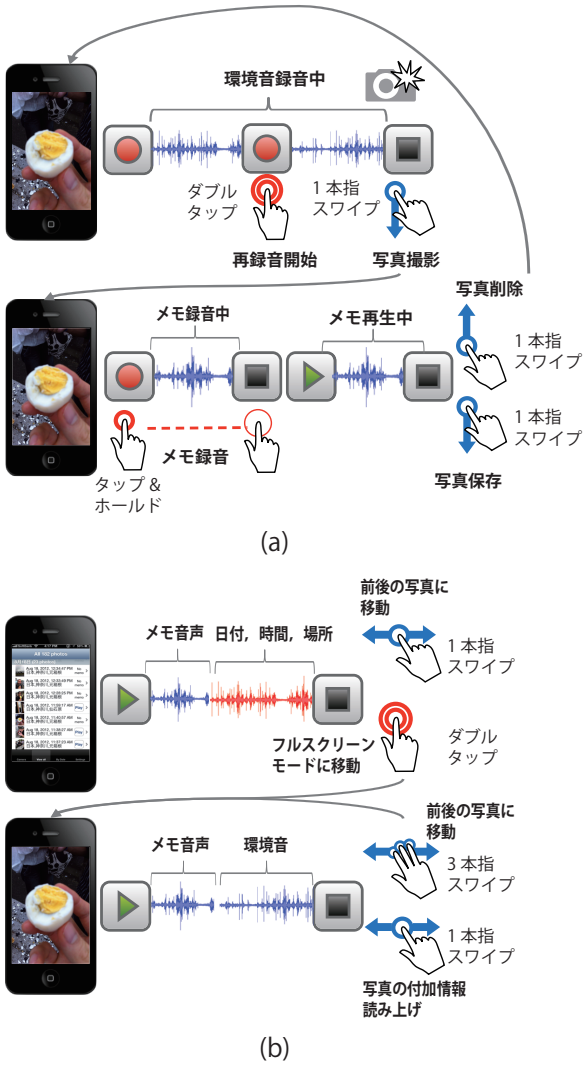


図 1. (a) カメラモードの操作方法，(b) アルバムモードの操作方法

4 ユーザ実験

5 名の視覚障害者被験者の協力により APA の有効性を評価した．実験では以下の 3 つの点に着目した．

(a) 写真の撮影，音声メモや周辺音の録音を素早く実行できたか?

(b) それぞれの写真の識別をしながら，素早く閲覧できたか?

(c) 晴眼者に対して主体的に写真を共有できたか?

4.1 被験者

被験者は，男性 3 名，女性 2 名，年齢は 25 歳から 36 歳までの 5 名で，謝礼を支払った (表 1)．

表 1. 被験者プロフィール.

	年齢	性別	視力	撮影経験	共有経験
P1	25	女性	弱視	有り	有り
P2	34	男性	全盲	無し	有り
P3	35	男性	全盲	無し	無し
P4	34	女性	弱視	有り	有り
P5	36	男性	全盲	有り	有り

4.1.1 写真撮影の経験

P1，P4，P5 は，カメラ付携帯電話で写真を撮った経験があり，画像のファイル名を変更することで写真を管理していた．P1 と P4 は，その方法に比較的満足していたが，P4 は，処理があまりにも単調であることから，日付や時刻で管理するようになり，結果としてあまり写真を撮らなくなった．P2 と P3 は，実験的に写真を撮った経験があるのみである．P3 は，自分のカメラで友人に写真を撮ってもらった経験はあった．全ての被験者が独力で良い写真が撮れるかどうか不安を持っていた．撮影経験のある被験者も全員が晴眼者のサポートを得ていた．

4.1.2 写真共有の経験

P1，P4，P5 は携帯電話あるいは PC から e メール添付により他者と写真を共有したことがあった．P4 は写真をブログに投稿した経験があった．P2 は自分で写真を撮った経験はないが，彼が共有したい写真を友人に撮ってもらいその場で SNS に投稿した経験があった．

4.2 実験手順

実験は 1 人 1 日 (約 8 時間) で行い，以下の 4 つのセッションで構成される．

1) 導入セッション

被験者は，まず筆者らの研究所にて事前インタビュー・APA の操作方法のトレーニング・実験手順の説明を受けた．この導入セッションは平均 1 時間半であった．

2) 写真撮影セッション

第 1 セッション終了後，被験者は実験者と共に写真を撮るためあらかじめ被験者が行きたいと興味を示した場所に出かけた．そのためそれぞれの被験者は異なる場所で写真を撮影した．これは可能な限り日常に近い状況可で実験を行うためである．目的地を決定後，実験者によって道順が確定される．これは単に美しい風景だけを写真に撮るのではなく，被験者が興味を持てるもの，例えば触って楽しめたり直接経験できるものに出会えるよう考慮したためである．実験者は，被験者に同行するだけでなく会話をしながら各地点を回り，周辺の環境の説明を行った．あらかじめ決められた訪問地はあったものの，いつ移動するかについては，被験者に決定させた．時には，元々の道順から大きく外れることになったが，被験者が別の興味を見つけたときにはそれを優先した．また写真を撮るタイミングも彼ら自身が決定できるよう考慮した．被験者は好きな場所で自分自身の写真を取ってもらうよう実験者に依頼できることとした．この場合，実験者が写真を撮り音声メモは被験者が録音した．被験者がカメラの向きを合わせる手助けを必要とした時には，実験者が手助けを行った．カメラの向きを手で修正する前に，まず言葉で向きの修正を手助けした．これは，実際に写真を取っているのは被験者自身であると感じられるように考慮したためである．

3) 写真共有セッション

約 5 時間の外での実験終了後，被験者は研究所に戻り，写真共有セッションを行った．写真の撮影経験に関する簡単な質問の後に，被験者は観察者（別の実験者）と撮影してきた写真を共有するためのセッションを実施した．被験者は，写真を共有しながらその日経験したことを観察者に共有する．音声メモを他人に聞かれることをためらった被験者がいたため，必要であればイヤフォンの使用を認め，被験者は観察者に聞かせたいと思ったときのみ音声を再生できることとした．観察者は被験者と写真を見ながら，その日の経験について自由に会話するよう指示された．被験者と外出した実験者も，写真共有セッションに同席した．観察者は，被験者が写真についての説明を始めるまでは，コメントや質問を待つよう指示された．これは，被験者が晴眼者の助けなしに，一人で写真の識別できるか確認するためである．観察者は，被験者が写真の識別が困難な時のみ，手助けするよう指示された．写真共有セッションは，平均 30 分前後を要した．

4) 実験後のインタビューセッション

1 日の実験の最後に，その日の経験や APA に関する被験者からのコメントを収集した．

4.3 実験結果

それぞれの実験の目的について，ユーザ実験により得られた知見を述べる．

(a) 写真の撮影，音声メモや周辺音の録音を素早く実行できたか?

VoiceOver の経験があったのは P3 のみであったが，他の被験者も，導入セッション中に APA を使用するために必要な基本的ジェスチャを習得することができた．写真撮影セッション中に撮影された写真は，13 枚から 20 枚であった (表 2).

7 段階のリッカート尺度によるアンケート結果を

表 2. 写真の撮影場所と枚数.

	場所	枚数
P1	ショッピングモール/レジャー施設	17
P2	ランドマークタワー/レジャー施設	18
P3	ワインショップ/レジャー施設	20
P4	史跡/公園	13
P5	中華街/ショッピングモール	15

表 3 に示す.「写真の撮影は容易であったか」との問いに対して，2 名の被験者が 7 を，その他は 5 で回答した.「音声メモの録音は容易であったか」の問いに対しては，2 名が 7，2 名が 6，1 名が 5 と回答した．携帯電話で写真撮影の経験があった 3 名の被験者に対して，以前の方法と，APA とを比較し，音声メモの詳細度，スピード，使いやすさについて評価を依頼した．その結果，3 つの項目全てに対して，全員が APA が優れていると評価した.

(b) それぞれの写真の識別をしながら，素早く閲覧できたか?

全ての被験者が問題無く閲覧操作を行え，撮影した写真を識別することができた．アンケートにおいても「閲覧は簡単だった」,「目的の写真を見つけることができた」,「撮影した状況を思い出せた」という問いに対して，P2 を除くすべての被験者が,「全くそう思う」または「そう思う」と返答した (表 3).

P2 だけは 2 枚の写真を混同したため,「目的の写真を見つけることができた」という問いに対して「どちらかというとそう思う」と返答している．この 2 枚は，シーンが同じで 1 枚は実験者が撮った被験者とランドマークタワーの写真であり，もう一方は被験者が撮ったランドマークタワーだけであったが音声メモがほぼ同じであったため，P2 はその区別ができなかった.

(c) 晴眼者に対して主体的に写真を共有できたか?

「写真共有セッションを主体的に行うことができたか」という問いに対して，P2 を除くすべての被験者が,「全くそう思う」または「そう思う」と返答した (表 3)．P2 は，前述の 1 組の写真の識別ができなかったことから,「どちらともいえない」と答えた．この判断ミスが，彼の APA 全般に関する評価にどの程度影響を及ぼしたかを判断することは難しいが，他のすべての写真については，問題無く詳細を説明できていた.

4.4 観察

周辺音の有効性

P5 が中華街の門で撮影した写真を共有していた時，周辺音として水しぶきが再生された.「自分のすぐ隣でだれかが水しぶきをあげていたことに気づいていなかった」と，新たな発見をコメントした．P1 と P4 は，当初観察者に音声メモを共有することをためらっていたが，2 枚の写真について周辺音を共有したところ，その価値に気付いたとコメントした．このときの周辺音は，噴水とブラスバンドの音であった．また P4 は，周辺音を利用すれば視覚障害者の友人とも写真を通して経験を共有できるとコメントした.

表 3. アンケート結果
(1:全くそう思わない，7:全くそう思う)

	P1	P2	P3	P4	P5
写真撮影セッション					
撮影は楽しかった	7	4	6	7	6
撮影は簡単だった	5	7	5	7	5
メモ録音は簡単だった	6	7	6	7	5
独力で写真を撮りたいと思うようになった	3	4	7	7	4
以前の方法との比較					
より簡単だった	7	-	-	6	7
より早くメモを残せる	7	-	-	6	5
より詳細なメモを残せる	6	-	-	7	6
写真共有セッション					
閲覧は簡単だった	7	7	6	7	7
目的の写真を見つけることができた	6	5	7	7	7
撮影した状況を思い出せた	7	6	6	6	7
写真を見せることは楽しかった	7	7	7	6	6
写真共有セッションを主体的にできた	6	4	6	6	7
実験全体について					
写真共有は楽しい	7	6	6	7	6
同様のアプリを使って写真を共有したい	6	7	6	7	3
実験前より写真を撮りたいと思う	6	6	6	7	7
実験前より写真を共有したいと思う	7	7	6	7	4

主要なコメント

すべてのコメントは紙面の制限から掲載できないため，以下に主要なものを紹介する.

P1:「音声メモによって，自分で写真の管理ができるようになった．これはすばらしい進歩である.」

P4:「とても楽しかった．うまく写真がとれているかどうか不安だったため，写真共有についての楽

しさの評価は 6 としたが，それを除けば素晴らしい経験だった．これまでは自分の経験を言葉でしか伝えられなかったが，写真を使うことで視覚的に情報が伝えられることの楽しさを知った．こんな経験は初めてだった.」

P5:「写真が表示されたとき，すぐにそれについての話ができることは素晴らしい．周辺音に録音された周りの声を聞くと，写真を超えた情報を得ることができたと感じられて，とても良い気分である.」

5 考察

5.1 主体的な写真共有

APA の目的の 1 つは，視覚障害者が主体的に晴眼者との写真共有を可能にすることであった．その点で P3 は最も主体的に，演出も交えて写真共有セッションを楽しんでいた．ある写真を見せる前に，まず観察者に音声メモが聞こえないようイヤフォンを装着し，次に画面が観察者に見えないよう iPhone を裏返した.「多分私がどんなボトルを偶然みつけたか想像できないと思いますよ」と話した後に，画面を元に戻し「コンコルド型のウィスキーのボトル」の写真を観察者に見せた．このように手に持つことのできるデバイスだからこそ可能な視覚的演出を，その効果を判断することのできない視覚障害者自身がとったという点で実験者らの予想を超えた行動であった．今後は，このような “演出” も含めた写真共有の表現力向上を目指したインタフェースの工夫や知見の共有が必要となるであろう．

5.2 ユーザインタフェースの改善

音声メモや周辺音をどのタイミングで録音するかは，APA のようなアプリケーションを設計する際に重要な要因となる．実験に使用した APA では音声メモと周辺音を別々に取り扱った．シャッターが押されるまでの “周辺音” には晴眼者によるカメラの向きの指示や，周囲の状況説明などの音声が含まれてしまったため，実際には撮影後の “音声メモ” のほうが有用であった．

そこで，実験後に APA のインタフェースを修正し，シャッター前録音を廃止し，シャッター後にユーザ自身が “タップ&ホールド” ジェスチャにより，周辺音と音声メモを自由に録音できるインタフェースに改良した．この場合周辺音と音声メモの記録上の違いはなく一つの音声ファイルとして保存される．これによりシャッターを押す，“タップ&ホールド” でメモと周辺音を録音するというシンプルなインタフェースに変更された．

5.3 音声メモとテキストメモ

これまで写真にテキストでコメントをつけていた被験者も，音声メモは簡単に，素早く，そして周辺音を含めた詳細な説明が付与できるとコメントした．確かに音声メモは作成が容易であるという大きな利点があるが，テキストには写真の検索が容易になる，ネットワーク上で共有するときに既存サービスと親和性が高いなどの利点がある．今後，音声認識の利用も含めてテキストを音声メディアに取り組む必要がある．

また，二人の被験者が音声メモを後で修正して詳細度を向上したいとコメントした．P5 は，観察者からのコメントを音声メモに追加することで，さらに写真の回想が容易になると述べていた．このように音声メモの編集に対するニーズは高く，検討の必要がある．

6 まとめ

本研究では視覚障害者による写真の撮影･閲覧･共有を可能にするスマートフォン・アプリケーション “Accessible Photo Album”(APA) を開発し，評価実験を行った．APA では写真撮影と同時に音声メモや周辺音を録音し，閲覧・共有時に自動的に再生することで視覚障害者が独力で写真を識別し共有することが可能になる．実験の結果，晴眼者とであっても主体的に写真の共有が可能になるだけでなく，独自の演出表現までも可能になることを観察できた．写真共有は晴眼者だけでなく視覚障害者にとっても価値がある活動になることがこの実験でも裏付けることができた．

今後，より豊かな写真体験を視覚障害者に提供するためにはメディアのギャップを埋めるためにメタデータの拡大が重要な課題の一つとなる．例えばユーザが室内・外どこにいるか，位置情報，コンパス，データベースを組み合わせて写真に写っている建物の詳細な解説などが付与できるであろう．さらにクラウドソーシングのアプローチにより写真の詳細な説明を付加することも可能になってきている．写真は「ソーシャルメディア時代の新たな言語である」とも言われるほど，その重要性が高まっている．視覚障害者もそのような時代に独自の方法論を蓄積してコミュニティ参加できるようにするための技術開発が今後も必要となるであろう．

参考文献

[1] Baber, C., Cross, J., Khaleel, T., and Beale, R. Location-based photography as sense-making. *P*roc. BCS-HCI 2008, British Computer Society (2008), 133–140.

[2] Bigham, J.P., Jayant, C., Ji, H., et al. VizWiz: nearly real-time answers to visual questions. *P*roc. UIST 2010, ACM (2010), 333–342.

[3] Frohlich, D. and Tallyn, E. Audiophotography: practice and prospects. *E*xt. Abstracts 1999, ACM (1999), 296–297.

[4] Frohlich, D.M. Audiophotography: Bringing photos to life with sounds. Kluwer Academic Pub, 2004.

[5] Harada, S., Sato, D., Adams, D. W., Kurniawan, S., Takagi, H., and Asakawa, C. Accessible photo album: enhancing the photo sharing experience for people with visual impairment. *P*roc. CHI 2013, ACM (2013), 2127–2136.

[6] Jayant, C., Ji, H., White, S., and Bigham, J.P. Supporting blind photography. *P*roc. ASSETS 2011, ACM (2011), 203–210.

[7] Ko, J. and Kim, C. Low cost blur image detection and estimation for mobile devices. *P*roc. ICACT 2009, IEEE Press (2009), 1605–1610.

[8] Many Ltd. Picle. http://www.picleapp.com/.

[9] Robinson, S. Com-Phone Story Maker. https://play.google.com/store/apps/details?id=ac.robinson.mediaphone.

[10] UCR/California Museum of Photography. Sight Unseen. http://www.cmp.ucr.edu/exhibitions/sightunseen/.

[11] Vazquez, M. and Steinfeld, A. Helping visually impaired users properly aim a camera. *P*roc. ASSETS 2012, ACM (2012), 95–102.

[12] White, S., Ji, H., and Bigham, J.P. EasySnap: real-time audio feedback for blind photography. *P*roc. UIST 2010, ACM (2010), 409–410.

[13] Luo Xu. Speature. http://itunes.apple.com/us/app/speature/id401887412.

[14] Blind with Camera School of Photography. http://www.blindwithcameraschool.org/.

[15] Apple - Accessibility - iOS - Voiceover. http://www.apple.com/accessibility/ios/voiceover/.

[16] StoryMark Life. http://www.storymarklife.com/.

[17] Voicepic. http://voicepic.me/.

[18] Storyrobe. http://storyrobe.com.

[19] Photos by Blind Photographers - Photo Essays - TIME. http://www.time.com/time/photogallery/0,29307,1897093_1883579,00.html.

[20] Seeing Beyond Sight: Photographs by Blind Teenagers. Book, video & TED Talk. http://www.seeingbeyondsight.org/home/.

[21] Blind With Camera. http://www.blindwithcamera.org/.

[22] oMoby. http://www.iqengines.com/omoby/.

[23] Seeing with Sound - The vOICe. http://www.seeingwithsound.com/.

[24] TapTapSee. http://www.taptapseeapp.com/.

[25] 伊藤 邦明． 七重，光をありがとう．河出書房新社，2000

未来ビジョン

APA を用いて作成された音声情報群を共有することで，他の視覚障害者の経験を楽しむメディアとして利用できる可能性がある．ユーザ自身が音声情報を写真・地理情報とあわせてソーシャルネットワーク上で共有することで，視覚障害者自身の手でソーシャルなオーディオデータベースを作り出すことができる．このアプローチをジオボイスタグと呼ぶこととする．類似したサービスはこれまでも様々な場面で議論されてきた．しかし，テキストタグやフォトタグと比較してあまり注目されていない．理由の一つはジオボイスタグを利用したアプリケーションが限られるためであろう．視覚障害者にとっては状況が異なる．例えば視覚障害者が晴眼者の友人とジオボイスタグを聞きながら歩けば，視覚的には得られない周辺情報を獲得することができ，より対等な立場から会話を楽しむことができる．晴眼者は意図的に周辺情報の説明をする必要がなくなり，視覚障害者はレストランやショップの情報，それらの評判についても主体的に会話をリードできるようになる．視覚障害者が一人で街中を歩く場合のナビゲーションサポートの可能性もある．このシナリオ実現には周辺音を妨げないでボイスタグを提示する必要がある．視覚障害者は通常，車の音をはじめとした周辺音を音声ランドマークとして街中を移動するためである．共有サービス自体のデザインと並行して，使用状況を考慮したインタフェースや，ソーシャルメディア上の情報との融合，高齢者等の他のユーザ層に向けたユースケースなどを検討する必要がある．

個人的な小さな幸せを実現するブラウザ上での動画編集・共有手法

中村 聡史[*]　**石川 直樹**[†]　**渡邊 恵太**[‡]

概要. ニコニコ動画では，動画を一緒に楽しむだけではなく，もともとある動画をもとに，それを改良し，さらなる動画とするN次創作がそのコミュニティの一体感と，動画の再発掘に有効に働いている．しかし，こうしたN次創作はある程度の創作が前提となっているものであり，例えば音量の変更や，音楽の変更や，部分的な音声の差し替え，音声の追加といった事だけでは創作とはなりえないものである．そうしたちょっとした編集だけを行った動画をアップロードすると，パクリとなじられるであろう．本研究では，そうした微細な編集によって個人的な満足感を達成可能とするだけでなく，その編集を共有することで新たな動画視聴体験を可能とする仕組みを実現する．ここでは,「ブラウザ中で視聴中の動画に対する音量増減や音声の付与などの編集手法」と「付与された動画の共有手法」を実現とする．これにより，部分的に音量を調整したり，音楽を差し替えたり，音声を足したり，部分的にスロー再生やスキップしたいなど個人的な嗜好に基づく動画編集および視聴を可能とする．

1　はじめに

YouTube[1]やニコニコ動画[2]に代表される動画共有ウェブサイトが爆発的に成長している．YouTubeでは2012年1月時点で1秒あたりに動画長にして1時間分の動画が投稿され[3]ている．また，ニコニコ動画でも，2012年11月20日の中村らの調査 [1] によると，1日あたりに約5500本の動画が，1秒あたりに動画長にして約47秒分の動画が投稿されている．YouTubeは2012年1月の時点で1日に40億視聴，ニコニコ動画では2012年第2四半期に1日あたり1億視聴，平均視聴者数は786万人で，ユーザの平均視聴時間は102.5分であるという[4]．

ニコニコ動画では，先人が創作した動画を活かし，新たな動画を創作するというN次創作 [2, 3] が多く見られる．N次創作では，元々VOCALOIDなどで作成および投稿された音楽に対して，それにマッチする映像を付けてさらにアップロードしたり，ピアノなどで曲を弾き直してみたり，曲はそのままに歌を自分が歌ったもので差し替えてみたり，踊りを考えて音楽に合わせて踊ってみたりなど，多様な発展が見られる．こうしたN次創作は，ニコニコ動画を発展させ，コンテンツをより良い物へと高めるとともに，愛されるコンテンツを生み出している理由の一つでもある．

Copyright is held by the author(s).

[*] 明治大学 総合数理学部, JST CREST
[†] 東京農工大学 工学部
[‡] 明治大学 総合数理学部, JST CREST

[1] http://www.youtube.com/
[2] http://www.nicovide.jp/
[3] http://jp.techcrunch.com/archives/20120123youtube-reaches-4-billion-views-per-day/
[4] http://dic.nicovideo.jp/a/ニコニコ動画

ここで，ニコニコ動画でN次創作に参加するハードルについて考える．オリジナルのコンテンツに対して「歌ってみた」「演奏してみた」「踊ってみた」などタグが付けられているような動画はN次創作の1つである．ちなみに，2013年8月29日時点で，ニコニコ動画上で「歌ってみた」で検索される動画は691,641件,「演奏してみた」動画は166,800件,「踊ってみた」動画は100,078件ある．YouTubeで「歌ってみた」で検索される動画は約904,000件,「演奏してみた」動画は約586,000件,「踊ってみた」動画は約540,000件ある．なお，海外での「歌ってみた」に相当する（吹き替えも含まれる）「Fandub」で検索すると，その件数は約1,960,000件となる．

N次創作は，オリジナルとなるコンテンツ（動画）を改良するものであるため，参入障壁は，コンテンツを最初から作成するよりは低いように感じられる．しかし，この「歌う」「演奏する」「踊る」という事であっても技量が要求され，ハードルが高いといえる．さらに，アレンジする場合は，その技術力も必要となる．映像をつけるのも同様のハードルの高さが存在するであろう．結果として，N次創作は行われているものの，そこに参入しているユーザの総数は限られる．また，動画編集ソフトウェアを用いた動画編集や，動画のエンコード，動画のアップロードなどにも障壁は存在する．

ここで，ユーザが動画を視聴している際に，その動画に対して微妙に編集したいと思うことがある．例えば，ユーザが気に入っている楽曲動画なのだけれど，音量が終盤でフェードアウトしておらず，ぷつっと切れてしまうためフェードアウトさせたいと感じることや，作業用BGMとして音楽のメドレーを聴いている際に，ある曲だけは苦手なのでスキップしたいと感じることなどもある．それ以外にも，折角

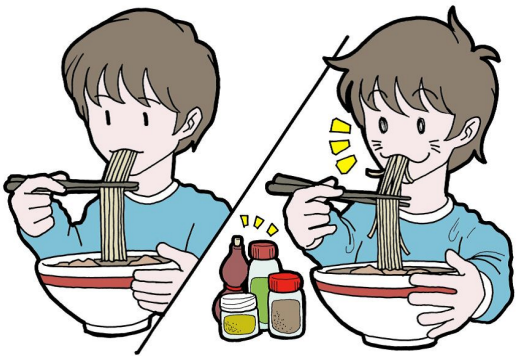


図 1.（左）調味料をたさずにそのまま食べる（右）調味料を足して自分好みに調整し食べる

の素晴らしい演奏動画なのだけれど普通に終わってしまうため拍手や歓声などを入れたいなどと感じることや，映像自体は好みなのだけれど BGM があまり好きなタイプではないため，BGM を差し替えたいなど様々である．しかし，仮に高価な動画編集ソフトウェアをもっており，編集能力が高いユーザであったとしても，他者が投稿した動画をダウンロードし，音量を微調整だけしたり，部分的に音を追加したり，部分的にスキップしたりといった，微細な変更だけでアップロードするということは世の中から N 次創作として許容されるとは考えにくい．そうした動画をアップロードすると，他の視聴者から「パクリ」と糾弾されてしまうこともあるであろう．

そこで我々は，上記のような編集を，オリジナルの動画に手を加えるのではなく，システムによって再生している動画自体をコントロールすることによってコンテンツを個人化する手法を提案および実装する．また，他者と共有可能にすることによって，編集も 1 つのコンテンツとして楽しむことが可能な仕組みを実現する．これは，ラーメン屋などにラーメンを食べに行った時に，自分の好みに合うように調味料を足すことに似ている（図 1）．

本提案手法によって，N 次創作とは違う，よりカジュアルに動画と係ることが可能な N 次装飾とも言える一つの動画の楽しみ方を提供する．

2　ブラウザ上でのカジュアルな編集と個人化

ニコニコ動画に見られる N 次創作は，その創作のもととなるオリジナルのコンテンツに敬意を払いながら，利用するという点でとても良いコンテンツ生態系を作り出していると言える．しかし，動画の作成者（N 次創作の作成者も含む）は，動画の視聴者数に比べたら圧倒的に少ない．これは，ニコニコ動画自体にアップロードされている動画の平均的な質が高いために，気後れしてしまっているという点もあるが，そもそも動画作成および N 次創作のハードルの高さも原因の 1 つでもある．

ここで，オリジナリティとクオリティが高い楽曲や映像の動画があり，多くの人から評価されているが，ユーザ個人としては「微妙に気持ち悪いから修正して欲しい」とか,「もう少しこうであったら良いのに」などと感じることがある．実際，そうしたコメントはニコニコ動画中のコメントで現れていることも多く，不満コメントに対し「あなたが作りなおしてアップロードしたら良いじゃないか」という旨のコメントが付くことも珍しくない．不満を述べるだけというのは実際生産的ではないが，元々の動画の投稿者以外のユーザが，その高い品質の動画をダウンロードし，ユーザの個人的な好みに合わせて微調整だけした動画をアップロードするというのは，N 次創作とは言いがたいものであり，他者からも受け入れられないであろう．

ここで，ユーザ自身がもつ，個人のための動画編集のニーズを調べるため，ニコニコ動画や YouTube などに投稿されている不満や要求に対するコメントを分類および整理し，ピックアップすると下記の通りとなる．

- 動画自体は質の高い映像コンテンツではあるものの，音楽が好みのものではないうえ，ユーザとしては映像と音楽が合っていないように感じている．そこで，映像はそのままに，音楽だけ好みのものに変更したい．
- ユーザが楽しんでいたストーリー性のある動画において爆発音が入っていたが，その爆発音の部分が好みに合わなかったため，他の爆発音と差し替えたい．
- N 次創作として「歌ってみた」や「演奏してみた」「踊ってみた」という動画は，オリジナルの同一楽曲を元に，各投稿者が歌ったり，演奏したり，踊ってみたりしたものである．このそれぞれの動画について，踊ってみた動画の映像と歌ってみた動画の音声を組み合わせたり，歌ってみた動画の音声とオリジナル動画の映像を組み合わせたり，ギターで演奏している動画とピアノで演奏している動画を組み合わせたりしたい．
- 作業用 BGM とタグ付けられている動画は，複数の楽曲をセットとして 1 動画とされているものであり，ユーザはこの動画を再生するだけで複数の楽曲を連続的に楽しむことができる．その楽曲の一部が気に入らないのでスキップしたり，ある曲だけ音量が低すぎてバランスが悪いため，その楽曲部分のみ音量を調整したい．
- 研究に関するデモビデオが無音であり面白みにかけていた．そこで，そのビデオに追加音声や，音楽を付けて面白くしたい．

- ある素晴らしい出来の動画があるが，音声がついていない．みんなで音声を持ちよって，この無声動画に音を付与したい．

上記のニーズは一部に過ぎないが，その多くはオリジナリティがあるとは言いがたく，独自に動画を投稿した時に許容されるかどうかは怪しい．つまり，上記のニーズを満たすには，動画として再編集およびアップロードするのではなく，個々の編集情報だけを記録し，その編集記録を元に動画を視聴可能な仕組みの実現が必要となる．我々は，こうした動画の個人化（自身の好みに変更すること）をブラウザ上で実現可能とし，その超個人化された編集情報を他者と共有することを可能とするものである．

3 プロトタイプシステム

3.1 必要要件

ユーザ（視聴者）の嗜好に基づき，微妙に動画コンテンツを編集し，個人向けにしたいという欲求を満たすため，我々はウェブブラウザ上で，実際に視聴しながら編集することが可能な編集手法を実現する．システムとして求められる要件を整理すると，下記の通りとなる．

- 視聴中の動画の任意の再生時間に対して，ボリュームを上げる，下げる，消すなどの音量に対するコントロールを可能とする．
- 視聴中の動画の任意の再生時間に対して，一部スキップや，スロー再生などの再生に対するコントロールを可能とする．
- 視聴中の動画の任意の再生時間に対して，任意の音声，音楽追加を可能とする．
- 上記機能を Web ブラウザ上で実現可能とする．
- 動画と編集情報による，新たなる動画編集 ID（URL）で動画を管理する．
- 動画編集 ID をベースとして，他者と共有することを可能とする．
- 共有された編集情報を元に，さらなる編集を可能とする．

3.2 実装方法

本提案システムを，Google Chrome の拡張としてプロトタイプシステムを実装した．本システムは，YouTube やニコニコ動画などのページで動画がロードされると，動画の ID を取得（YouTube やニコニコ動画上で動画を一意に識別する文字列）し，その ID に合致する編集情報の有無を確認する．編集情報が存在しない場合は，動画編集のためのコントロールを提示し，編集情報が存在する場合は，その情報に基づきコントロール上に編集情報を提示する．ここで，コントロールの表示は，ページ中にコントロール用のコードが含まれる DOM を挿入することで実現している．



図 2. システム導入前（左）と導入後（右）

音量の変更や，動画のスキップ，再生速度変更といった再生に関するコントロールは，ニコニコ動画や YouTube の動画を再生するコントローラの API を利用して操作している．また，付与される音声については，その音声がある URL 上の音楽／音声ファイルの場合は，そのファイルをロードし，時刻に合わせて同期的に再生させる．さらに，YouTube などの動画を追加音声として利用する場合は，別のタブとして動画をロードし，タイミングを合わせて再生するよう実装している．つまり，システムとしては同時に複数のコンテンツを再生しているだけである．

なお，システムはニコニコ動画および YouTube では，動画プレイヤをコントロールする API を利用しているだけである．そのため，ラッパーを作成するだけで，他の動画共有サイトなどにも適用可能となる．

3.3 操作方法

図 2 の左は，本プロトタイプシステムを導入する前の YouTube 上のオリジナルのインタフェースで，右が本プロトタイプシステムを導入した後の YouTube 上のインタフェースである．

編集可能インタフェースは，音量コントロールグラフとシークバー，スキップ，再生速度，共有という 3 つのボタン，そして「Drop a media file here」という点線の四角形で囲まれた領域からなる．YouTube では動画下部に編集インタフェースが追加され，ニコニコ動画ではシステムの都合上動画上部に編集インタフェースが追加される．

音量コントロールグラフでは，ユーザはグラフ上の点をマウスのドラッグ･アンド･ドロップ操作によって上下させることにより，その再生時間における音量を変更することが可能となっている．また，シークバーは，通常の動画のシークバーのインタフェースと同じく，動画の再生に応じて自動的に動いていくうえ，ユーザの操作によって任意の再生時間に移動することが可能である（図 3）．

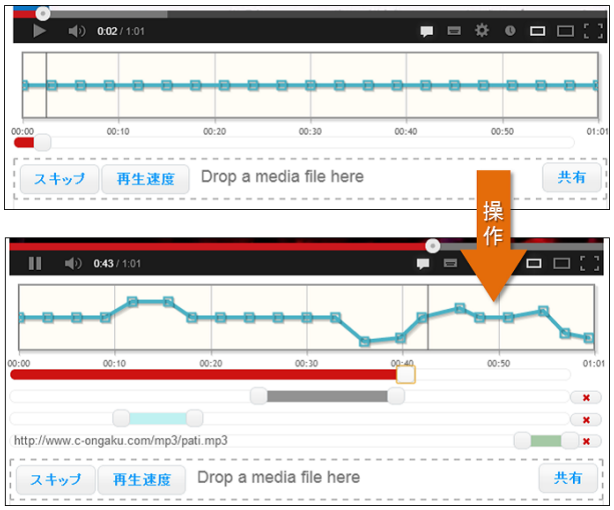


図 3. 編集インタフェースのみを拡大した状態

ユーザがスキップボタンを押すと，編集インタフェースに 1 つの時間軸が追加され，レンジスライダー（灰色背景の部分）が現在の動画の再生場所を左端に，そこから 10 秒後を右端にした状態で追加される．このレンジスライダーが指定されている部分が自動にスキップされる区間となる．ここで，スキップ開始場所はレンジスライダー左端のドラッグ・アンド・ドロップ，スキップ修了場所はレンジスライダー右端のドラッグ・アンド・ドロップによって変更することが可能である．

ユーザが再生速度ボタンを押すと，編集インタフェースに 1 つの時間軸が追加され，スキップボタンの時と色が違うレンジスライダー（水色背景の部分）が追加される．現在は，この再生速度ボタンによって指定された部分は，0.5 倍速で再生される（プロトタイプシステムでは固定）．

「Drop a media file here」という点線で囲まれたスペースに，ユーザがローカルのメディアファイルや，ウェブ上のメディアファイルの URL をドラッグ・アンド・ドロップすると，編集インタフェースに 1 つの時間軸が追加され，レンジスライダーが追加される．レンジスライダーは，現在の動画の再生場所を左端に，そこからドロップしたファイルの長さ分後を右端にした状態で追加される．ここでは，ユーザがレンジスライダーの左端，右端どちらをドラッグ・アンド・ドロップしても，メディアファイルの開始位置が変更される．また，ドロップされたファイルがどこから取り込まれたのかという情報は，時間軸インタフェース上に合わせて提示される．これにより，リファレンス情報が明確になる．

ボタンの押し下げや，ファイルのドロップなどによる編集内容は，追加された時間軸の右端に登場する×ボタンを押すことにより削除することが可能となっている．また，対象となる動画をユーザがあらかじめ編集している場合は，その編集情報が自動的

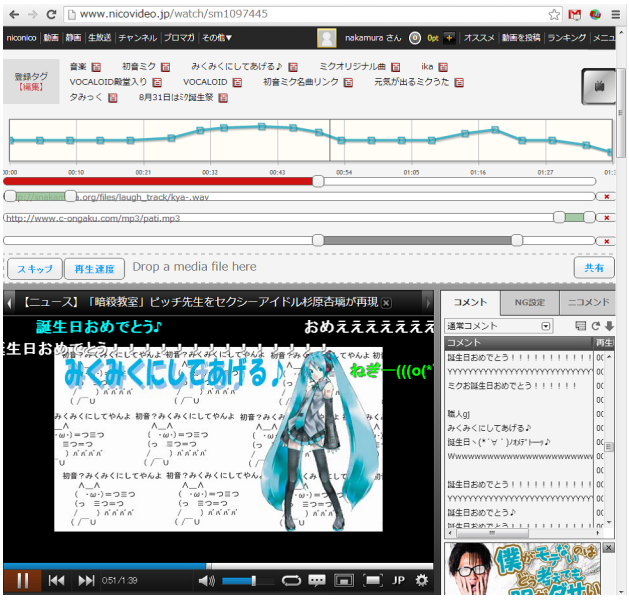


図 4. ニコニコ動画上のインタフェース

にロードされる．

なお，編集情報はローカルに保存されるが，共有ボタンを押すことによって，編集情報のみを他者と共有することが可能となっている．

3.4 利用例

図 4 は，本システムをニコニコ動画上で利用している様子である．ここでは，動画の最初の部分に大歓声を入れ，通常再生しサビ部分に向けて音量を徐々に上げている．また，1 つ目のサビ終了後，2 つ目のサビは飛ばしてすぐに音楽の終盤に飛び，音楽の最後では音量を下げるように指定されている．さらに，動画の最終盤に拍手と歓声が再生されるようになっている．

図 5 は，本システムを YouTube 上で利用している様子である．動画自体は，ある研究者が 20 個のテーマを 45 秒ずつ（トータル 15 分）で発表している様子である．ユーザがこの発表の動画を他者に紹介しようと思った際に，10 分しか紹介に使える時間がなかったため，紹介相手が興味をもってもらえなさそうなテーマについてはスキップするように指定する一方，あっという間に進んでしまってわかりにくい部分については，再生速度を変更しスロー再生することによって時間を長めにとっている．また，発表者の音量が大きかったり小さかったりとぶれがあったため，なるべく一定になるように音量コントロールインタフェースで変更している．さらに，全体的に音声が盛り上がりに欠けていたため，笑って欲しい所にラフトラックを入れたり，最後に盛大な拍手を入れることによって，盛り上がりを演出している．



図 5. YouTube 上でのインタフェース

4　考察

本提案手法により，演奏してみた動画の最後に拍手を入れたり，踊ってみた動画の最後に歓声を入れたり，歌ってみた動画の途中途中に合いの手の声を入れたり，科学系の動画の途中途中に解説の説明を音声として入れたり，わかりにくい部分はスロー再生または停止して説明したりといった事が可能となった．この既にある動画に対する編集（装飾）は，その動画コンテンツを変質させる．例えば，ストーリー性のある動画について，途中途中にラフトラック（笑い声）を挿入すると，より面白く感じるという事が実際のコンテンツ制作より観察された．今後は，その有効性をユーザ実験などによって明らかにしていく予定である．

動画の再生時間長が長い場合に，スキップや挿入した音声などがどこを指しているのかがわかりにくく操作しづらいという問題があった．この問題については，動画編集ツールにありがちな時間軸を拡大縮小するようなインタフェースを導入することである程度解決可能であると考えられる．

今回提案している手法は，様々な所から集めてきたリソースを 1 つの動画としてコンテンツ内に取り込むのではなく，部品として外部に存在させつつも 1 つの動画であるかのように振る舞わせることができるということである．つまり，どの音声，どの音楽がどこから来たのかというリファレンス情報が明確に存在しているということであり，そのリファレンス情報を元にオリジナルコンテンツへのアクセスも可能にするものである．ウェブ上の動画などで，リソースの取得元を明記していないケースも多いうえ，仮に明記されていたとしてもアクセスは容易ではない．そういった点で，オリジナルの作成者も尊重しつつも，リファーしながら新たなコンテンツを生み出すことができる事が可能となっている．

これまで，ニコニコ動画など動画共有サイトでは，オリジナル動画の作成者と，N 次創作をする動画の作成者，動画の視聴者という 3 つのタイプの利用者が存在していた．ここで，オリジナル動画の作成者や N 次創作をする動画の作成者に比べ，動画の視聴者の数の方が圧倒的に多い．我々の手法は，この圧倒的に数が多い動画の視聴者自体に，編集という形で動画に積極的に関与することを可能にするものであり，動画コンテンツを中心とした生態系がより成長するのではないかと期待される．

一方，本提案手法はオンライン上での編集および視聴を可能とするものであるため，公式サイトから配信されているような動画コンテンツも，本手法を使うことによって擬似的に編集し，視聴することが可能となる．例えば，公式のチャンネルとして，アニメなどが配信されているものについて，オープニングやエンディングを飛ばしたり，ダイジェストシーンのみにするよう編集しておき，ダイジェストのみ視聴といったことも可能となる．また，そうした公式の動画に対して，音声を付与したり，別の音楽に差し替えたりといった事も可能となる．また，本手法ではある公式動画の映像と，ある公式動画の音声の組み合わせを動画を違法にダウンロードして編集するのではなく，同期的に再生することで組み合わせとして楽しむことが可能となる．このような動画の使用は，コンテンツの世界を広げるものと期待される．

また，本提案手法を利用することによって，ユーザが好きなアーティストの楽曲を CD から取り込んで mp3 化し，ウェブ上のお気に入りの動画の映像に合わせて再生する（その動画の音声自体は消音にする）ということも可能となる．ここでは，ユーザはそうした mp3 ファイルをウェブ上にアップロードして他者と共有するわけではなく，個人的に楽しむだけであるため，著作権の侵害にも当たらず，楽しむことが可能となる．

現在は，YouTube とニコニコ動画のみに対応しているが，システムの処理としては YouTube やニコニコ動画の動画コントロールのための API を利用しているだけである．そのため，そうした API さえ存在し，ラッパーとなるプログラムを書けば，他の動画サービスでも本手法を適用可能となる．例えば，Ustream などに保存されている，WISS の登壇発表などに対して，説明の音声を付与したり，発表と無関係なシーンを飛ばしたりといったことが動画

を編集すること無く実現可能となる．

5 関連研究

Seek Rope[4] は，YouTube などの動画共有サイトにおいて，シークバーをロープのように取り出し，ロープを丸めることで繰り返し再生したり，ロープを繋げることで複数の動画をつなげたり，ロープを格子状に配置することでランダム再生を行ったりすることを可能とするシステムである．ウェブ上のコンテンツを活用するという点で類似性があるが，本研究は再生のコントロールのみならず，音声や音楽，他動画などによって装飾可能であるという点で違いがある．

ウェブ上のコンテンツを組み合わせて新たなコンテンツを作り出す試みとして Massh![5] がある．Massh!では，楽曲同士をインタフェース上でつなぐことによってマッシュアップ音楽を作成可能としている．ウェブ上のコンテンツを組み合わせ，利用するという点で類似はあるが，我々の手法は個人的な欲求を満たすというものであり，その表現方法や目的は異なる．

MixBit[7] は，短い動画クリップ（16 秒）を組み合わせることで，新たな動画を作成可能とするシステムである．MixBit では，ユーザ自身が動画を撮影せずとも他者が撮影した動画を活用して動画制作が可能となっており，動画制作のハードルを下げている．ただ，MixBit は閉じたコミュニティの中の動画を再生時間に沿って順に繋ぎ合わせるものである．我々の手法は，コミュニティに閉じること無くすべてのオンラインコンテンツを対象としており，また映像と音声などの組み合わせも可能としている．

ラジヘぇ[8] は，ラジオを聞きながらその音声に対して音声コメントをボタン 1 つで入れる事を可能にするものであり，コンテンツの他者との楽しみを強調するものである．本研究で提案する手法の一部は，ラジヘぇの音声付与に似ている．本提案手法は，編集を目的としており，ラジヘぇのようなリアルタイムでのコミュニケーションは対象としていない．

SMIL[6] はウェブ上でのマルチメディアコンテンツの振る舞いを記述するマークアップ言語である．SMIL では，同期的にマルチメディアコンテンツを再生するといった記述も可能であり，異なるフォーマットの動画ファイルを並べて再生することなども可能である．我々の手法は，SMIL の仕様とは異なり，ブラウザ自体がウェブ上のリソースおよびプレイヤを直接操作する事が可能となっているため，多様な表現が可能である．ただ，SMIL の記述スタイルは参考になる部分が多いため，今後編集情報の共有などにおいて参考にする予定である．

6 まとめ

本研究では，ウェブ上の動画に対して，ブラウザ上で音量コントロールや，スキップ操作，音声や音楽の付与といった編集を可能とし，その編集情報に基づき視聴可能な仕組みを実現した．また，編集情報を他者と共有可能な仕組みを実現し，他者の編集を参考に視聴したり，他者の編集を利用して再編集したりといった事を可能とした．

今後は，本プロトタイプシステムをウェブ上で公開し，多くのユーザに利用してもらうことでフィードバックを集める予定である．また，現在の操作インタフェースおよび機能は，まだ十分に練られていないため，ユーザからのフィードバックなどを元に改良を行っていく予定である．

ユーザによる編集情報と，ニコニコ動画のコメントにはある程度の相関がある可能性がある．そこで，ユーザの編集情報と，ニコニコ動画のコメントを相互分析することによって，自動編集を可能とする仕組みも実現する予定である．

参考文献

[1] 中村聡史，山本岳洋，後藤真孝．濱崎 雅弘: 視聴者反応と音楽的特徴量に基づくサムネイル動画の生成手法, 情報処理学会論文誌（トランザクション）データベース（TOD58）, Vol.6, No.3, pp.148-158 (2013-06-28).

[2] 濱崎雅弘，武田英明，西村拓一：動画共有サイトにおける大規模な協調的創造活動の創発のネットワーク分析-ニコニコ動画における初音ミク動画コミュニティを対象として-，人工知能学会論文誌, Vol. 25, No. 1, pp. 157-167 (2010).

[3] 濱野智史: アーキテクチャの生態系 — 情報環境はいかに設計されてきたか，NTT 出版（2008）．

[4] 佐藤剛, 宮下芳明: SeekRope: 曲げて切って結べるシークバー, インタラクション 2010 論文集, pp. 197-200 (2010).

[5] 徳 井 直 生: Massh!, http://www.sonasphere.com/mash/ (2008).

[6] W3C: Synchronized Multimedia, http://www.w3.org/AudioVideo/.

[7] mixbit: https://mixbit.com.

[8] 加藤由訓, 苗村健: ラジヘぇ：ラジオ聴取時における感想共有システム, インタラクション 2013, pp.32-39 (2013).

HirakuReader : 行間を拡張する電子書籍

中村 将達　　西田 健志*

概要.

本を読む時には，書かれている内容を鵜呑みにするのではなく批判的思考を持ち，関連する情報や異なる立場からの意見に触れるなどの学習を通じて広い視野を獲得する「能動的読書」を行うことが理想的だと言われているが，そのような読書習慣を身につけて持続させていくことは容易ではなく，誰しもができるものではない．こうした中，電子書籍は書き込みを読書間で共有できる機能や関連情報が素早く得られるという点から能動的読書を促す効果が期待されている．だが，電子書籍の書き込み共有機能が堅い意見交換の場となってしまうことで堅い場が苦手な人々が避けがちになってしまうという問題があり，能動的読書の習慣のない人々はなおさらそのような傾向に陥る可能性がある．そこで，本研究では娯楽性を考慮したデザインによって堅い意見交換が苦手な人にも本への気軽な書き込みを促し，本に対する積極的な態度を喚起する HirakuReader を開発した．HirakuReader は，行間を拡張するインターフェースを用いることで本の中に存在する行間をメディアとして機能させることができ，読者による投稿によって書籍コンテンツが変容していく電子書籍リーダーである．行間を活用した情報提示によって，読者の投稿を促し，関連情報への素早いアクセスを実現する．本論文では能動的読書を実現するアプローチと，そのアプローチに基づいた HirakuReader のデザインと実装について述べる．

1　はじめに

本を読む時には，本に書かれている内容を鵜呑みにするのではなく批判的思考を持ち，関連する本や情報，異なる立場からの意見に触れるなどの学習を通じて広い視野を獲得する「能動的読書」を行うことが理想的だと言われている [9]．しかし，そのような読書習慣の獲得が高等教育での目標のひとつとされていたり，読書法に関する啓発本が多数出版されていたりすることからもわかるように，そのような良い読書習慣を持つことは容易ではない．

それに対して能動的な読書を促す方法として注目されているのが，読んでいる本とともに様々な情報を提示することができる電子書籍である．たとえば，Kindle や Kobo[6] といった主要な電子書籍リーダーでは，書籍中の特定箇所のハイライトや，特定のフレーズに対するメモや感想を他の読者と共有することができるようになっている．また，本の感想を共有するための SNS サービスも数多く存在している [15][14].

しかし，そのような既存のシステムは，ユーザを「変なことを投稿してはいけない」という気持ちにさせるなど堅苦しさを感じさせている可能性がある．電子書籍へ注目が集まる一方で読書よりもテレビやインターネットの手軽なコンテンツを好みがちな若い世代の読書量の低下が問題視されている中[1]，電子書籍が堅い意見交換の場となってしまってはますます読書から遠ざかる人が出てしまう恐れがある．

我々は，コメントの重畳表示などに見られる利用者の交流を最優先するニコニコ動画 [13] の精神に電子書籍も倣うべきだと考える．交流最優先の精神が生み出した弾幕やコメント職人といった新たな遊びは堅苦しさを払拭することで多くの人々を巻き込み，その結果として政治のような堅い話題までもが語られる場を作り出すことに成功している．ニコニコ動画自身も，同じようなインタフェースで漫画や小説を楽しめるサービスを提供しているが [12]，文字が主体の情報にコメントを重畳するのはあまりに読みづらいため，電子書籍に対しては同様の精神を受け継ぎながらも異なるインタフェースをデザインする必要があると思われる．

その一試みとして本研究では，書籍中の行間を簡単なタッチジェスチャで広げて空間を作り出し，その空間を利用して関連情報を提示したり様々な情報を書き込んで共有したりできる電子書籍リーダー HirakuReader を開発した．特に，行間に書き込まれる内容や量が変化する様子を徐々に太くなっていく線や燃え上がるエフェクトなどで視覚化することで，コメントの内容うんぬんよりもとにかくコメントする楽しさ，他の読者と気楽にやり取りする楽しさを前面に出し，堅苦しさをなるべく感じずに交流できるように配慮している．読者の投稿によってインタラクティブに変化していく行間は文章そのものの可読性を落としうるが，読者がそのような破壊的

Copyright is held by the author(s).

* Masatatsu Nakamura，Takeshi Nishida，神戸大学大学院 国際文化学研究科

[1] 読書離れに関しては諸説あり [8][4]，読書離れが起きているとは一概には言えない．

な行動を行うことは本から受けがちな権威的な印象を和らげ能動的読書を促す効果があるものと期待される．本論文では，まずデザイン目標を議論し，その後に HirakuReader の実装および関連研究を紹介し，最後に今後の課題を述べる．

2 デザイン目標

本研究では行間を拡張するインターフェースのデザイン目標として以下のものを掲げている．

1. 本は与えられたまま読むものだという意識の破壊
2. 能動的な読者による書籍そのものの変化
3. 読書時の快適な情報量調節

2.1 本は与えられたまま読むものだという意識の破壊

電子書籍上での意見や注釈の共有機能は，付箋などといった紙の本のメタファーに則ったものとして実装されているものが多く，なるべく中身を汚さないようにするという紙の本の感覚が誘発され，本に対する受け身な態度を助長してしまう．

そこで我々は，読者の働きかけによってコンテンツが変容していき，書き込み行為自体が娯楽性を持つような本を実現することで読者の積極的な読書態度を促すことができるのではないかと考えた．

本研究では行間というスペースを活用することで，読者の働きかけが本のコンテンツを演出するような機能デザインを行った．HirakuReader では，本の中にある行間を「開く」ことで書き込みを行うことができ，そうして行われた書き込みは読者間で共有される．書き込みが行われた行間は光ったりアニメーションしたりすることで情報提示を行い，文章の見た目を変容させる．投稿によって変化する行間の表現性や娯楽性によってさらなる投稿を促し，より読者を積極的に本に介入させることを狙った．

2.2 能動的な読者による書籍そのものの変化

本には著作という概念が存在するため，第三者が内容を自由に書き換えることができないが，本研究では行間というスペースを用いて読者の投稿と情報提示を実現しているため，コンテンツそのものの書き換えは行なっておらず，著作物を破壊することなくコンテンツのさらなる充実を可能にしている．

本は時が立つにつれて書き換える必要がある部分などが出てくるが，そういった本は，概ね改訂版で内容が更新されたり，新版のまえがきなどに説明が付け足されたりする．電子書籍では，読者たちの投稿によって改訂版を待たずして内容の訂正などが可能になるが，そのためには読者たちに積極的な情報投稿を促す必要がある．本研究では書き込み行為に娯楽性を持たせることで，読者たちの積極的な態度を喚起しており，積極的な読者が増えることで内容が日に日に更新されていくような本のあり方を実現することを狙った．

2.3 読書時の快適な情報量調整

本を読んでいる際，元の文章だけでなく関連する情報を閲覧する場面がある．単語の意味，文章内に出てくる用語の詳細な説明といったものに限らず，他の読者たちの投稿した情報を参照することがあるだろう．その際，情報閲覧までにかかるステップが煩雑であるほど読書行為が阻害されてしまうため，なるべく素早く，読書の邪魔をしないような情報提示が求められる．

従来の電子書籍アプリケーションにおけるインターフェースでは，単語部分の長押しや，画面上に現れるアイコンをタッチするなどといった操作を行い，ポップアップウィンドウや画面端の領域に情報を提示するという手法が採用されている．しかし，そういった情報提示が行われる領域は文章表示の邪魔をしないために一定のサイズが保たれており，読者がもっと情報を得たいと感じた時には画面遷移を伴う必要がある．

HirakuReader のインターフェースでは，二本指で単語や行間を挟み込むことで選択し，そのまま指を広げていくことで文章のレイアウトを変化させて空間を作りだし，そこに様々な情報提示を行うことで関連情報へのアクセスを可能にしている．選択から検索までをスムーズに行うことができ，もっと情報を見たいと感じた際にはそのまま指を広げていくことでさらに多くの情報提示することができるため，閲覧する情報の量を快適に調整することができる．

3 HirakuReader

本章ではタブレット端末向けのアプリケーションとして実装した HirakuReader の機能について説明を行う．

3.1 行間を広げるジェスチャインターフェース

行間を広げる操作は，二本の指で画面にタッチしている状態からそのまま指を広げるという動作で実行できる．このインターフェースは読書中でも素早く行うことができ，各種情報へのアクセスを素早く，容易にしている．行間を広げると文書のレイアウトが上下に動いていき，ユーザが指を広げれば広げるほどそれに応じて行間が広がっていく．このインターフェースでは，二本指で画面に触れた際の指の真ん中の位置がポインティング位置となっており，二本指の真ん中で単語や行を狙うことで選択を行う（図 1）．



図 1. 行間や単語の選択の際，二本指でタッチしている位置の真ん中がポインティング位置となっている．

3.2 行間を広げて関連情報を検索する機能

文章内の単語は二本指で挟み込むことで選択することができ，そのままピンチアウトすることでその単語が含まれる行を挟んで上下の行間を広げられる（図 2）．広げられた行間にはその単語に関する情報が提示され，辞書や各種 Web サービスでのその単語による検索結果を表示することができる．単語に関する情報は指をどれほど広げるかによって提示される内容が変化し，少し広げると辞書，もう少し広げると画像検索の結果や Wikipedia のページなどを閲覧することができ，指を広げるほど多くの情報を得ることができる．

3.3 広げた行間に情報を追加し，共有する機能

文章内に存在する行間も単語同様に二本指で挟み込むことで選択することができ，選択された行間を挟んで上下に広げていくことができる．誤って単語の選択が行われてしまわないよう，行間の選択にはページ横の余白のスペースを利用し，その余白の中で任意の行間を二本指の真ん中で狙うという操作を行うことで選択を行えるようにした（図 3）．

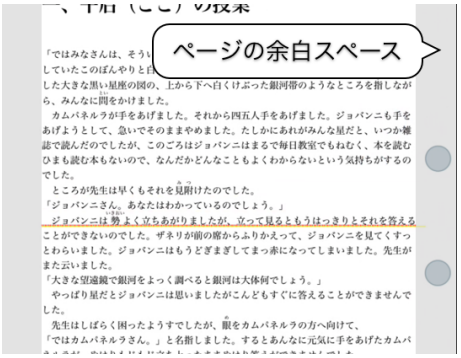


図 3. ページ横の余白領域（グレーの色がかかった領域）の中で任意の行間を二本指の真ん中で挟み込むことで行間の選択を行う．

この広げられた行間には，意見や内容の補足などといった情報を画像や動画などを交えて投稿することができる．こうして投稿された情報はアプリケーション間で共有され，読者の手によって行間に含まれる情報は充実していく（図 4）．投稿の際には文字のサイズや色，表示される際の位置などを指定することができ（図 5），提示される情報に表現の幅を持たせた．



図 5. 投稿用のフォーム．文字の大きさや位置，色を設定することができ，画像も投稿できる．

3.4 行間へ追加された情報を視覚化する機能

読者によって情報が投稿されている行間は，閉じられている状態では線が引かれ，光る・アニメーションするなどの表現を行い，情報が埋め込まれていることを知らせる．投稿が少ない状態では，行間は細い点線で表示され，投稿が増えていくとしだいに線が太く点線が細かくなっていき，最終的には太い直線になる（図 6）．

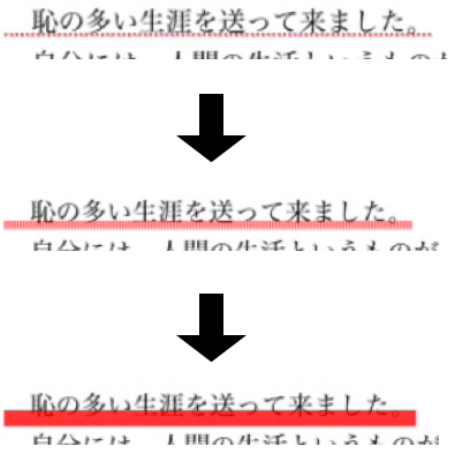


図 6. 行間は投稿が少ない状態では細い点線で表され，投稿が増えるほど点線が細かく，太くなってゆき，行間が強調される

読者が設定できる投稿の情報には，文字のサイズや色，位置などがあり，投稿された書き込みの色に応じて行間の色が変化するなど，設定が直接反映されるような表現を実現している．行間の表現には燃え上がるなどの演出機能も用意しており，こうした強力な表現は本を破壊的に読むメタファーとして機能する（図 7）．

これらの機能によって読者の働きかけでインタラクティブに変化する行間を実現し，投稿する行為に

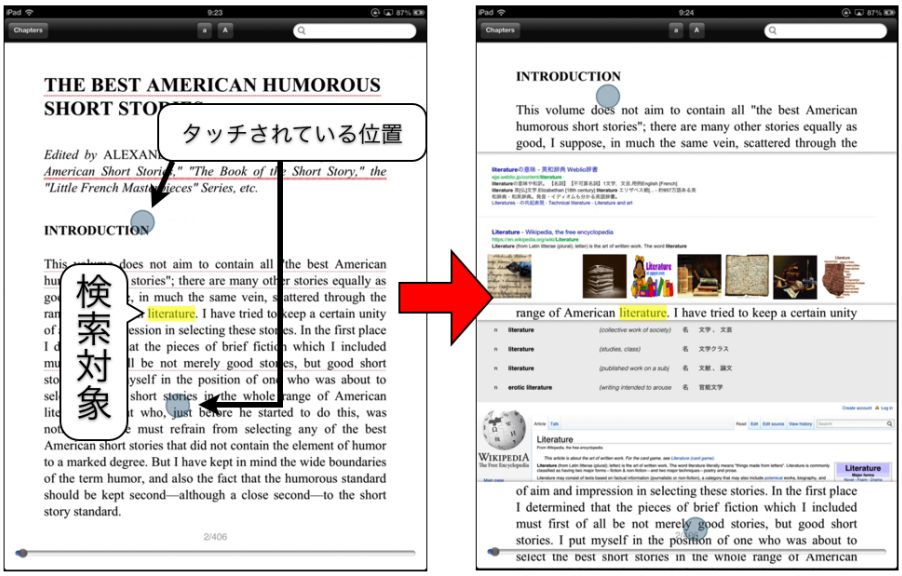


図 2. 文章中の単語が選択され，検索されている様子．文章内の単語は二本指で挟み込むことで選択することができ，選択中の単語は黄色くハイライトすることでわかりやすくした．

娯楽性を持たせることでより積極的に本と関わろうとする読者の態度を誘発することを狙った．

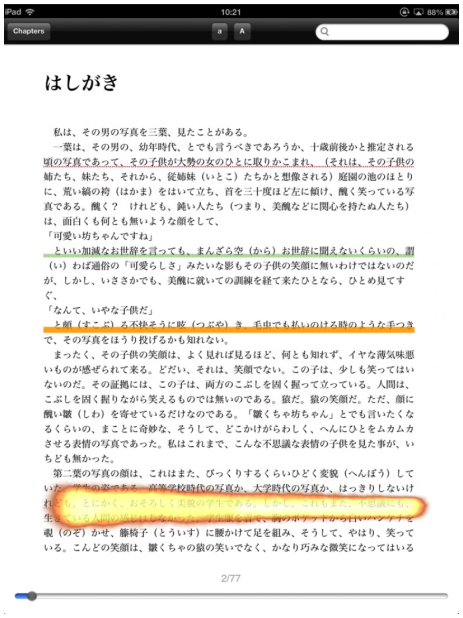


図 7. 情報が投稿された行間の表現．投稿の量や設定によって表現は変化し，行間が燃え上がるようなエフェクトもある．

3.5 実装

本システムはサーバークライアントシステムとして実装を行った．クライアント側では，iPad タブレットを利用し iOS アプリケーションとして実装した．電子書籍のフォーマットには HTML ベースの標準的な EPUB3 を対象としており，CSS や JavaScript を用いて書籍のデザイン調整などを行った．行間の表現などのダイナミックな部分は HTML を用いずに，iOS の API を用いて実現している．サーバ側は Ruby を用いて実装を行っており，データベースには MySQL を用いた．

3.6 ユーザフィードバック

本システムのプロトタイプを同じ研究室の学生に使用してもらい，出てきた声をまとめる．

- 難しい本や退屈な本でも面白い書き込みがあることで読み進めようという気持ちになった
- 一度読んだ本でも，自分の後から読んだ人の書き込みが行われることでまた読みたくなった
- 単語を選択して素早く検索できるため，本を読み進めるのが早くなった，またそれに伴って英語の本が読みたくなった
- 単語を検索して出てくる画像などが面白く，本を読み進める上で参考になる場面があった
- 画面上下の端において，行間を広げる際に狙った箇所とは違う行間を広げてしまうことがあった
- 文字部分を長押しすることで選択する従来の文字選択手法に比べて，二本指で挟む手法は選択部分が指で隠れずに済み素早く選択する

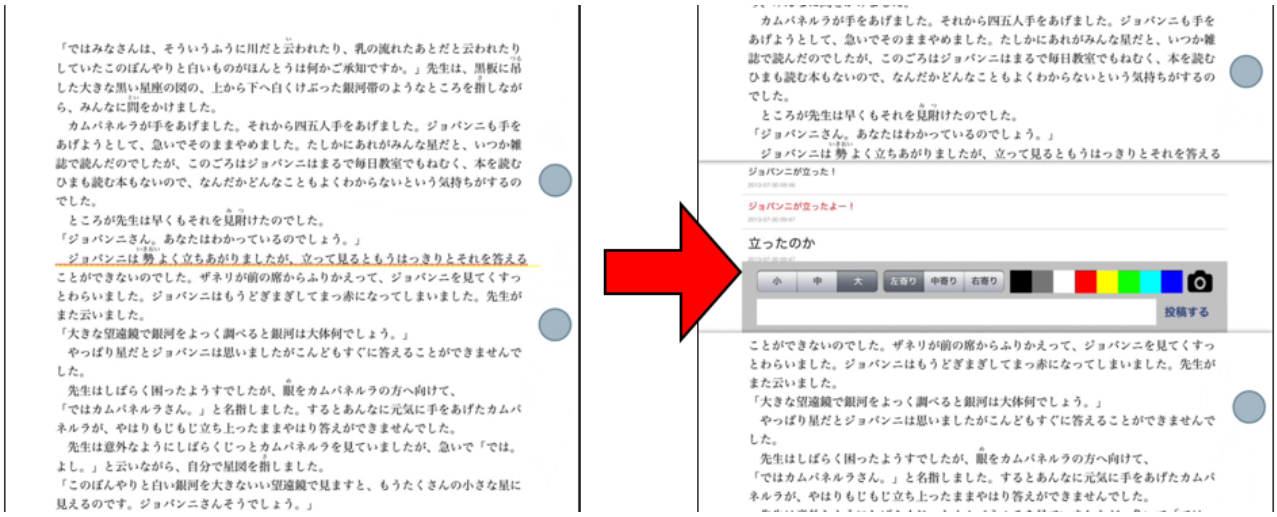


図 4. 情報が投稿されている行間を選択して，情報を閲覧している様子．文章中の行間も単語同様二本指で挟み込んで選択することができ，選択中の行間は黄色くハイライトしている．

ことができたが，文字の細かい選択範囲が操作できないことに不便を感じた

面白い書き込みがあることで読み進めようという気持ちが促進されることや，検索の素早さから関連情報を得やすくなり本の理解度が増すなどの声がある一方で，行間を広げる操作を間違えてしまうなどの声もあり，インターフェースに改善の余地があることを示唆している．

4 今後の課題

本システムは，肝となる行間の表現や，文字選択機能や行間に書き込まれた情報の共有機能に関して課題がある．

4.1 行間の表現

読者の投稿によって変化する行間の表現は娯楽性を引き出す上で重要であるが，どのような情報が投稿されるとどのような行間の演出が現れるかといった点については未だ検討中であり，これからの課題となっている．例えば，燃え上がるなどの表現（図 7）は強力な演出となりうるが，いかにして読者の投稿と絡めて効果的に用いるかは検討していく必要がある．

4.2 投稿の共有

電子書籍はデバイスの解像度や画面サイズ，設定された文字サイズによって行の長さが変化してしまうため，特定の行間に書き込んだとしても，他のデバイスや文字サイズで閲覧する際には書き込みが表示される位置が変化してしまう問題がある．

また，行間にしか書き込みを行えない現状のシステムでは文章中の特定の単語や節に対してアノテーションしたいというニーズには応えることはできない．書き込み箇所を曖昧にしておくほうが気軽な書き込みを誘発できると我々は考えているが，アノテーションとの兼ね合いを考慮した書き込み箇所の指定方法についても考察する必要があるだろう．

4.3 文字選択

行間を広げるインターフェースは，二本指の真ん中で単語や行間を狙って選択を行うが，文字範囲の選択に関して一つの課題がある．英語などの元々分かち書きされている言語を対象とする場合には，一つの単語を容易にポインティングすることができるが，日本語などの分かち書きされていない言語の場合は困難が生じる．

5 関連研究

Schilit らの [9] は紙の本で行われる注釈やマーキングといった書き込みをデジタルの文書でも行えるようにした．ペンによる自由な形式での注釈やマーキングをつけながらの読書は，文章の理解度を促進する効果があるが [18]，デジタル文書は紙の文書と違ってレイアウトが柔軟に切り替わるためにそうした情報は管理しにくく [3]，また大量の読者による情報共有は実現しづらいため，本研究ではフォームを用いた情報の投稿を採用した．

電子書籍上での注釈情報の共有は Johnson らによって教育上の効果があることが確認されており [5]，電子書籍のこうした機能には教育ツールとしての可能性が期待されている [7]．

このようなユーザ間での注釈情報の共有は，書籍よりも Web ページを対象としたものが積極的に試みられている．その中には Web ページ上に重畳表

示する手法がいくつかあり，ペンデバイスを用いた手書きのもの [17] や付箋メタファーを用いたもの [2] など，本に対して行われる書き込みを模倣した手法が多く，電子書籍での注釈共有を考える上で参考になる．

ニコニコ動画 [13] の動画を見えにくくするようなコメントシステムは，ユーザによるコメント投稿に娯楽性を持たせており，コメントそのものが動画コンテンツの一部と化している．同じ動画共有サイトの YouTube[11] とニコニコ動画では，ユーザの行動パターンに顕著な差が見られ [16]，ユーザの働きかけを積極的に引き出すシステムによって齎されるコンテンツの受容態度の変化が見て取れる．ゲームなどの娯楽性のあるシステムを用いることでユーザから効果的に情報を収集する研究 [10] もあり，娯楽性や表現性といったものがユーザの積極性を引き出すという効果には期待ができる．本研究の電子書籍システムも，行間への書き込み行為が娯楽性を持つような機能やデザインを試みている．

本研究では，行間を広げることで情報提示を行ったが，文章のレイアウトを柔軟に変化させることで効果的な情報提示を行った研究として Chang らの [1] がある．本の脚注などの情報を見るためには，読んでいる部分から離れた位置まで視線を移動させる必要があり，読書の負担となる．Chang らは本文中に情報提示を行う空間を作りだすことでそうした問題を解決している．我々の手法では指を広げるという単一の動作で，レイアウトを変化させて情報提示空間を作り出す機能と情報提示空間を広げていく機能を実現しており，情報への素早いアクセスが可能となっている．

6 まとめ

本研究ではなるべく多くの人が能動的読書を達成するための試みとして，娯楽性によって読者を引き込んで本への積極的な態度を促すというアプローチを提案し，電子書籍リーダー HirakuReader を開発した．HirakuReader では行間をメディアとして活用することで読者による本への書き込みに娯楽性を持たせ，関連情報へ素早くアクセスできるようにした．今後は正式なユーザスタディを行い，娯楽性が読書態度をどれほど改善するのかといった点や，行間を広げるインターフェースの効果の調査等を行いたい．

電子書籍の浸透がきっかけとなり，今後の本のあり方に関して様々な議論が交わされているが，本研究が本のあるべき姿に一つの方向性を示すことを期待している．

参考文献

[1] B.-W. Chang, J. D. Mackinlay, P. T. Zellweger, and T. Igarashi. A negotiation architecture for fluid documents. In *Proc. ACM '98*, UIST '98, pp. 123–132, New York, NY, USA, 1998. ACM.

[2] diigo. http://www.diigo.com.

[3] G. Golovchinsky and L. Denoue. Moving markup: repositioning freeform annotations. In *Proc. ACM '02*, UIST '02, pp. 21–30, New York, NY, USA, 2002. ACM.

[4] 日本人の読書離れは進んでいるのか. http://research.goo.ne.jp/database/data/000618/.

[5] T. E. Johnson, T. N. Archibald, and G. Tenenbaum. Individual and team annotation effects on students' reading comprehension, critical thinking, and meta-cognitive skills. *Computers in Human Behavior*, 26(6):1496 – 1507, 2010.

[6] Kobo. http://kobo.rakuten.co.jp.

[7] D. G. Lebow and D. W. Lick. HyLighter: An effective interactive annotation innovation for distance education. In *20th Annual Conference on Distance Teaching and Learning*, pp. 1–5, 2005.

[8] 若者だけじゃない、各年代で進む読書離れ 30・40 代女性の半数は 1 か月の本代 1,000 円未満. http://moneyzine.jp/article/detail/206052.

[9] B. N. Schilit, G. Golovchinsky, and M. N. Price. Beyond paper: supporting active reading with free form digital ink annotations. In *Proc. SIGCHI'98*, CHI '98, pp. 249–256, New York, NY, USA, 1998. ACM Press/Addison-Wesley Publishing Co.

[10] L. von Ahn, R. Liu, and M. Blum. Peekaboom: a game for locating objects in images. In *Proc. SIGCHI '06*, CHI '06, pp. 55–64, New York, NY, USA, 2006. ACM.

[11] YouTube. http://youtube.com.

[12] ニコニコ静画. http://seiga.nicovideo.jp.

[13] ニコニコ動画. http://nicovideo.jp.

[14] ブクレコ. http://bookreco.jp.

[15] ブクログ. http://booklog.jp.

[16] 沖松健太郎, 松本淳. 動画共有サイトにおけるユーザー参加型のオープン・クリエーションモデル– YouTube とニコニコ動画の比較分析, 日本社会情報学会 第 23 回全国大会. 2008.

[17] 浜口拡輝, 加藤直樹, 山崎謙介. Web 上への手書きメモが共有可能なブラウザ PerowserEx の開発. 情報処理学会研究報告, 2009, 2009.

[18] 浩成 野崎, 彩奈 吉橋, 恭子 梅田, 撤郎 江島. テキストへの自由な書き込み行為が文章理解に及ぼす影響. 日本教育工学会論文誌, 29:49–52, mar 2006.

SuprIME: IMEによるテキスト編集機能の統合

中園 翔　　増井 俊之 *

概要．計算機上で様々なテキストエディタが利用されているが，システムごとに編集方法や編集機能が異なっているのが不便である．本論文では，各国語入力のためにOSに用意されているIME(Input Method Editor)機能を利用することにより，あらゆるテキストエディタにおいて同じ操作によるテキスト編集を可能にする方法を提案する．我々の手法を利用すると，異なるテキストエディタ上での編集操作が共通化されるだけでなく，テキスト編集時に便利な様々な機能をあらゆるエディタで利用することが可能になる．

1　はじめに

計算機上でテキストを編集するために様々なテキストエディタが利用されている．文書を作成するときはワープロを利用し，メールを書くにはメールクライアントを利用し，文字端末でのプログラム開発にはvimやEmacsを利用し，IDEを利用した開発では付属のエディタを利用し，ネット上でテキストを扱うにはブラウザのテキストフォームを利用するといったように，場合に応じて様々なエディタが利用されている．

エディタの機能や操作体系はエディタごとに異なっているのが普通である．ブラウザやワープロでテキストを1行消したい場合はマウスで行全体を選択してから削除キーを押せばよいが，vimでは「d」キーを2回タイプして消すのが普通であり，EmacsではCtrl-Kキーが利用される．Emacsに慣れたユーザがワープロ上でもCtrl-Kで行を消去したいと思っても，そのようなカスタマイズはできないのが普通であるし，機能拡張が可能なシステムを利用している場合でも，操作体系を完全に同じにすることは難しい．

あらゆるエディタの操作を統一することは難しいが，様々なエディタで共通に利用できるソフトウェア層をユーザとアプリケーションの間に置くことができれば，異なるエディタの編集操作をある程度共通化できる可能性がある．現在のパソコンにはIME(Input Method Editor)と呼ばれる文字入力機構が用意されており，様々な言語のテキスト入力に利用されている．IMEはエディタなどとは独立したソフトウェアであり，ユーザのすべてのキー入力を受け取って各国語に変換した結果をアプリケーションに送出する．IMEはあらゆるアプリケーションで共通に利用されるので，たとえば日本語入力用のIMEを利用する場合，Emacsでもブラウザでも IDEでも同じ操作で日本語を入力できる．IMEは一般には各国語入力のみのために利用されているが，テキストの挿入/移動/削除といった編集操作もIMEが受け持つようにすれば，様々なエディタ上で同じ操作で編集を行なうことが可能になると考えられる．

このような考えにもとづき，Mac上の様々なテキストエディタにおいて同じキー操作によるテキスト編集を可能にするSuprIMEシステムを試作した．本論文ではSuprIMEの実装と利用方法について述べ，柔軟でユニバーサルなテキストエディタの構築について考察する．

2　SuprIME使用例

2.1　日本語入力

SuprIMEは，MacRuby[1]で記述された日本語IMEである「Gyaim[2]」に様々な編集機能を追加したものである．図1はSuprIMEによる日本語入力の例である．IMEはアルファベット以外の文字を入力するときだけ有効にするのが普通であるが，SuprIMEは，編集操作を行うために常に有効であることを想定して実装した．そのため、英語入力と日本語入力の両方が可能である．

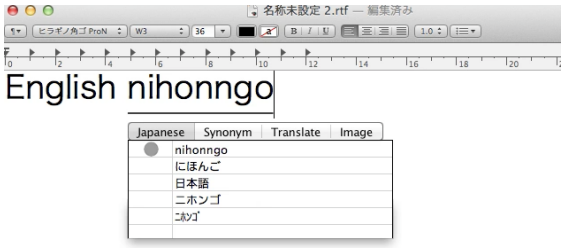


図 1. SuprIMEによる日本語入力の例

Copyright is held by the author(s).

* 慶應義塾大学環境情報学部 環境情報学科，慶應義塾大学環境情報学部

[1] Rubyを Mac用に拡張したもので，あらゆるMacのAPIをRubyから利用することができる．

[2] https://github.com/masui/Gyaim

2.2 ブロック移動

テキストの一部を別の場所に移動する操作はテキストエディタの重要な機能のひとつであるが，エディタによって操作方法が大きく異なっている．たとえばEmacsでテキストを移動させたい場合は，移動する領域をキー操作によって指定してから削除/コピーし，カーソルを移動してからペーストするという手順を利用するのに対し，ブラウザの編集領域でテキストを移動させたい場合は，マウスで領域を指定した後で選択領域をドラッグして別の位置に移動することが多い．このように，テキスト移動のような基本操作でもエディタごとに操作が異なっているのは不便であり，操作ミスをしがちであるが，SuprIMEを利用するとあらゆるエディタにおいて同じ操作でテキストを移動することができる．

図2はMac OSに標準登載されている「テキストエディット」でテキストを編集しているところである．

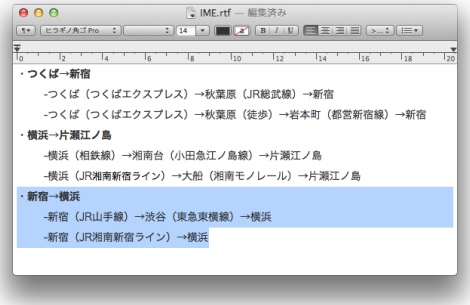


図 2. ブロック移動前の状態

ここでShift+↑キーを押すとテキストは図3のように変化する．

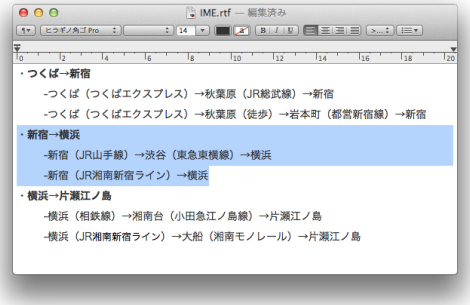


図 3. Shift+↑キーを押した後の状態

図4はブラウザ上でGoogle Docsのテキストを編集しているところである．ここでShift+↓キーを二度押すと，テキストは図5のように変化する．

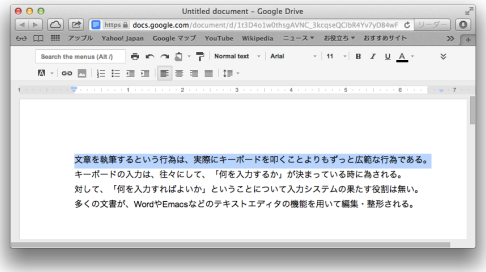


図 4. ブロック移動前の状態

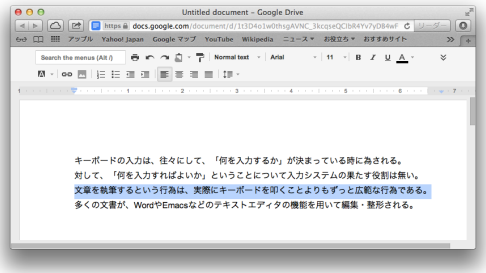


図 5. Shift+↑キーを押した後の状態

このように，あらゆるエディタにおいてShift+矢印キーという共通の操作でブロック移動を行なうことができる．

2.3 連続インデント

前述したブロック移動は，インデントを調整する機能も備えている．プログラミングを行なっている時だけでなく，普通のテキストを編集している場合でも行頭の空白やタブの量を調整するインデント処理は頻繁に行われるが，自動的にインデントを行うエディタもあれば，コマンドを打ち込まなければならないものもあり，そのような調整機能を持っていないエディタも多い．SuprIMEを利用すると，あらゆるエディタや入力フィールドでブロック移動と同時にインデントが自動で行われる．

図6は，Xcode上でPythonプログラムを書いている例である．Xcodeは，Objective-CやRubyなどの言語で自動インデントの機能を備えているが，Pythonには対応していない．たとえば，図6のようなテキストに対して図7のように移動させたいブロックを選択し，Shift+↓キーを入力することで，テキストは図8のように変化する．

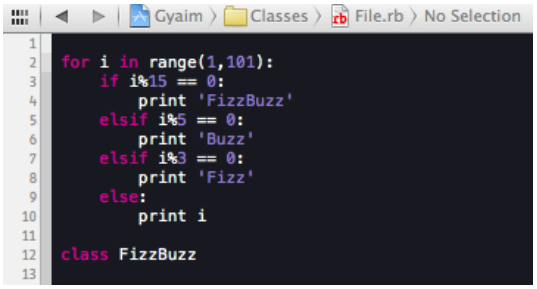


図 6. ブロック移動前の状態

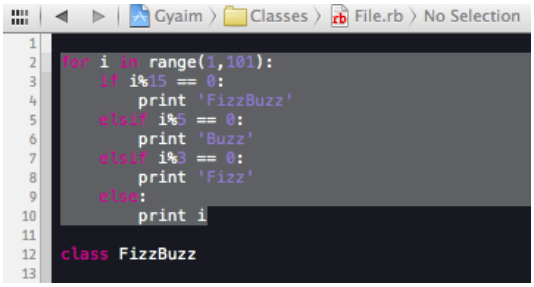


図 7. テキストを選択した状態

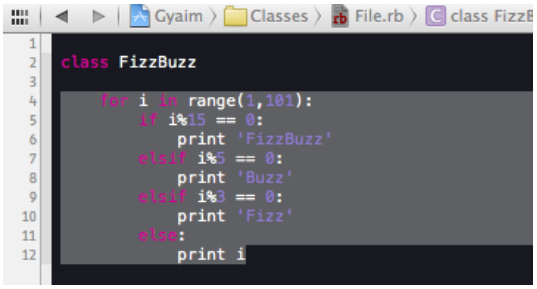


図 8. ブロック移動後の状態

このように，あらゆるエディタにおいて，ブロック移動を行う際に適切な空白とタブを自動で補完することができる．

2.4 単語置換

多くのエディタが単語の検索/置換機能を持っているが，その操作方法はシステムごとに異なっている．例えば，コンソール上で検索と置換を行う時は grep や sed コマンドを用いるが，多くのエディタソフトでは Control+F や Command+F が用意されている．

SuprIME を利用することでこの操作を共通化する例を挙げる．

あらかじめ割り当てたファンクションキーを押すと，図 9 のように SuprIME の単語置換ウィンドウが起動する．



図 9. 単語置換ウィンドウの例

図 10 のように，入力語と置換語をテキストエリアに打ち込み，Replace ボタンを押すと，

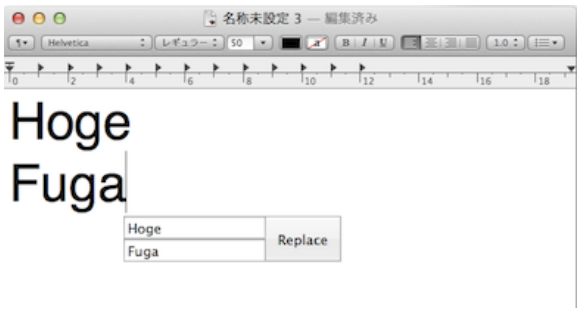


図 10. 置換ウィンドウに単語を入力した例

テキストは図 11 のように置換される．

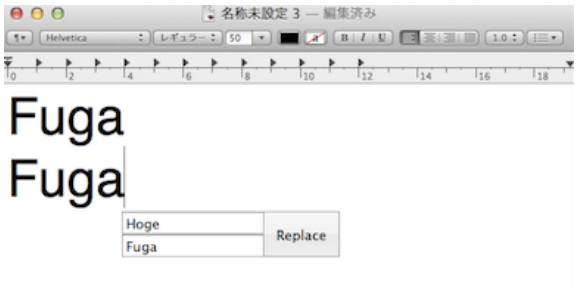


図 11. 単語置換後の状態

このように，あらゆるエディタ上で検索と置換が可能である．

2.5 Dynamic Macro

Emacs では lisp プログラムでエディタの機能を拡張することが出来るが，これらのスクリプトを IME 上に実装することで，他のエディタ上でも Emacs と同様の拡張機能を動作させることができる．

以下に SuprIME で Dynamic Macro[7] を実現した例を挙げる．Dynamic Macro は，入力の繰り返しを自動化する Emacs 拡張であるが，SuprIME ではこれと同様の機能を Ruby で実装している．図 12 のように，ユーザの入力操作が繰り返しになってい

るとき，任意に設定したファンクションキーを押すと，テキストは図 13 のように編集される．

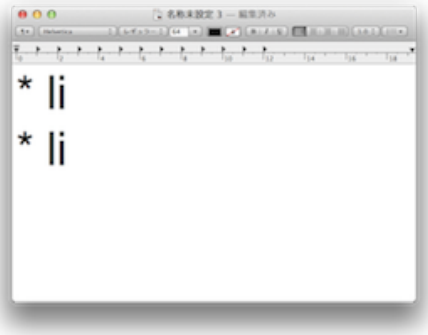


図 12. Dynamic Macro 機能を使用する前のテキスト

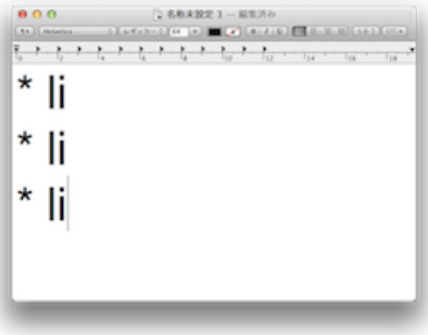


図 13. Dynamic Macro 機能を使用した後のテキスト

以下は SuprIME の Dynamic Macro 機能をブラウザのアドレスバー上で実行した例である．図 14 のように，ユーザの入力に”abc”が繰り返されているとき，任意に設定したファンクションキーを押すと，テキストは図 15 のように編集される．



図 14. Dynamic Macro 機能を使用する前のテキスト

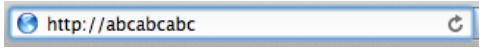


図 15. Dynamic Macro 機能を使用した後のテキスト

Emacs 上に実装された Dynamic Macro は Emacs 上でしか利用できないが，SuprIME 上に実装された Dynamic Macro はあらゆるテキストエディタの上で利用することができる．

3 実装

SuprIME は MacRuby で記述された Mac 用の IME であり，ソースが 500 行程度とコンパクトであるにもかかわらず，他の IME に見られない機能を実装しており，本論文のような実験も容易である．

3.1 MacRuby による実装

MacRuby は，Mac OS 用のアプリケーションを開発するために拡張された Ruby 実行環境であり，Mac OS の Objective-C ライブラリを Ruby で扱うことが出来る．SuprIME では，Mac OS の IME フレームワークである InputMethodKit Framework を MacRuby から呼び出すことによって基本的な IME の機能を実装している．

3.2 修飾キーの扱い

エディタが編集機能のために用いる，Control や Command などの修飾キーは IME がハンドルしないのが普通であるが，SuprIME では、一部のファンクションキーと修飾キーの入力を受け取り，共通化されたテキスト編集操作を実現している．

SuprIME のように，修飾キーの一部をハンドルする実装を行った場合，修飾キーの入力はエディタに送出されないため，エディタが同じキー入力を編集操作の一部に割り当てている場合，それを利用することは出来なくなる．

3.3 テキストデータの取得

ブロック移動やインデントの処理といったテキスト編集の機能を，あらゆるテキストエリアで行うためには，テキストエリアに入力されている全文を IME が取得する必要があり，InputMethodKit はこの機能を備えていない．SuprIME では，ブロック移動やインデントのコマンドが入力された際に，Mac OS に実装されている，テキストエリアの全文を選択状態にするコマンドを AppleScript により送信している．

Mac OS では，多くのアプリケーションにおいてテキストエリアは NSTextField Class のオブジェクトとして実装されており，この NSTextField に対してテキスト編集や入力の操作を行うための NSTextInput Protocol が存在する．このプロトコルには標準の実装として Command+A キーによる全文選択の機能が備わっている．

SuprIME では，このキー入力を AppleScript によりテキストエリアに送信することにより，テキストエリアの全文を選択状態にしている．SuprIME では，NSTextInput の標準実装である Command+C キーによる全文コピーの機能を AppleScript により送信することで，Mac OS のクリップボードにテキストエリアの全文を保存し，それを InputMethodKit

Framework が読み込むことで，テキストエリアの全文を IME が取得している．

3.4 実装手法の問題点

しかし，このような手法で実装できる編集機能や，IME としてのパフォーマンスには問題がある．たとえば，対象とするテキストエリア自体が NSTextField Class を利用していない場合や，対象とするテキストエリアがテキストの全選択やコピーに割り当てるキーバインドを変更している場合，テキストエリアの全文を取得することが出来ない場合があり，SuprIME の持つ編集機能は有効に機能しない．また、どこまでの編集機能をエディタやテキストエリアから切り離し，どれだけの拡張性を持たせることができるか，という点において，IME の層での実装には問題がある．各国語入力のための枠組みである IME は，テキストエリアよりも先にユーザのキー入力をハンドルすることができるが，エディタが想定するキー入力を横取りして、エディタの機能をを打ち消す実装になる可能性がある．前述した Dynamic Macro の例のように、編集操作をマクロ化して行うことは多くあるので，IME とテキストエリアとの間に，テキスト編集のための枠組みを用意し，機能拡張が容易に可能な OS と API 設計のレベルでの実装の修正が求められる．

4 議論

4.1 キー操作の統一の必要性

Xerox PARC の Larry Tesler が 70 年代に発明した [10] 「コピー/ペースト」は現在広く普及しており，ほぼすべてのアプリケーションにおいて同じ操作 [3] でテキストをコピーしたりペーストしたりできるようになっている．つまりコピー/ペーストに関しては 1 章で述べたような問題がほぼ存在せず，ユーザは混乱せずに様々なアプリケーション上でテキスト操作を行なうことが可能になっていることになる．コピー/ペーストの操作がこのように標準化されているのに対し，テキスト移動のような処理が標準化されていないのは問題であるが，それが大きな問題だと認識されておらず，ユーザが様々なエディタの使い方を覚えることをシステムが強制していることはさらに大きな問題だといえるだろう．SuprIME のような手法を利用することによってシステム全体を統一的に利用できるようにする工夫はまだまだ重要だと考えられる．

4.2 ユニバーサルな入力/編集

パソコンやスマートフォンで利用されている様々なテキスト入力システムは変換方式も使い勝手も全く異なるのが普通になっているが，POBox[6] のような単純で柔軟な入力方式を利用すると，パソコンでもスマートフォンでもほぼ共通の入力を行なうことが可能になる．Gyaim はこのような思想にもとづいて作成された IME であるが，Gyaim を拡張した SuprIME を利用すると編集操作も共通化することができるため，広い範囲の機器における入力と編集の共通化が期待できる．

4.3 高度な文書編集補助

SuprIME を利用すると，単純な文書編集作業が共通化できるだけではなく，より高度な編集補助を行なうことも可能である．たとえば，単語の言い換えを補助するために，図 16 のように類語を検索して入力することが出来る．

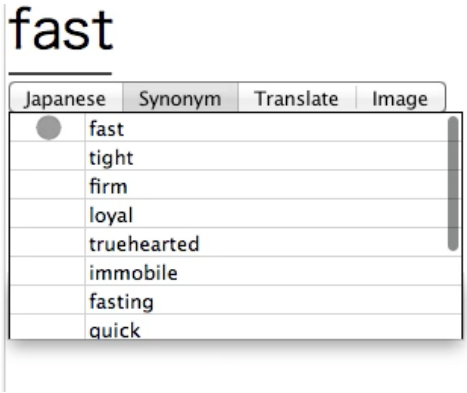


図 **16. SuprIME** の類語変換機能の例

図 17,18 は SuprIME の画像入力の例である．ここでは、「知らない」という入力語句について画像を検索し，入力候補として表示している．画像が埋め込めるテキストエリアであれば画像を埋め込み，そうでなければ URL が入力される．

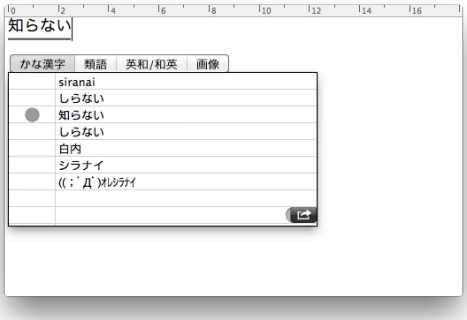


図 **17. SuprIME** で「知らない」という単語を入力した例

[3] Mac の場合は Command-C と Command-V



図 18. SuprIME の画像入力ウィンドウ

SuprIME を利用すると，このような高度なテキスト編者操作もあらゆるアプリケーションで共通に利用できる．

4.4 テキスト入力/編集以外への応用

本論文ではテキストエディタに絞った説明を行なったが，IME はユーザのすべてのキー入力を直接受け取る窓口になっているため，キー操作をテキスト編集と関係無い仕事に割り当てることも可能である．たとえばシステム音量や画面の明るさなども SuprIME から制御することが可能であり，システムで用意されたショートカット設定などが不要になる．

4.5 本手法の限界

IME は，テキストエリアが実装されたアプリケーションとは個別に実装されているため，現在の SuprIME の実装ではアプリケーションの内部状態によって動作を変えたり，アプリケーションの振る舞いを制御したりすることはできず，表に出ているテキストの編集操作しかできない．

既存のシステム自体は変更せずに，皮をかぶせる形で機能を拡張する手法はある程度有用ではあるが，問題の根本的な解決が必要な場合には限界がある．テキスト編集の場合は根本的に解決しなければならない問題は多くないので，本論文の手法はとりあえず有効だといえるが，根本的な解決のためには，テキスト入力の枠組みである IME に加えて，各コンピュータがテキスト編集操作を一元化するための枠組みを用意する必要があるだろう．

5 結論

OS に標準装備された IME 機能を活用することにより，様々なテキストエディタにおける入力/編集操作を共通化する手法を提案した．IME 機能はアプリケーションから独立しているため，アプリケーション内のデータを操作することはできないが，様々なアプリケーションにおける入力/編集操作の多くの部分を共通化することができた．テキストの入力や編集は計算機利用における最も重要な仕事のひとつであることは間違いないので，操作を簡単化/共通化するための優れた枠組みの開発は重要な課題だと考えられる．

参考文献

[1] TextEditors.org. http://texteditors.org/.

[2] P. H. Dietz, B. Eidelson, J. Westhues, and S. Bathiche. A practical pressure sensitive computer keyboard. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, UIST '09, pp. 55–58, 2009.

[3] C. Harrison, D. Tan, and D. Morris. Skinput: appropriating the body as an input surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pp. 453–462, 2010.

[4] F. C. Y. Li, R. T. Guy, K. Yatani, and K. N. Truong. The 1line keyboard: a QWERTY layout in a single line. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, UIST '11, pp. 461–470, 2011.

[5] I. S. MacKenzie, R. W. Soukoreff, and J. Helga. 1 thumb, 4 buttons, 20 words per minute: design and evaluation of H4-writer. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, UIST '11, pp. 471–480, 2011.

[6] T. Masui. An efficient text input method for pen-based computers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '98, pp. 328–335. ACM Press/Addison-Wesley Publishing Co., 1998.

[7] T. Masui and K. Nakayama. Repeat and predict – two keys to efficient text editing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '94, pp. 118–123, 1994.

[8] T. Murase, A. Moteki, N. Ozawa, N. Hara, T. Nakai, and K. Fujimoto. Gesture keyboard requiring only one camera. In *Proceedings of the 24th annual ACM symposium adjunct on User interface software and technology*, UIST '11 Adjunct, pp. 9–10, 2011.

[9] J. Rick. Performance optimizations of virtual keyboards for stroke-based text entry on a touch-based tabletop. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology*, UIST '10, pp. 77–86, 2010.

[10] L. Tesler. A personal history of modeless text editing and cut/copy-paste. *interactions*, 19(4):70–75, July 2012.

[11] D. Wigdor and R. Balakrishnan. TiltText: using tilt for text input to mobile phones. In *Proceedings of the 16th annual ACM symposium on User interface software and technology*, UIST '03, pp. 81–90, 2003.

CapStudio: プログラムの実行画面の録画映像を利用した統合開発環境

深堀 孔明　　坂本 大介　　加藤 淳　　五十嵐 健夫 *

概要. 通常，プログラム開発は，コードエディタ上でプログラムのソースコードを編集することによって行われる．しかし，ゲーム開発においては，キャラクタ画像やボタンの配置など見た目に関わる部分はゲームの実行画面を見ながらリアルタイムに編集できた方が直感的である．ゲームエンジンのシーンエディタといった，プレビュー画面上でパラメータを編集できるシステムもあるが，フレームワーク上で決められた制約があり，ゲーム画面の描画などのプリミティブな処理をプログラマが編集できず，コーディングの自由度が低くなるという問題がある．そこで，本稿ではゲーム画面の録画映像をベースとしたエディタ（以下スクリーンキャスト）を提案する．これをコードエディタと連携させることで，プログラマはプログラム全体をコードエディタで記述しつつも，ゲーム画面のデザインに関わる定数値やリソースファイルをスクリーンキャスト上でリアルタイムに編集できる．具体的には，定数値やリソースが編集されると即座にその結果がスクリーンキャストに反映される．反対に，スクリーンキャストのゲームオブジェクトから対応するコード片やリソースファイルへアクセスしたり，ゲームオブジェクトの位置やサイズを直接操作してソースコード中の定数値を調整できる．本稿では，スクリーンキャストを組み込んだ統合開発環境 CapStudio を実装し，本手法の可用性を検証した．

1　はじめに

ゲーム開発など，ビジュアルな要素を多く使用するプログラミング開発においては，キャラクタ画像やボタンの配置といった，見た目に関わる部分の調整は非常に重要である．キャラクタ画像は配色を微妙に変えるだけで大きく印象が変わるし，ボタンなどの GUI 部品の配置方法によってゲームの操作性は大きく変化する．しかし従来のコードベースの開発環境においては，このようなビジュアル要素の開発を支援する枠組みはほとんどなかった．このため，たとえばプログラムの編集中にはゲーム画面は表示されないため，キャラクタ画像やボタンの表示位置を修正するたびにプログラム全体をビルド・実行してゲーム画面を表示し，見栄えを確認する必要があった．

一方，Unity [9] などの近代的なゲームエンジンでは，プログラムの編集中にゲームのプレビュー画面が常に表示される．プログラマはパラメータの修正結果をリアルタイムに確認したり，画面上のゲームオブジェクトを直接動かすことで位置パラメータを調節することができる．しかし，これらのシステムにはフレームワーク上で決められた制約があり，ゲーム画面の描画といったプリミティブな処理をプログラマが編集することができず，コードベースの開発環境に比べてコーディングの自由度が下がってしまうという問題がある．

本稿では，従来のコードベースの開発環境に対してゲームエンジンのような対話可能なプレビュー画面を提供することを目的とする．そのためのシステムとして，対話可能なゲーム画面の録画映像であるスクリーンキャストを組み込んだ統合開発環境 CapStudio を提案する (図 1)．プログラムの実行中，CapStudio はゲーム画面をバックグラウンドで録画する．実行後，録画映像はスクリーンキャストとして表示される．スクリーンキャストは一般的な動画プレイヤと同様に，映像を再生したり，再生時刻をシークバーで変更できる．

ただし，実際にシステムが記録する情報は，ゲームオブジェクトの描画位置や大きさといった描画履歴であり，システムはこれを用いて録画映像を生成している．また，この描画履歴を用いることで，スクリーンキャストとコードエディタを協調して動作させることができる．具体的には，プログラマがソースコードや画像などのリソースファイルを編集する

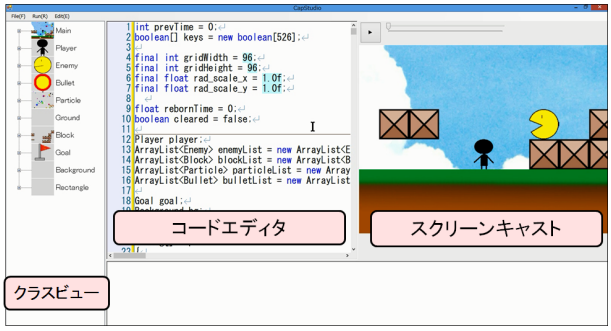


図 1. CapStudio のインタフェース.

Copyright is held by the author(s).

* Koumei Fukahori, Daisuke Sakamoto, Jun Kato, Takeo Igarashi, 東京大学

と，即座に編集結果がスクリーンキャストに反映される．反対に，スクリーンキャスト上のゲームオブジェクトを選択することで，対応するコード片やリソースファイルへアクセスできる．さらに，ゲームオブジェクトの位置や大きさを直接操作して，ソースコード中の対応する定数値を書きかえることができる．本稿では作成したシステムについて述べ，その実装を説明したあと，3つのサンプルプログラムを用いて本システムの可用性を議論する．

2 関連研究

本研究は 2012 年の Bret Victor によるデモビデオ“Inventing on Principle”[1] に強く触発されている．このビデオでは，本システムのようにプログラムの実行画面を開発環境に埋め込み，コードエディタと連携させる手法が紹介されている．ただし，これはあくまでコンセプトビデオであり，システムの具体的な実装については言及されていない．一方で，本稿では，このような実行画面とコードエディタの連携機能を，スクリーンキャストを用いることで実現し，その具体的な実装を示している．

実行画面の録画映像をベースとしたインタフェースを持つシステムとして，Whyline for Java [5] というデバッグツールがある．Whyline は実行時に記録した実行履歴を解析し，プログラムの挙動の原因となるコード片を特定する．たとえば，プログラマが録画映像上のゲームオブジェクトを選択することで，そのオブジェクトの色や位置を決定したコード位置へジャンプすることができる．しかし，Whyline は本システムとは違い，ソースコードの編集を目的としたツールではない．そのため，コードの編集結果を即座に録画映像に反映したり，映像内のゲームオブジェクトを動かしてソースコード中のパラメータを書きかえることはできない．

プログラムを明示的に再実行せずにソースコードの修正結果を計算する手法として，バックグラウンドでプログラム全体を再ビルド・再実行して再計算を行う方法がある [3, 4]．しかしこの方法は，外部プロセスと通信を行うプログラムを開発する場合，暗黙的な再ビルド・再実行によって，意図せずに外部プロセスの状態を変えてしまう恐れがある．一方，本稿で提案する手法では必要な情報をすべて実行時に記録し，記録した情報のみを用いて再計算を行うため，予期しない外部プロセス・リソースの改変は起きない．また, プログラムの編集状態と実行状態を区別しない Live Programming 言語 [2, 6, 8] も，コード編集の結果を再実行せずに実行画面に反映することができる．しかし，上記の手法と同様に、プログラムの編集中に予期せず外部プロセスに影響をおよぼす恐れがある。

また，本システムともっとも関連のあるシステムとして，Unity [9] や Unreal Engine [10] といったゲームエンジンのシーンエディタや Visual Studio [11] などの GUI エディタがある．これらの開発環境では，プログラムの編集中，実行画面のプレビューが常に表示され，パラメータの修正結果をリアルタイムで反映したり，プレビュー画面上のゲームオブジェクトを直接動かして位置パラメータを調節できる．しかし，これらのシステムにはフレームワーク上で決められた制約があり，ゲーム画面の描画といったプリミティブな処理をプログラマが編集できないため，コードベースの開発環境に比べてコーディングの自由度が下がってしまうという問題がある．また, プレビュー画面内の時間の流れは不可逆で，本システムのスクリーンキャストと違い，シークバーを動かして時間を巻き戻すことはできない．

3 提案システム CapStudio

CapStudio の実行画面を図 1 に示す．プログラマはまずコードエディタにソースコードを記述し，その後，スクリーンキャストを利用してゲーム画面のデザインを修正する．

提案手法の最大のポイントは，ゲームエンジンのシーンエディタのようなパラメータ調整用のエディタを，コードベースの開発環境に対して提供することにある．一般に，描画処理を含めたすべてのソースコードがプログラマによって記述される場合，プログラムの挙動を静的に予測してプレビュー画面を生成することはできない．そこで本研究では，実行時に出力されたゲーム画面を擬似的なプレビュー画面に見立てる．まず，システムはゲーム画面の描画履歴を実行時に記録し，これを利用してゲーム画面の録画映像をスクリーンキャストとして生成する．さらに，記録した描画履歴を利用してスクリーンキャストとコード・リソース間の連携機能を実現する．

3.1 スクリーンキャストの生成

プログラマはまず，コードエディタに Processing 言語 [7] でソースコードを記述する．CapStudio のコードエディタは公式の開発環境 Processing IDE（PDE）と互換性がある．PDE で実行できるコードは CapStudio でも実行でき，その逆も真である．メニューバーの実行ボタンを押すと，プログラムがビルド・実行されゲーム画面が表示される．

プログラムの実行時，システムはバックグラウンドでゲーム画面の描画に関わった関数 (以下，描画関数) の呼び出し履歴を記録する．たとえば，図 2 のようにゲーム画面上に画像がひとつだけ表示されるプログラムでは，ゲーム画面の大きさを指定する *size()*, 画像データをファイルから読み込む *loadImage()*, 画面をクリアする *background()*, 画像を描画する *image()* のいずれかが呼び出されるたびに，そのコード位置，返り値，仮引数の式に含まれる変数の値といった関数呼び出し情報を記録する．これによ

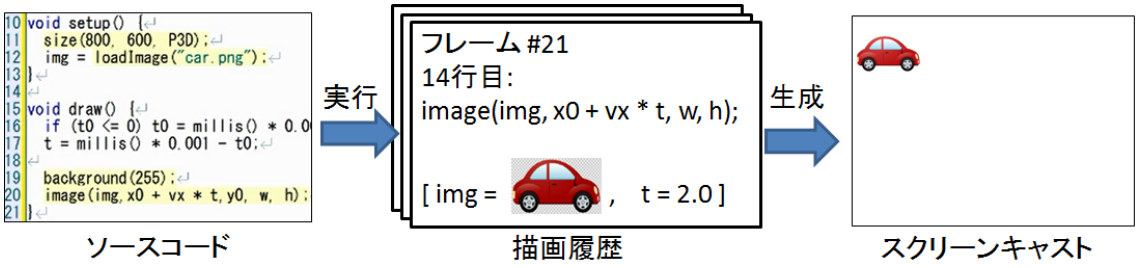


図 2. スクリーンキャストの生成

り，描画されたゲームオブジェクトの位置や大きさ，およびそれを描画したコードが特定できる．ただし，ゲーム画面の描画には Processig が公式に提供している API 関数だけが使われると仮定している．つまり，ソースコード中のどの関数が描画に関わっていて，どの引数が位置や大きさを指定するのかという API の仕様はシステムがあらかじめわかっている．もし OpenGL や Java Swing など別の描画 API を使う場合は，各 API 関数の仕様を別途システムに組み込む必要がある．

プログラムの実行が終了すると，システムは記録された描画履歴から，ゲーム画面の録画映像がスクリーンキャストとして生成される (図3)．スクリーンキャストは動画プレイヤのように，再生ボタン（▷）を押して録画映像を再生したり，シークバーを動かして再生時刻を変更できる．

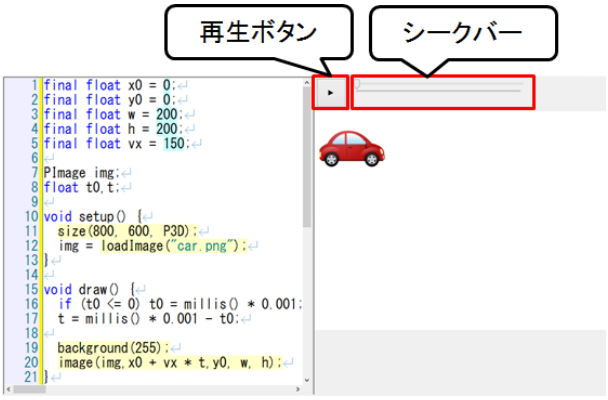


図 3. CapStudio で開発したプログラムの例．車の画像が水平方向に等速移動する．

3.2 コード・リソース編集結果のスクリーンキャストへの即時反映

プログラムの実行後，システムは，ゲーム画面のデザインに関わるソースコードをハイライトする．ゲーム画面のデザインに関わるソースコードとは，本稿では，1) 描画関数呼び出し文の引数，2) 定数宣言式の右辺，の 2 つを指す．たとえば，図 3 のコードでは，*loadImage()*，*image()* といった描画関数の引数と，x0，y0 といった定数の値がハイライトされる．プログラマがこれらのいずれかを書きかえると，システムは再度ソースコードの構文解析を行い，新しい引数式・定数値および前節で記録した描画履歴にしたがってスクリーンキャストを再描画する (図 4-a)．同様に，プログラマが画像などのリソースファイルを編集すると，システムは編集されたリソースを読み込み直して，スクリーンキャストを再描画する (図 4-b)．

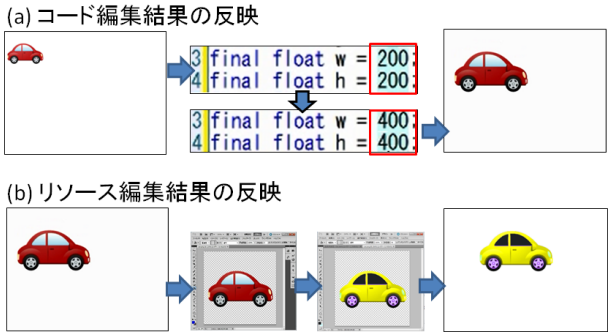


図 4. コード・リソース編集結果のスクリーンキャストへの反映．

3.3 スクリーンキャストを用いたコード・リソースの編集

スクリーンキャスト内のゲームオブジェクトはすべてクリッカブルであり，クリックすることで対応するコードやリソースへアクセスできる．左クリックすると，それを描画した関数呼び出しが赤くハイライトされ，そこへカーソルが移動する (図 5-a)．右クリックすると，対応する画像ファイルを開くためのメニューが表示され，選択した画像ファイルが外部の画像編集ソフトで開かれる (図 5-b)．前述したように，これらのコードや画像ファイルが編集されると，即座に結果がスクリーンキャストに反映される．

また，スクリーンキャストからコード内の定数値を調節できる．スクリーンキャスト上のすべてのゲームオブジェクトは位置や大きさをドラッグ操作で変更できる (図 6-a)．こうしてゲームオブジェクトの位置・

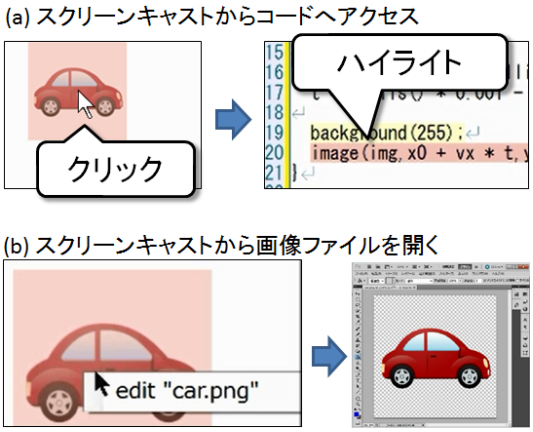


図 5. スクリーンキャストからコード・リソースへのアクセス

大きさを変更すると，同様の結果になるようにソースコードが自動で書きかえられる (図 6-b)．デフォルトでは対応する関数呼び出し文の引数に定数値が加えられる．書きかえられた引数を右クリックすると，別の書きかえパターンが列挙される (図 6-c)．列挙された各パターンは，引数の式に登場する定数を書きかえ対象とする．たとえば，*image(img,x0+vx*t,y,w,h)* という関数呼び出しによって描画された画像を，左端の X 座標が 150 になるように移動したとする．ただし，x0，vx, t の値はそれぞれ 50，25, 2 で x0, vx が定数である．このとき，まず第二引数に目的値 (150) と現在の値 (100) との差分 (50) が加算され，*rect(img,x0+vx*t+50,y,w,h)* と書きかえられる．その後，式 *x0+vx*t+50* が右クリックされると，定数 x0, vx を書きかえ対象としたパターンがそれぞれ「*final float x0 = 100*」「*final float vx = 50*」と表示される．たとえば前者のパターンを選択すると，第二引数が書きかえ前の式 *x0+vx*t* に戻り，x0 の宣言式が *final float x0 = 100* に書きかえられる．

ただし，引数に現れる定数の内，次の 2 つの条件を満たすものが書きかえ対象となる．1) 描画関数の引数のみに影響を与える．2) 定数がどんな値になるべきか計算可能．たとえば if 文の条件式に使われる定数は書きかえ対象にはならない．また，現在の実装では，引数式が定数に関する線形式の形になっている場合のみ目標値を計算できる．つまり，引数の式が a * c1 + b / c2 で c1，c2 が定数だとすると，c1 のみが書きかえ対象となる．

4 実装

CapStudio は Processing 2.0 beta 8 の開発環境である．インタフェース部分は C#の WinForm アプリケーションとして実装されており，Windows 7 および Windows 8 で動作している．以下，前章で

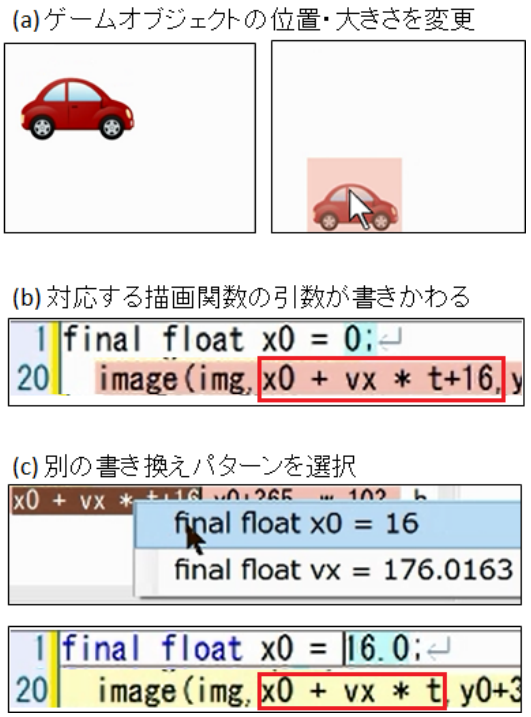


図 6. スクリーンキャストを用いたコード編集

紹介した機能について，実装アルゴリズムを述べる．

4.1 描画履歴の記録・スクリーンキャストの生成

実行時に描画履歴を記録するため，プログラムがビルドされる前に，描画関数の呼び出し文に対してコード書きかえが行われる．ソースコード中の描画関数呼び出し文の前後に，その呼び出し文のコード位置，呼び出し時の返り値および引数の式に登場する変数の値を記録するコードが自動挿入される．

記録された描画履歴を用いて，現在シークバーが示している時点（フレーム）のゲーム画面を再現する．まず，プログラムの実行が終わると，システムは実行時に描画関数によって読み込まれたリソースファイルをすべて読み込む．次に，コードエディタに記述されたソースコードを構文解析し，定数の値と描画関数の引数式を抽出する．そして，実行時に描画関数が呼び出されたときの引数値を，引数式に定数値と実行時に記録された変数値を代入することで求める．その後システムは，読み込んだリソースと求めた引数値を使って，目的のフレームの間に呼び出された描画関数を順に再実行することで，スクリーンキャストを描画する．

4.2 コード・リソース編集結果のスクリーンキャストへの即時反映

プログラマがコードやリソースファイルを書きかえると，スクリーンキャストは実行時に読み込まれたリソースファイルをすべて読み込み直す．その後，

ソースコードの構文解析を再実行し，定数の値と，描画関数の引数式を更新する．その後，新しい定数値・引数式を用いて，各時点における描画関数の引数値を再計算する．その後，現在のフレームで呼び出された描画関数を順に実行しなおして，スクリーンキャストを再描画する．

4.3 スクリーンキャストを用いたコード・リソースの編集

プログラマがスクリーンキャストをクリックすると，現在のフレームで発行された描画関数の呼び出しをすべて走査し，マウスカーソルと接触しているゲームオブジェクトを探す．見つかったら，それを描画した関数呼び出し文をハイライトする．あるいは引数にリソースオブジェクト (PImage 型の画像オブエジェクトなど) が与えられている場合は，その元ファイルのファイルパスを表示する．

同様に，プログラマがスクリーンキャスト上のゲームオブジェクトをドラッグすると，まずすべての描画関数呼び出しを走査して，ドラッグされたゲームオブジェクトを探す．次に，マウスの移動量に応じてゲームオブジェクトの新しい座標・大きさを求める．その後，その位置・大きさで描画されるように対応する描画関数の引数の値を求め，現在の引数の値との差分を引数式に加算する．同時に，別の書きかえパターンについて，書きかえ対象とする定数の目標値を，その定数に関する一次方程式を解くことで求める．

4.4 スクリーンキャスト生成手法の応用

２章で解説したコード・リソース編集のための機能に加えて，スクリーンキャストの生成手法を応用して，次の３つの機能を実装している (図 7)．1) あるクラスのメンバ関数の中で呼び出された描画関数のみを選択して再実行することで，そのクラスのサムネイル画像を生成する．2) 他のフレームで描画されるゲームオブジェクトを現在のフレームに重ねて描画することで，ゲームオブジェクトの軌跡を残像として表示する．さらに，3) ゲームオブジェクトの描画位置をずらすことで，視点を二次元的に移動させ，画面外に描画されていたゲームオブジェクトを表示させることができる．

5 サンプルプログラム

作成したシステムの可用性を検討するため３つのサンプルプログラムを作成した (図 8)．１つ目の例ではゲーム画面のレイアウト，２つ目の例ではリソース編集における，スクリーンキャストの有用性を議論する．３つ目の例では，実行時に外部プロセス・リソースと通信するプログラムの開発において，本システムが有効な性質を持つことを示す．

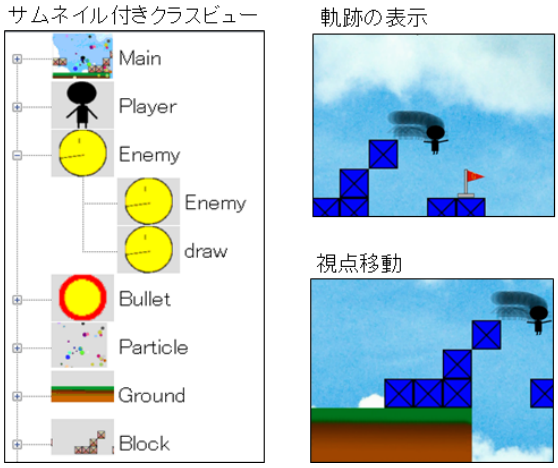


図 7. スクリーンキャスト生成手法の応用

5.1 ゲームのタイトル画面

図 8-a はゲームのタイトル画面に模した GUI プログラムである．タイトル・メニュー項目がそれぞれラベル・ボタンとして表示され，それぞれの位置と大きさはソースコード中で定数として指定されている．これらのオブジェクトを一度実行時に表示させてしまえば，あとはスクリーンキャストでオブジェクトをドラッグして位置・大きさを調節できる．もちろん，既存 GUI エディタやゲームエンジンでも，このように GUI 部品のレイアウトを直接操作することはできる．しかし本研究では，ソースコードがすべてプログラマによって書かれたプログラムに対して同様の操作を可能にしたことが重要であると考えている．

5.2 2D アクションゲーム

図 8-b は簡単な 2D 横スクロールアクションゲームである．ユーザはキャラクタを操作し，道中の敵を弾丸で倒しながらゴールを目指す．キャラクタや敵といったゲームオブジェクトはすべて，外部ファイルから読み込まれた画像を使って描画される．これらの画像を編集するためには，通常はファイルマネージャから画像を保存しているフォルダへ移動し，目的の画像を開く必要がある．一方，本システムでは，スクリーンキャスト上のゲームオブジェクトから直接目的の画像を開くことができ，ユーザにとっての負荷が大きく軽減されている．特に，弾丸のように一瞬しか現れないゲームオブジェクトは，再生時間を巻き戻せない映像内で捉えることは難しく，ゲームエンジンのシーンエディタを利用した場合でも，ゲームオブジェクトから画像を直接開くことは難しかった．しかし本システムでは，シークバーを動かして再生時刻を調整することで，そのようなオブジェクトもスクリーンキャスト上で簡単に捉えることができる．



図 8. サンプルアプリケーション.

5.3 ネットワーククライアント

図 8-c はネットワーク通信を行う簡単なメディアアート作品である．Kinect から取得された骨格データを用いて，人間の手からパーティクルを放出する．ただし，Kinect の公式 SDK は Processing に対応していないので，骨格データは外部の C#プロセスに取得させ，TCP 通信で Processing プロセスへ送信している．プログラムの実行を終了すると，C#プロセスとのネットワーク接続が遮断され，Processing プロセスは新たな人間の骨格データを取得できなくなる．このとき，関連研究の章で紹介したバックグラウンドでプログラム全体を再ビルド·再実行する手法では，プログラマがソースコードを編集したときに正しく編集結果を計算できない．一方本システムは，実行時に骨格データの情報が描画履歴として記録されており，これを使って計算を行うため，コードの編集結果がスクリーンキャストに正しく反映される．

6 まとめと今後の課題

本稿では，対話可能なプログラムの録画映像であるスクリーンキャストを組み込んだゲーム開発のための統合開発環境 CapStudio を提案した．スクリーンキャストは従来のコードエディタと協調して動作し，プログラマは全てのプログラムコードをコードエディタで記述しつつも，ゲーム画面のデザインをスクリーンキャスト上で直感的に修正できる．

本提案の問題としては，スクリーンキャストがソースコードの特定箇所の変更結果しか反映できないことが挙げられる．現状の実現方法では，描画関数の呼び出し文と定数の定義式の変化のみトラッキングしている．したがって，ゲームオブジェクト同士の衝突判定といったロジックに関わるコードを書きかえても，スクリーンキャストには反映されない．本システムはあくまでゲームオブジェクトの配置や色合いといった，画面デザインの修正を支援することを目的としている．

現時点では，Processing の描画 API の内，特に現在は，画像読み込み·画像描画といった，2D ゲームの開発に必要な描画関数のみに対応している．したがって今後は，3D モデルの描画を含めた Processing の描画 API のフルセットへの対応を予定している．それと並行して，スクリーンキャスト上のゲームオブジェクトを三次元的に操作するためのインタフェースも実装する．さらに，OpenGL や Java Swing といった，より低レベルな描画 API への対応も検討している．

参考文献

[1] Bret Victor - Inventing on Principle. http://vimeo.com/36579366.

[2] J. Edwards. Subtext: Uncovering the Simplicity of Programming. In *Proc. OOPSLA '05*, pp. 505–518, 2005.

[3] B. Hartmann, L. Yu, A. Allison, Y. Yang, and S. R. Klemmer. Design as exploration: creating interface alternatives through parallel authoring and runtime tuning. In *Proc. UIST '08*, pp. 91–100, 2008.

[4] J. Kato, S. McDirmid, and X. Cao. DejaVu: integrated support for developing interactive camera-based programs. In *Proc. UIST '12*, pp. 189–196, 2012.

[5] A. J. Ko and B. A. Myers. Debugging reinvented: asking and answering why and why not questions about program behavior. In *Proc. ICSE '08*, pp. 301–310, 2008.

[6] S. McDirmid. Living it up with a live programming language. In *Proc. OOPSLA '07*, pp. 623–638, 2007.

[7] Processsing. http://processing.org.

[8] Welcome to Self - Self - the power of simplicity. http://selflanguage.org/.

[9] Unity3D Game Engine. http://unity3d.com.

[10] Game Engine Technology by Unreal. http://www.unrealengine.com/.

[11] Microsoft Visual Studio. http://msdn.microsoft.com/ja-jp/vstudio.

順解析と逆解析を相互に利用する打楽器のインタラクティブなデザインインタフェース

山本 和彦　　五十嵐 健夫

概要.

本稿では, 打楽器の形状と音色のデザインを効率的に行うために, コンピュータによって物体形状からそれを叩いたときの音色を計算する順解析と音色からそれを再現可能な物体形状を推定する逆解析を交互に行い, その支援を行うインタフェースを提案する. 本インタフェースでは, ユーザが物体形状を編集した場合には音色を, 音色自体を編集した場合には対応する物体形状をそれぞれインタラクティブにユーザにフィードバックする. 従来の打楽器を始めとする生楽器の設計では形状データの作成, 解析による確認, 出力による評価の 3 段階のプロセス間の非常に煩雑なトライ&エラーを繰り返す. これを効率化するために, 形状モデリングと同時にその音を解析して, 結果をインタラクティブにフィードバックするインタフェースも過去に提案されているが, 提示された情報を基にユーザはさらにどのような操作を加えれば目的とするものに近づけていくことができるのか, ということを判断するのは困難であるという問題がある. 本システムでは提示された音色情報自体をさらにユーザが編集してその影響を物体形状に反映させることを繰り返すことによってこの問題を解決する. 本稿で提案するような順問題と逆問題を相互に解いてデザインの支援を行うアプローチは楽器のみならず多くの一般的な製品の設計においても役立つ可能性がある.

1　はじめに

現在, 打楽器を含む多くの生楽器の設計プロセスは, 形状モデリング, 解析, 出力/評価の 3 段階から成り立っている. このプロセスでは, まず最初に CAD で形状データを作成する. その後, それを有限要素法等で解析してその音の情報 (モード周波数, 振幅) や実際にそれをものとして出力しても破綻が無いかなどを確認, 最後にそれを実際に物体として出力して今度はその実際の音や演奏感を評価する. しかしながらこの 3 つのプロセスはそれぞれが乖離しており, プロセス間を横断したトライ&エラーは非常に労力がかかる. 例えば, 解析の結果が気に入らなかった場合には再び形状を修正し直してから再解析を行うということを繰り返す. また, 解析プロセスを一度パスできたとしても, その後に出力して実際の出音が所望のものでなかった場合には再び形状モデリングと解析のプロセスを繰り返すことになってしまう. これは非常に煩雑な作業であり, 設計者はプロセス間を横断したトライ&エラーに労力が割かれてしまい, 本来の創造的なデザインに集中することが困難になってしまっているという問題がある.

この煩雑さを解消するために, 梅谷ら [1] は形状モデリングと同時にそれを叩いたときの音の基本周波数の情報を音としてユーザにフィードバックすることによって形状モデリングと解析のプロセスを統合した. しかし, これだけではフィードバックされた情報を基にユーザは, さらにどのような形状変形操作を加えれば目的とする音に近づけていくことができるのか, を判断することが困難であるというさらなる問題がある.

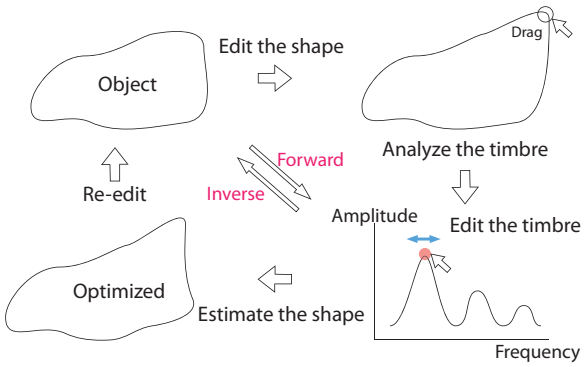


図 1. 順解析と逆解析を繰り返して形状と音色両面から打楽器をデザインする

そこで, 本稿では, 打楽器の形状と音色のデザインを効率的に行うために, コンピュータによって物体形状からそれを叩いたときの音色を計算する順解析と音色からそれを再現可能な物体形状を推定する逆解析を相互に行い, その支援を行うインタフェースを提案する.　本システムではユーザが物体形状を編集した場合には音色を, 音色自体を編集した場合には対応する物体形状をそれぞれインタラクティブにユーザにフィードバックする. これを使ってユーザは, 形状編集操作に応じて提示された音色情報自

Copyright is held by the author(s).
Kazuhiko Yamamoto, Takeo Igarashi, 東京大学

体をさらに編集してその影響を物体形状に再び反映させることを繰り返すことによって, 形状と音色両面から打楽器をデザインしていくことが可能である (図 1).

本稿で提案するような順解析と逆解析を交互に解いてデザインの支援を行うというアプローチは, ユーザが解析手法独特のノウハウ等の余計な問題に煩わせられることなく常にその時々で着目しているデザイン要素にのみ集中することができるようになるという利点がある. これは音だけではなく例えば剛性や熱伝導性能といった問題にも置き換えることで, 楽器のみならず多くの一般的な製品の設計においても役立つ可能性がある.

2 関連研究

コンピュータによって物体の振動音を計算するための最も代表的な手法は, 物体を有限要素法で離散化して固有値解析を行い, 算出されたモード (固有振動) の線形和としてその振動を近似するものである [3]. 固有値解析は非常に計算コストが高く一般的には前計算としてのみ行われる処理であり, インタラクティブな速度で計算を行いデザインツールに利用するためには高速化のアルゴリズムが必要である.

梅谷ら [1] は求める固有値を最低次の固有値のみに限定することで, ユーザが物体形状をモデリングすると同時にインタラクティブな速度で固有値解析を行い, その結果をフィードバック可能にした. この研究ではユーザが鉄琴の形状モデリングを行うと同時にその基本周波数の情報を音としてフィードバックすることによって, 複雑な形状の鉄琴であっても容易にその音高を調整することができる. しかし, 提示される情報が基本周波数のみであり, 楽器としての評価を行うためには不十分であった. C.B.Maxwell ら [4] は固有関数が既知である物体形状からの微小な変形による影響をモード空間に射影して自由度縮小を行うことで準インタラクティブな速度で基本周波数だけではなく高次の周波数成分も計算しており, 音色の評価をすることも可能である. しかし, 元の形状からの変形が大きくなると計算が破綻してしまうという問題がある.

コンピュータによるインタラクティブな物理シミュレーションをデザインツールに利用する研究は近年盛んに行われているが [6][13], そのほとんどが順解析か逆解析のどちらかを単独で利用するものである. 順解析とは, 原因を与えてそこから結果を計算する問題であり, 物体の運動シミュレーション等多くの問題がこれに属する. その逆が逆解析であり, 結果から原因を推定する問題である. 本研究で扱う音色からの物体形状の推定もこの逆問題に属する.

Design by Dragging[5] は順解析と逆解析両方を同時にデザインに利用できるように一つの環境に統合した最初の研究である. この研究では医療機器の

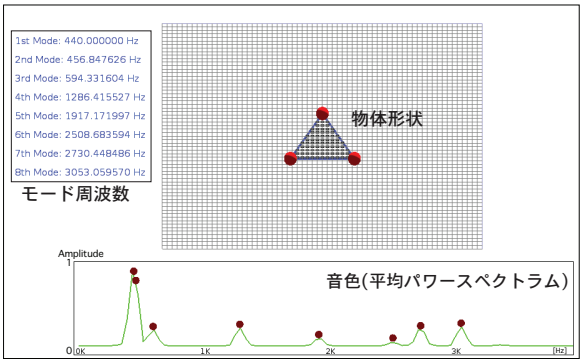


図 2. プロトタイプ画面

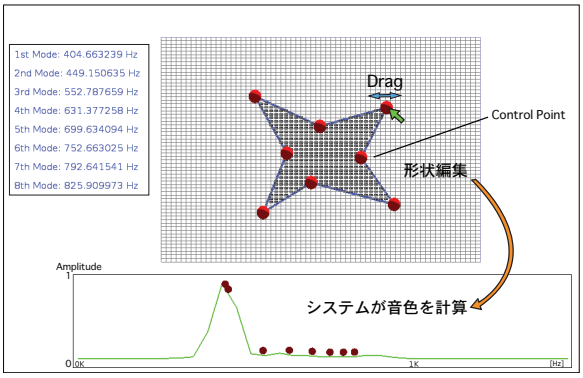


図 3. マウスによる物体形状モデリング

形状を予め細かくパラメータ化しておき, 前計算したシミュレーション結果の中からユーザが現在編集している状況に近いものを検索, 補間することで, ユーザが物体形状を編集した際には構造への負荷が, 結果として表示された構造への負荷状況を編集した場合にはそれによる形状への影響がインタラクティブにフィードバックされる. これによってデザイナーは解析ノウハウを一切気にするこの無く本来のデザイン作業に集中できる. しかし, 予め細かくパラメータ化された形状は編集の自由度に乏しく, デザイナーはツール制作者の想定を超えるような形状を作ることができないという問題がある.

3 ユーザインタフェース

図 2 が本研究で提案するインタフェースのプロトタイプ画面である. これによって打楽器の形状と音色を編集する. 中央に編集する打楽器の形状が表示してあり, 左側に現在の形状でのモード周波数を低周波数から順に表示, 下側には実際に音色を鳴らしたときの音のパワースペクトラムの時間方向の平均値およびその編集画面が用意されている. パワースペクトラムは横軸に周波数 (リニアスケール), 縦軸に正規化されたパワーを示している. 横軸は一番低い周波数のピークと高い周波数のピークが操作しや

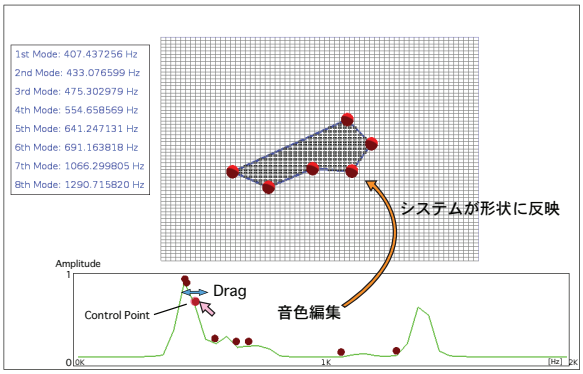


図 4. 周波数領域での音色編集



図 5. テーブルをあたかも打楽器のように叩くことで現在編集している音色を鳴らして確認する.

すいように全体がスケーリングされている.

ユーザは中央の画面において物体の周囲の制御点 (赤い点) をマウスでドラッグして 2 次元的に物体形状を変形する (図 3). また, 制御点を結ぶ辺の中間をクリックすることで物体周囲の任意の位置に制御点を追加することも可能である. ユーザが物体形状を変更するとそれを叩いたときの音色が再計算され, モード周波数とパワースペクトラムの表示が更新される. 一方, 下側のパワースペクトラム表示画面には周波数軸上で現在の形状でのモード周波数の位置に制御点 (赤い点) が設けられており, ユーザはこれをマウスでドラッグすることによりモード周波数の位置を移動させて音色自体を編集することができる (図 4). 音色が編集されるとシステムは現在の形状からできる限り近い形状でその音色を再現できるような形状を計算して画面中央に表示する. 音色編集はあまりにも大幅な変更を許してしまうと計算が収束しないか時間がかかってしまうので現在のモード周波数の近傍に編集範囲を制限する. ユーザはこの形状操作と音色操作の 2 つの操作を繰り返して最終的な目標の形状と音色を同時に作成していく.

ここで, ユーザにとってモード周波数を直接操作することは音色操作の方法として決して最適では無い. しかし音色と形状を結びつけるパラメータであるモード周波数を操作することは, 提案手法を実現する上で最も直接的であるためこの方法を採用した. また, モード周波数の位置はパワースペクトラムのピーク位置近傍となる場合が多いが, 減衰が早いモードではピークとして表れてこない場合もあり, ユーザにとって操作しにくくなってしまうこともある. この点については今後の改善が必要である.

さらに, 音色の確認についてはモード周波数やパワースペクトラムを視覚的に表示するだけでは不十分であり, 実際に音として耳で聴いて評価する必要がある. しかし, 楽器としてその出音が音楽的かどうかという評価を行うためには規則的に音を鳴らしたり, マウスでボタンをクリックしたときに音を鳴らす, というだけでは不十分であり, 実際にユーザがその音を使って音楽を演奏して評価するのが妥当である. そこで本研究では Possessing Drums[10] の手法を音色フィードバックの手段として採用する. これは音響信号処理によって, マイクから入ってきた音 (例えばテーブルを叩いた音) から駆動音 (テーブルの振動特性を差し引いたもの) を抽出した後に, 任意の音響伝達特性を再現できるような共振器に通して音色変換をすることによって身の回りの物体に任意の音を割り当てる研究である. これによって現在編集している物体の伝達特性 (それぞれのモード周波数を持つ共振器) にマイクからの音を通すことで, PC を使って設計を行っているユーザは目の前のテーブルを叩いてシミュレーションされた音色で音楽を試奏することでその評価を行うことが可能となる (図 5).

4 アルゴリズム

4.1 音色のシミュレーション

物体の振動音をシミュレーションするためには有限要素法で線形弾性体として離散化して固有値解析を行うのが一般的である. しかし, 従来固有値分解によって最低モードだけではなく高次モードも含めて求めることは非常に計算コストが高く, インタラクティブな計算に利用することは困難であった. そこで本研究では部分構造合成法 [2] を応用した前処理付きの高速な固有値解析手法を提案する.

有限要素法で離散化された自由度 N の固有値問題は次のような一般化固有値問題に帰着することができる.

$$\boldsymbol{K}\boldsymbol{\phi} = \lambda \boldsymbol{M}\boldsymbol{\phi} \tag{1}$$

$\boldsymbol{K}$ は $N \times N$ の剛性行列, $\boldsymbol{M}$ は質量行列, λ は固有値, $\boldsymbol{\phi}$ は固有ベクトルである. ここで任意の数 $n(n << N)$ だけ固有ベクトルを並べた行列 $\boldsymbol{U}$ によって $\boldsymbol{K'} = \boldsymbol{U}^T\boldsymbol{K}$, $\boldsymbol{M'} = \boldsymbol{U}^T\boldsymbol{M}$, $\phi' = \boldsymbol{U}\boldsymbol{\phi}$ と定義すると式 1 は,

$$\boldsymbol{K'}\boldsymbol{\phi'} = \lambda \boldsymbol{M'}\boldsymbol{\phi'} \tag{2}$$

と変形することができ, 本来の問題に比べて非常に小さな自由度の簡単な問題として解くことができる.

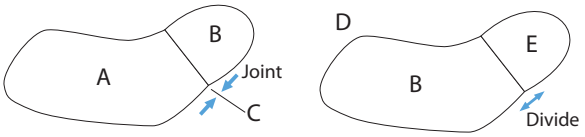


図 6. 部分構造の接合 (左) と分割 (右)

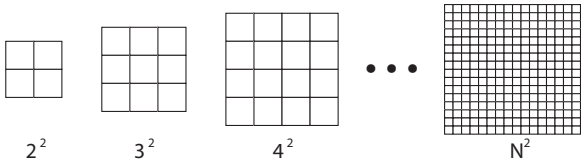


図 7. 前計算を行うグリッド



図 8. 変更後の形状は, 変形されていない部分 + (変形された部分を覆う正方形領域 D - 変形された部分を除きかつ変形されていない部分を除く D 内領域 (青色)) となる.

この問題を部分構造合成法を使って解く.

部分構造合成法 [2] は解析対象の構造体を複数の分系に分割し, 分系ごとに自由度縮小して重ねあわせることによって高速に解析を行う手法である. その中でも特に差分モード法 [7] ではある分系の固有関数が既知であった場合にその分系に関しては自由度縮小を行ったまま他の分系と合成することで計算コストを大幅に抑えることができる.

まず, 図 6 左のように物体 A に物体 B を接合し, その共有する節点群を添字 C で表すとする. ここで, 分系 A の固有関数 $\boldsymbol{U}_A$ と分系 B の固有関数 $\boldsymbol{U}_B$ がそれぞれ既知であるとすると, 全体の剛性行列 $\boldsymbol{K'}_{total}$ は自由度を縮小することができ, AB 全体の構造を一度に計算するより大幅に省メモリかつ少ない計算コストで処理することができる.

$$\boldsymbol{K'}_{total} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{U}_{A_C}^T \boldsymbol{K}_{CC} \boldsymbol{U}_{A_C} & \boldsymbol{U}_{A_C}^T \boldsymbol{K}_{CB} \boldsymbol{U}_{B_C} \\ \boldsymbol{0} & \boldsymbol{U}_B^T \boldsymbol{K}_{BC} \boldsymbol{U}_{A_C} & \boldsymbol{U}_B^T \boldsymbol{K}_{BB} \boldsymbol{U}_B \end{bmatrix} + \begin{bmatrix} \boldsymbol{U}_A^T \boldsymbol{K}_{AA} \boldsymbol{U}_A & \boldsymbol{U}_A^T \boldsymbol{K}_{AC} \boldsymbol{U}_{A_C} & \boldsymbol{0} \\ \boldsymbol{U}_{A_C}^T \boldsymbol{K}_{CA} \boldsymbol{U}_A & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \quad (3)$$

各行列の添え字は節点の属する分系を表す. また, $\boldsymbol{U}_{A_C}$ は分系 A の固有関数の C 節点要素, $\boldsymbol{U}_{B_C}$ は分系 B の固有関数の C 節点要素である. 質量行列に関しては同様の形となるため省略する.

本研究では, 図 6 左における部分構造 A はユーザが編集する前の固有関数が既知の形状, B をユーザが編集した後に加わった部分形状と考えることができる. しかし新たに付け加えた形状 B の固有関数は未知であるためこの方法をそのまま利用することはできない. そこで本研究では未知の構造 B の固有関数を高速に求めるために前計算法を導入する.

まず, 前処理として計算領域全体を 2 次元の場合には正方形格子の有限要素法で離散化する. これは 3 次元の場合には立方体格子となる. あらかじめ十分な大きさ N をとり, 2^2, 3^2 4^2...N^2 節点の大きさの正方形形状それぞれにおいて十分な数の固有ベクトルを求めておく (図 7). ここで, 差分モード合成法 [7] では図 6 右のように固有関数 $\boldsymbol{U}_D$ が既知である構造 D から構造 E を取り除いたときの構造 B の自由度縮小された剛性行列 $\boldsymbol{K}_{B'}$ は,

$$\boldsymbol{K}_{B'} = \boldsymbol{U}_D^T \boldsymbol{K}_B \boldsymbol{U}_A = \boldsymbol{U}_D^T \boldsymbol{K}_D \boldsymbol{U}_D - \boldsymbol{U}_D^T \boldsymbol{K}_E \boldsymbol{U}_D \quad (4)$$

と表すことができる.

つまり, 図 6 右における未知の構造 B を覆う構造 D の特性が分かっていれば構造 B の自由度縮小された行列 $\boldsymbol{K}_{B'}$ を構造 B 単独の固有値分解を行うことなく求めることができる. この未知の構造 B を覆う構造 D として前計算した正方形領域の結果を使うと, 図 8 のように変形後の形状は, 変形されていない部分 + (変形された部分を覆う正方形領域 D - 変形された部分を除き, かつ変形されていない部分を除く D 内領域 (図 8 の青色部分)), となる. この第一項は既知であり, 第二項も差分モード合成法の分割処理で計算すれば固有値分解は不要であるのでこれは最終的に自由度が大幅に縮小された後の全体の式 3 の固有値分解を一度解くだけで計算が完了する. これによって従来より高速な固有値解析が可能となる.

ここで, 計算精度に関してはそれぞれの部分構造においてどれだけ自由度を残しておくかに依存する. これはユーザが形状をどれだけ大きく変更したかということと計算コストのトレードオフで決定する.

4.2 所望の音色を実現するための形状最適化

いくつかのモード周波数, つまりは与えられた音色からそれに近いモード周波数を実現するような形状を推定する手法について説明する.

本研究では, ユーザが現在編集している形状からできる限り近い形状で所望のモード特性が得られる解をレベルセット法による形状最適化で求める. レベルセット法 [8] とは二つの相で構成される閉じた領域においてそれぞれの相の形態と形状を表現する方法である. 本研究の場合は物体表面からの距離関数 $\phi(x)$ がこれに相当し, 物体の内部で正, 外部で負の値を取り, 表面で零になる. レベルセット法に基づく構造最適化手法 [11] は, トポロジー最適化法に代表される他手法 [12] で問題となる形状の波打ち問題やグレースケール (物体と空洞の中間領域) 問題が発生せず, 計算コストも少ないという利点がある. ここで, 通常, 構造最適化問題では, 得られる構造が

初期構造に大きく依存してしまうことが問題として挙げられるが, 本研究ではユーザが現在編集している形状から最も近い最適形状を得ることを目的としているためにこの点はむしろ利点となる

最適化はレベルセットの関数である目的汎関数 $F(\phi(x))$ を最小化することによって行う. 線形弾性体の固有モード最適化問題においては目的汎関数は, 現在の固有値と目的とする固有値間の残差となる [11].

$$Minimize \quad F(\phi) = \sum_{k=0}^{s+1} \left\{ w_k \left(\frac{\lambda_k - \lambda_{target,k}}{\lambda_{target.k}^2} \right) \right\} \quad (5)$$

ここで λ_k は k 番目の固有値, $\lambda_{target,k}$ は目標とする k 番目の固有値, w_k は重みパラメータである. s は目標とする固有値モードの数であるが, $s+1$ まで考慮しているのは, 最適化の過程で固有ベクトルの形状が変化して次数が入れ替わってしまうことを防ぐためである. この最適化はレベルセット境界に適切な法線方向速度参照領域 $V_N(x,t)$ を与えて陽解法で移流させることで行う.

レベルセットの時間発展を表すレベルセット方程式の弱形式は参照領域 D において

$$\int_D \frac{\phi(x,t+\Delta t) - \phi(x,t)}{\Delta t} \Psi d\Omega + \int_D V_N(x,t) \Psi d\Omega = 0 \quad (6)$$

となる. ここで, Ψ は試行関数である. 形状最適化問題においては左辺第二項は目的汎関数のレベルセット関数に関する変分

$$\begin{aligned} \int_D f(\phi(x)) \psi d\Omega = & \int_D \sum_{k=0}^{s+1} \left\{ 2 w_k \left(\frac{\lambda_k - \lambda_{target,k}}{\lambda_{target.k}^2} \right) \right. \\ & \times (\epsilon(u_k) : \boldsymbol{D} : \epsilon(u_k) - \lambda_k \rho u_k \cdot u_k) \\ & \left. \times \delta(\phi(x)) H(\phi(x)) \right\} \psi d\Omega \end{aligned} \quad (7)$$

で表される. この $f(\phi(x))$ は目的汎関数の形状感度の被積分関数であり, ϵ はひずみテンソル, u_k は k 番目の固有ベクトル, $\boldsymbol{D}$ は弾性テンソル, $H(\phi(x))$ と $\delta(\phi(x))$ はレベルセット関数が $-h \leq \phi \leq h$(h は任意の実数) の領域で値を持つ Heavisible 関数と Dirac Delta 関数 [11] である. 式 6 は有限要素法で離散化して,

$$\phi^{t+\Delta t} = \phi^t - \Delta t \boldsymbol{E}^{-1} \boldsymbol{V}_N^t \quad (8)$$

$$\boldsymbol{E} = \bigcup^{n_e} \int_{V^e} \boldsymbol{N}^T \boldsymbol{N} d\Omega \quad (9)$$

$$\boldsymbol{V}_N^t = \bigcup^{n_e} \int_{V^e} f(\phi(x,t)) \boldsymbol{N} d\Omega \quad (10)$$

に従って目的汎関数が収束するまで各節点のレベルセット関数を更新することで解く. ここで, $\boldsymbol{N}$ は形状関数, $\bigcup$ は要素の重ね合わせを表す.

アルゴリズムの手順としてはユーザが現在編集している固有関数が既知の形状を初期条件とし, 式 10 にてレベルセットを更新する. 更新の度に 4.1 章の方法によって固有値解析をやり直し, 現在の固有値と固有ベクトルを設定し直す. また, レベルセット境界を更新すると, 境界から離れたところのレベルセットの値が無効になるので Fast Marching Method[9] 等を用いて全体を再初期化する必要がある.

ここで, レベルセット法では物体形状を距離関数によって陰的に表現しているため, 計算が収束した後にユーザがさらに形状を編集できるようにするためにはレベルセット法から明示的に物体表面のメッシュを生成してやる必要がある. 本研究では陰関数からのメッシュ生成には Marching Cube 法 [14] を用いた. さらに Marching Cube 法で生成されたメッシュはレベルセット格子と同程度の解像度となっているために細か過ぎてユーザが頂点を操作しにくくなってしまうという問題があるがスムージングを行い, 荒いメッシュに近似することでこれを解決した.

5 結果

5.1 システムの速度と精度

本稿で提案する手法で形状変形を行いながらモード周波数を算出していった場合と, それと同じ形状で従来の固有値解析を行った場合の節点数とおよその処理時間の関係を表 1 に示す. 括弧内の数字は変形前後で追加された節点数である. 評価に使ったシステムは, CPU: Intel Core i7 Quad 3.4GHz, RAM: 32GB で C++で実装し, 各形状において最低次から 8 つの固有周波数を求めた. 既存手法では節点数の増加に応じて計算時間が大幅に増加しているが, 提案手法では僅かな増加で済んでいる.

表 1. 固有値解析の時間 [ms]

節点数 (追加節点数)	既存手法	提案手法
6588(1314)	2000	400
9026(2475)	18000	840
13576(3081)	37000	1030

一方, 形状最適化の精度に関しても評価も行った. 最低次から 8 つのモードを対象として目標値を定め, 初期状態から最適化を行った結果が表 2 である. 結果の 6 次モードに新たなモードが入ってきており, 次数が変わっているが, 目標とする固有周波数群自体はほぼ得られてることが分かる. 8 次の目標値も達成するためには最適化の対象にするモードの数をさらに増やす必要がある.

5.2 エンジニアによる試用

本システムを実際に CAD や解析ソフトウェアの使用経験のあるエンジニア数名に試用してもらった後, インタビューを行った. 被験者には自分の好みの打楽器を作ることを想定してデザインを行うよう

表 2. 固有周波数最適化の結果 (単位 [Hz])

	初期状態	目標値	結果
1st	440	409	410
2nd	674	738	764
3rd	758	967	955
4th	1119	1367	1312
5th	4305	4776	4633
6th	8656	9500	5342
7th	9912	10649	9408
8th	11279	12567	10610

に指示した.

彼らは普段の設計では既に過去に実績のあるCADデータを雛型として, それに編集を加えていくことで新しい製品を作っていくことが多いとのことである. そのため, 本システムのアプローチは目的となる音色側からも操作を加えていくことができると楽器音に不快な成分が含まれていた場合に抑えるのが容易になったり, または新たな楽器音を探していく際にも非常に便利である, と意見を述べた. また, 逆解析の結果得られた形状そのものを再びインタラクティブに修正していけるため, 普段使っている解析ツールにも本システムが取り入れられると作業が容易になるとの意見を得た.

しかし, 実際の楽器は対称な形状のものがほとんどであるので, 対称性を前提にしたインタフェースにしたほうがよいとの意見も得られた. さらに, 生楽器では形状を自由に設計するという需要はあまり無く, 例えばスピーカ筐体等の設計に利用する方がより有用であるとの意見も多かった. これらは今後の課題とする.

6 まとめと今後の課題

本稿では, 物体形状からそれを叩いたときの音色を計算する順解析と音色からそれを再現可能な物体形状を推定する逆解析を相互に行うことで形状と音色両面からの効率的な打楽器デザインを行うことのできるインタフェースを提案し, 実際にエンジニアにも試用してもらうことでその創造的なデザインに役立つ可能性があることを確認した.

しかし, 本システムの有用性を評価するためにはソフトウェア上での評価だけでは不十分であり, 今後は実際に作成した打楽器を実物として出力して評価を行っていく必要がある. さらに, 実際の楽器では数多くの部品が組み合わされているため, 本稿の手法だけでは実用的とは言えない. 複数の部品の組み合わせを考慮することは今後の課題である. また, 音色からの形状推定ではユーザの操作次第では物理的に実現不可能な音色が設定されてしまうことがあるという問題もある. このため予めユーザに編集可能な範囲をシステム側から提示できることが好ましいが, 最適化を行う前の段階ではその音色が実現可能がどうかを知ることは難しく, 今後の課題である.

参考文献

[1] 梅谷信行, 高山健志, 三谷純, 五十嵐健夫. 実時間固有値解析による対話的な鉄琴のデザイン. 情報処理学会論文誌, Vol.52, No.4, pp. 1599–1607, 2011.

[2] 長松昭男, 大熊政明. 部分構造合成法. 培風館, 1991.

[3] Changxi Z., and Doug L. J. Toward High-Quality Modal Contact Sound. In *ACM Transactions on Graphics, 30(4)*, 2011.

[4] Cynthia B. M., and David B. Modal Parameter Tracking for Shape-Changing Geometric Objects. In *Proc. of the 10th Int. Conference on Digital Audio Effects (DAFx-07)*, 2007.

[5] Dane C., Chi-Lun L, Arthur G. E., Daniel F. K. Design by Dragging: An Interface for Creative Forward and Inverse Design with Simulation Ensembles. In *Transactions on Graphics 27(3)*, 2008.

[6] Greg S., Manfred L., Mitani J., and Igarashi T. SketchChair: An All-in-one Chair Design System for End-users. In *the fifth International conference on Tangible, Embedded and Embodied Interaction*, pp. 260, 2011.

[7] Mochizuki T., and Ogawa M. Efficient Approach for Structural Modi cation Analysis. In *Review of Automotive Engineering, Vol.27, No.1*, pp. 75–82, 2006.

[8] Osher S., and Fedkiw R. Level Set Methods and Dynamic Implicit Surfaces. In *Springer*, 2002.

[9] Sethian J. A. Level Set Methods and Fast Marching Methods. Cambridge Monographs on Applied and Computational Mathematics. In *Cambridge University Press, 2nd edition*, 1999.

[10] Yamamoto K. Possessing Drums: An Interface of Musical Instruments that Assigns Arbitrary Timbres to Personal Belongings. In *Journal of Information Processing, Vol.21 No.2*, pp. 274–182, 2013.

[11] Yamasaki S., Nishiwaki S., Izui K., and Yoshimura M. A structural optimization method based on the level set method using a new type of re-initialization scheme. In *International Conference on Computational Methods*, pp. 260, 2007.

[12] Yang R. J., and Chung C. H. Optimal Topology Design using Linear Programming. In *Computers and Structures*, Vol.53, pp. 265–275, 1994.

[13] Umetani N., Igarashi T., and Niloy J. M. Guided Exploration of Physically Valid Shapes for Furniture Design. In *ACM Transactions on Graphics 31(4)*, 2012.

[14] W.Lorensen, H.Cline. Marching Cubes: A High Resolution 3D Surface Construction Algorithm. In *Computer Graphics, 21 (4)*, pp. 163–169, 1987.

SensingFork：フォーク型センサを用いた食行動改善手法の研究

門村 亜珠沙　李 爭原　陳 嚴章　塚田 浩二　朱 浩華　椎尾 一郎*

概要. 『ピーマンも食べてね』と小さな子供に言葉で説得するのは難しい．もし，嫌いな食材を食べたときや夕ご飯を全部食べたときに，やる気の出るフィードバックを返したら，好き嫌いの多い子供の食行動も楽しく改善できるかもしれない．そこで我々は，子供の食行動の改善を目的に，食行動（食事中の動作と大まかな食材の種類）を検出するフォーク型デバイス：センシングフォークを開発した．本論文では，センシングフォークのプロトタイプ設計と実装，このセンシングフォークと連動し，子供の食習慣改善を目指すスマートフォン用アプリケーション：腹ペコパンダ，および，本システムの性能評価について述べる．

1　はじめに

バランスの良い食事は健康を維持する上で重要な要因である．現在，世界各国で食事の摂取方法・食材の選び方・食に関する文化など，広い視野から食について教育すること，つまり食育が重要な課題となっている．特に，先進国では，政府が web サイト上でも食にまつわる多くの情報を公開するなどして，食育を推進している[1] [2]．日本でも，2005 年に，バランスの悪い栄養摂取や野菜摂取量の不足などの不健康な食習慣の増加を危惧し，人々が活発で健康的な生活を楽しむことができるように，食育基本法が制定された[3]．その結果，現在では，多くの幼稚園では食育活動を行っており，例えば，教諭らは子供たちに様々な種類の食材を食べることを促している．しかし，食育活動を引き続き家庭内で行うのは，時間・知識不足などの理由から親子にとって難しい [11]．そこで，我々は小さな子供にとって身近な食器であるフォークに着目し，家庭内での親子の食育支援を行うシステムを構築する．フォークのような身近な日用品を HCI 技術で拡張することで，子供たちが楽しく食習慣を自然に改善することを目指す．

本研究では，食事中の動作と大まかな食材の種類を検出するために，カラーセンサ／電極／マイコン／Bluetooth／バッテリーなどをフォーク自体に組み込んだデバイス：センシングフォークの設計と実装を行った．さらに，センシングフォークと連動し，子供の食習慣改善を目指すスマートフォン用のアプリケーション：腹ペコパンダを実装した（図 1，[4]）．



図 1. システムの外観

2　関連研究

近年，HCI の分野において，食に関する多くの研究が発表されている．Grimes らは，食事をコンピュータで拡張する，Human-Food Interaction の可能性について調査・議論を行っている [3]．Parker ら [9] は，食事のオンライン広告に，栄養素や適切な食べ合わせを併せて示した場合の，子供たちへの影響を調査している．Mansour ら [7] は，栄養バランスの良悪を楽しみながら学ぶことができる子供や若者向けのコンピュータゲームを開発している．これらの研究以外にも，食育支援のための様々なサービス[4] [5] が存在するが，その多くは PC やスマートフォン上のみで動作するものだった．

一方，人の食行動に対応したインタラクションを実現するために，食事の場にセンサやコンピュータを組み込むなど，ユーザの食行動に関する研究も多数行われている．Playful Tray[6] は，皿の横に取り付けたディスプレイから食事の進行に合わせて視覚的フィードバックを与えることで，食べるのが遅い子供を支援している．Dining Presenter[8] は，テーブル上方にプロジェクタを設置することで，事

Copyright is held by the author(s).

* Azusa Kadomura and Itiro Siio, お茶の水女子大学, Cheng-Yuan Li, Yen-Chang Chen and Hao-Hua Chu, 国立台湾大学, Koji Tsukada, 公立はこだて未来大学

[1] The U.S. Department of Agriculture: http://www.choosemyplate.gov

[2] British Nutrition Foundation: http://www.nutrition.org.uk/

[3] 日本内閣府: http://www8.cao.go.jp/syokuiku/index.html

[4] Fatworld: http://www.fatworld.org/

[5] Shokuiku Town: http://www.glico.co.jp/shokuiku/game/

前に記入した絵やメッセージをテーブルに映し出し，食事を楽しくする効果や食育支援を目指している．Hapifork[1] は，食材を口に運ぶ頻度から早食い行動を感知し，振動で知らせるフォーク型デバイスである．また，本研究に先立って我々は，ユーザが食材を食べた瞬時に音フィードバックを返すことで，食行動改善を目指すフォーク型デバイスを開発してきた [5]．これらのシステムは，少数のセンサにより特定の食行動を検出しているため，アプリケーションが限定されていた．そこで本研究では，人の食行動（食事中の動作と大まかな食材の種類）を検出する多様なセンサを搭載したフォーク型デバイスと，これを用いたアプリケーションを開発した．

3 システム概要

本システムは，センサを組み込んだフォーク：センシングフォーク（図 2 左）と，フォークと連携して動くスマートフォンアプリケーション（図 2 右）から構築される．スマートフォン上のソフトウェアには，主に 3 つのモジュールがある．①フォークのセンサデータからユーザの食行動を推定する「食動作検出部」，②フォークと接触した食材を大まかに判別する「食材判別部」，および，③子供の食習慣の向上を目指した食育ゲーム：腹ペコパンダである．これらを組み合わせて，ユーザが「何を」「どのように」食べているのかを検出し，適切なフィードバックを返すことで食育を支援する．次に，センシングフォークのプロトタイプの設計と実装について述べる．

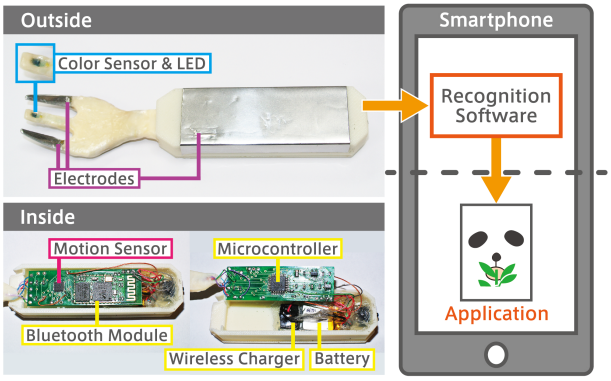


図 2. システム構成
左：センシングフォーク本体のプロトタイプ
右：スマートフォン上のソフトウェア

4 センシングフォーク

図 2 に示すように，センシングフォーク本体は，様々なセンサやハードウェア部品を内蔵している．得られたセンサデータは，スマートフォン上に送信され，そこで食動作と食材の検出が行われる．フォークの持ち手サイズは L80 × W25 × H15 mm に収めることで，小さな子供でも容易に握れるように設計した．以下，各モジュールについて詳細に説明する．

4.1 ハードウェア

4.1.1 電極

センシングフォークは，先端部分と握り部分に計 3 つの電極を備える（図 2 左上）．先端部分の両端 2 つの電極は，食材とフォークが接触したことを検出し，同時に食材の電気抵抗値を測定している．電気抵抗値は食材により異なるため，食材の推定に利用できる可能性がある．一方，握り部分の電極は，先端部分が食材を介してユーザの口に触れたとき，握り部分→手→腕→口→食材→先端部分と微弱な電流を流す．ここで，握りと先端の電極間の電気抵抗値を測定することで，フォークと食材がユーザの口に接触したこと（＝摂食動作）を検出できる．以上の電極部分の構成と，周辺回路について，図 3 に示す．

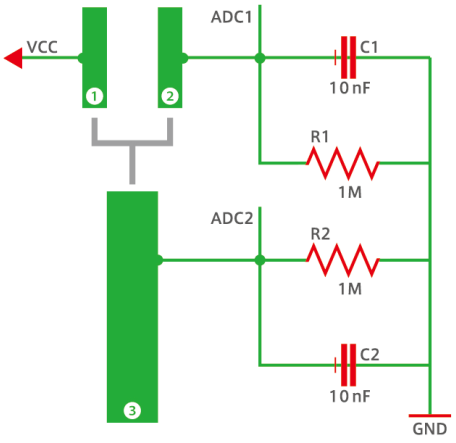


図 3. 電極とその周辺回路図
①と②はフォーク先端部分の電極を示し，
③はフォーク握り部分の電極を示している

4.1.2 カラーセンサ

フォークの先端部分中央には，カラーセンサと白色 LED を並べて組み込んだ（図 2 左上）．カラーセンサ（Avago ADJD-S311-CR999）は L2.2 × W2.2 × H0.76 mm と超小型であり，フォークに接触した食材の色（RGB 各色と輝度）を 1024 段階で検出し，I^2C を介してマイコンに送信する．フォークが食材の中に差し込まれると，白色 LED が点灯し，食材からの反射光をカラーセンサで測定する．なお，先端部の電極・およびカラーセンサ基板は，白色レジンでフォークの形状に固定した後，透明レジンで表面をコーティングした．レジンは共に歯科用の素材を利用することで，食の安全と防水に配慮した．

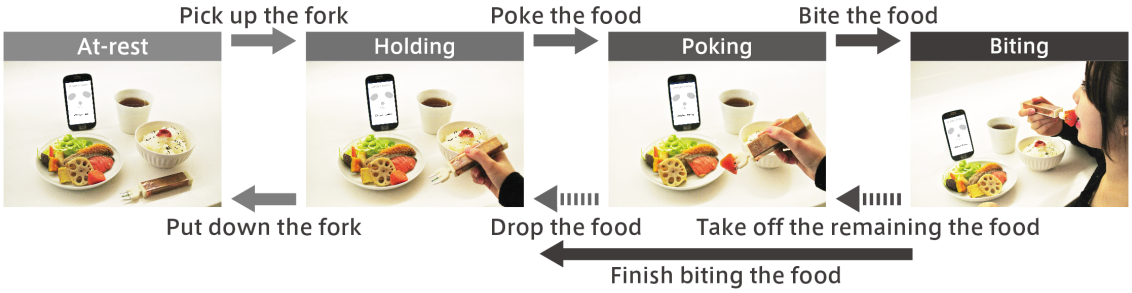


図 4. フォークの状態から導かれる食動作の遷移

4.1.3 モーションセンサ

フォークの握り内部のメイン基板上には，モーションセンサ（3 軸加速度＋ 3 軸ジャイロセンサ: InvenSense MPU-6050）を設置した．これらのセンサデータはフォークの動きを検出し，I^2C を介してマイコンに送信され，食動作の推定に活用する．食動作検出の詳細は，4.2.1 で説明する．

4.1.4 その他のモジュール

握り内部のメイン基板上にはマイコン，Bluetooth モジュール（SeeedStudio Serial port bluetooth module），リチウムイオン充電池，およびワイレス充電機構を搭載した（図 2 左下）．マイコンは Atmel ATMega328 を利用し，Arduino 互換機として実装した．Bluetooth モジュールは，上述したセンサデータをまとめてスマートフォンに送信する．リチウムイオン充電池はサイズを考慮して 120mAh のものを選択し，毎秒 10 回の無線送信を行っても 2 時間以上連続動作することを確認した．バッテリーを内蔵し，無線通信機能を備えることで，通常のフォークとできるだけ同じように扱えるように配慮した．

4.2 ソフトウェア

4.2.1 食動作検出部

食動作検出は，フォークから Bluetooth 経由で送信されるセンサデータを元に，スマートフォン上のアプリケーションで判定する．本システムでは，フォークの状態から以下の 4 種類のユーザの食動作を推定する．図 4 に食動作の一覧と遷移を示す．以下，食動作検出と推定方法について述べる．

- 休止状態（At-rest）：
  休止状態とは，ユーザがフォークを握っていない（フォークが静止している）状態のことである．ユーザが使う食器がフォークだけと仮定すると，ユーザが食事のための動作をしていないと推定できる．フォークのモーションセンサが，一定時間動きを検出しない場合，休止状態にあるとシステムは判定する．
- 把持状態（Holding）：
  把持状態とは，食材を刺していないフォークをユーザが手に持っている状態のことである．この状態からは，ユーザがフォークを持ち上げていて，何かしらの食材をフォークで食べようと考えていると推定できる．フォークのモーションセンサがフォークの動きや手振れを感知した場合，把持状態にあるとシステムは判定する．
- 刺突状態（Poking）：
  刺突状態とは，ユーザが食材をフォークで刺した状態のことである．この刺突状態からは，ユーザが食材を刺したが，まだ口に運んでいないことを推定できる．フォーク先端の 2 つの電極間に一定の電流が流れることで，刺突状態にあるとシステムは判定する．なお，刺突状態をトリガーとして，次項（4.2.2）で説明する食材判別を開始する．
- 摂食状態（Biting）：
  摂食状態とは，フォークに刺さっている食材がユーザの口に接している状態のことである．この摂食状態からは，ユーザがフォークに刺さっている食材を食べていることが推定できる．フォーク握り部分の電極と先端部分の電極間に微弱な電流が流れることで，摂食状態にあるとシステムは判定する．

これら 4 つの食動作の判定をより正確にするために，次の 2 つの仕組みを導入した．まず，各状態の間の遷移に図 4 のような制約を設けることで，誤認識を軽減する．すなわち，休止状態からは把持状態へしか遷移せず，把持状態からは刺突状態か休止状態へしか遷移しない．例えば，フォークで食材を刺さない（＝刺突状態にならない）限り，把持状態から摂食状態へは移行しない．これにより，例えば子供が食材を刺さずにフォーク先端をなめる動作と摂食動作を区別することができる．次に，摂食状態においては，フォーク握り部分と先端部分の電極に人体を介して微弱な電流が流れるため，フォークの把持者と摂食者が同一であることが判定できる．

4.2.2 食材判別部

本研究では，ユーザにバランスの良い食事摂取を促す目的で，2 通りの手法を試作した．1 つ目は食材種類判別手法であり，食事前に初期化プロセスを設け，各食材の色と電気抵抗値を登録する．次に，ユーザが食事中，食材を刺す度に，初期データと照合することで食材種類の判別を行う．機械学習アルゴリズムとしては，サポートベクターマシーンを利用し，LIBSVM[2] を用いて実装した．2 つ目は食材色判別手法であり，初期化は行わず，ユーザが食事中に食材を刺す度に，その食材の色を記録する．なお，ノイズを軽減するために，過去 15 回の色データを平滑化して利用する．

食材色判別手法では，食事中に取得した色データを随時読み取ることで，色の絶対値から食材の大まかな種類を推定する．本手法では厳密に食材を判別することは難しくなるが，見た目の色の違う多くの食材を食べることで食育を支援する食材 5 色バランス健康法 [10] も提案されており，色ベースとした判定手法には一定の有効性があると考える．また，初期登録の手間がないことから，導入や運用も容易である．よって，本論文では，食材色判別手法を用いて，後述の食育ゲームを実装した．

5 食育ゲーム：腹ペコパンダ

本章では，健康的な食生活のために様々な色の食材を摂取することを促す，センシングフォークと連動した食育ゲーム「腹ペコパンダ」について説明する（図 1，図 5）．腹ペコパンダでは，スマートフォン画面上に表示されたパンダが，食材を把持したり，摂食したりなど，ユーザ（子供）と同じ食行動を行う．ここで，お腹を空かせたパンダに子供が共感することで，実際の子供の行動に影響を与えることを目指した．仮想ペットとしてパンダを題材とした理由は，アジア諸国の子供に親しまれているキャラクターだからである．腹ペコパンダは「偏食」「注意散漫」という子供の一般的な食の問題[7]の改善を目指す．まず，子供が別の色の食材を食べると，ゲームの得点が増える．これにより，高得点を得るために，より多くの食材摂取を動機付ける．次に，子供が食事時間中に一定時間摂食行動を停止したときに，パンダが悲しむことで，子供に食事に戻ることを促すフィードバックを返す．

以下に，フォークの状態と腹ペコパンダの画面（図 5）の関係について説明する．

1. まず，お腹を空かせて悲しんでいるパンダが登場し，子供が食べ始めるまで，『僕は腹ペコパンダ、お腹が空いたよ。』と話す.【フォークの状態：休止・把持状態】

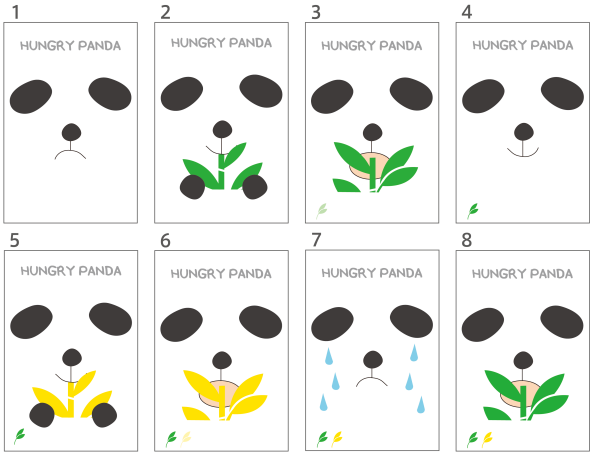


図 5. 腹ペコパンダの画面例

2. 子供が食材を刺すと，パンダも子供が刺した食材と同じ色[8]の笹を持つ.【刺突状態】
3. 子供がその食材を摂食すると，パンダもその笹を食べ『おいしい！』と話す.【摂食状態】
4. フォークの状態が把持状態に移行すると，パンダは笹ポイントを取得し，笑顔になる.【把持状態】
5. 同様に，子供が別の食材を刺すと，パンダは子供が刺した食材と同じ色の笹を持つ.【刺突状態】
6. 子供がその食材を摂食すると，パンダもその笹を食べ,『これもおいしいね！』と話す.【摂食状態】
7. 子供が食事を中断すると（＝一定時間，休止・把持状態が続くと），パンダは泣きながら，『もっと食べたいよ！』と子供がまた食べ始めるまで訴えかける.【休止・把持状態】
8. 子供が再び最初の食材を摂食すると，パンダは最初の笹と同じ色の笹を食べるが，笹ポイントは増えない.『他の食材も食べたいよ！』と話すことで，多様な食材の摂食を促す.【摂食状態】

最後に，終了ボタンを押すと，腹ペコパンダが獲得した笹ポイントの数に対応したアニメーションが表示される．つまり，万遍なく食べることで得られる笹ポイントの数に従って，最後に表示されるアニメーションが変わる．これにより，子供がより多くの食材を摂食する動機付けを行っている．

[7] 厚生労働省：2005 年度乳幼児栄養調査結果 http://www.mhlw.go.jp/toukei/list/83-1.html

[8] 現システムでは，カラーセンサが検出した RGB 値を予め設定した閾値で分類して，5 色バランス健康法に基づく 5 色【赤・白・黄・緑・黒】の笹にマッピングしている．

表 1. 各食動作認識の成功回数と各認識に要した平均時間

	User A	User B	User C	User D	User E	User F	各状態の平均
Holding 1【把持状態 1】	5(0.3sec)	5(0.5sec)	5(0.3sec)	5(0.5sec)	5(0.6sec)	5(0.4sec)	5.0(0.4sec)
Poking【刺突状態】	5(0.6sec)	5(0.3sec)	5(0.9sec)	5(0.6sec)	5(0.6sec)	5(0.5sec)	5.0(0.6sec)
Biting【摂食状態】	5(0.3sec)	4(0.9sec)	3(0.9sec)	4(1.2sec)	4(1.1sec)	3(0.7sec)	3.8(0.9sec)
Holding 2【把持状態 2】	5(1.1sec)	5(1.8sec)	5(1.5sec)	5(1.1sec)	5(0.5sec)	5(1.2sec)	5.0(1.2sec)
At-rest【休止状態】	5(1.5sec)	5(1.4sec)	5(1.5sec)	5(1.5sec)	5(1.1sec)	5(1.4sec)	5.0(1.4sec)
各ユーザの平均	5.0(0.8sec)	4.8(1.0sec)	4.6(1.0sec)	4.8(1.0sec)	4.8(0.8sec)	4.6(0.8sec)	

このように，本ゲームはセンシングフォークによって推定・検出されたユーザの食行動を基に，視覚・聴覚的なフィードバックを返し，食事時間中に食育を支援することを目指す．すなわち，子供は楽しく食行動を改善でき，親はパンダの獲得した笹ポイントを見ることで，日々の子供の食行動を継続的に把握することができる．付随的な効果としては，子供の食行動が改善された場合には，本ゲームによるご褒美（笹ポイントと最後のアニメーション）に加えて，親がお菓子などのご褒美を子供に与えるなどして，親子が食育に楽しく取り組むきっかけにつながると期待できる．

6 評価実験

センシングフォークの性能を評価する基礎的な評価実験を行った．本章では，食動作検出と食材判別の精度について述べる．

6.1 食動作検出について

食動作検出の精度を確認するために，6 人のユーザ（男女各 3 名，21〜28 歳）にセンシングフォークで食材（一口サイズのチーズ）を 5 回食べてもらった．一連の食動作の認識精度を確認するため，各ユーザには 1 回食べるごとに，フォークを机に置くように指示した．すなわち，一度食材を食べる度に【把持状態 1 →刺突状態→摂食状態→把持状態 2 →休止状態】の順で一連の食動作が行われた．なお，全てのユーザはセンシングフォークを初めて利用し，システムの機能も事前に説明しなかった．

実験では，各ユーザの動作とスマートフォン上の実験用ソフトウェア画面を同時にビデオ撮影した．実験用ソフトウェアには，現在の食動作の状態とセンサデータがテキストで表示されている．実験終了後，実際のユーザの行動とソフトウェア上に表示された食動作を PC 上でフレームごとに目視で比較した．なお，ユーザの食動作の遷移は，【把持状態 1 ：フォークが机から離れた瞬間】【刺突状態：フォーク先端が食材に触れた瞬間】【摂食状態：ユーザが食材を噛んだ瞬間】【把持状態 2 ：ユーザの口からフォークが離れた瞬間】【休止状態：フォーク本体が机に置かれた瞬間】というフレームで判定した．

まず，各食動作が認識された回数と，システムが各認識に要した平均時間を表 1 に示す．ユーザ毎に 25 回の食動作（上述した 5 食動作を 5 セット）が行われた．その結果，全体としては 95%の認識率だった．一方,【摂食状態】において 23%の誤認識があり，【刺突状態→摂食状態】と認識されずに【刺突状態→把持状態 2 】と誤認識されていた．認識されなかったユーザの【摂食状態】をビデオ動画から観察したところ，誤認識された摂食時間の平均は 0.1 秒であり，正しく認識された場合の平均は 0.9 秒であった．したがって，フォークを一瞬しか口につけなかった場合に，誤認識が発生した．現在の実装では，バッテリーの消耗を考慮して，スマートフォンとの通信頻度を秒間 10 回（0.1 秒間隔）に抑えており，この結果，ユーザが素早く行動した場合に十分な測定が行えなかったと考えられる．センシング頻度を上げることで，食動作の認識は改善できると考えている．参考として，ユーザ C の食動作遷移例を図 6 に示す．

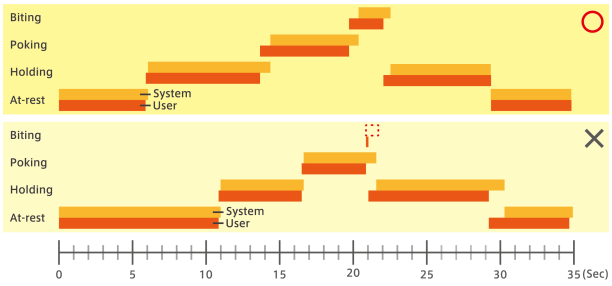


図 6. ユーザ C の食動作遷移の一例
上段が認識成功時，下段が認識失敗時の事例．各 2 列の横棒は上がシステムの認識結果，下が実際のユーザの食動作を示す．下段では,【摂食状態】が 0.07 秒と短時間で行われため，認識されていない．

また，表 1 に示す通り，各動作認識に要した時間の平均は 0.9 秒であった．各動作の中でも,【休止状態】の認識は平均 1.4 秒と時間がかかった．この理由は，1 秒以上フォークが静止したときに【休止状態】と判断しているからである．こちらも待機時間を調整することで，認識速度を改善できる可能性はあるが，一般にフォークを机に置く／取る行為が高速で繰り返されることは少ないので，実用上問題ないと考えている．また,【把持状態 2 】も平均 1.2 秒

と時間がかかっているが，これは実験プログラムの不具合（摂食時に 2.2 秒程度のアニメーションが表示され，終了まで状態表示が更新されなかった）が原因であり，実際の認識は【把持状態 1 】や【刺突状態】と同程度の時間で完了できると思われる．

6.2 食材判別について

4.2.2 で述べた食材色判別手法の性能を確認するために，市販のお弁当から色が異なった食材を 5 種類（ハンバーグ【赤】，米【白】，かぼちゃ【黄】，ブロッコリー【緑】，しいたけ【黒】）用意した．なお，【】内の色は前述の食材 5 色バランス健康法 [10] での分類を示す．実験者は，各食材の同じ面の違う箇所を 10 回ずつセンシングフォークで刺し，RGB 値を記録した．まず，同一食材毎の検出誤差（標準偏差）としては，米が平均 0.9，かぼちゃが平均 6.5 と安定していたのに対し，ハンバーグは平均 14，しいたけは平均 15，ブロッコリーは平均 25 と誤差が大きくなった．これは，ハンバーグの上のケチャップや，ブロッコリーの葉部分の隙間，シイタケの表面／内部の色の濃淡の違いといった食材の特性に影響を受けたと考えられる．

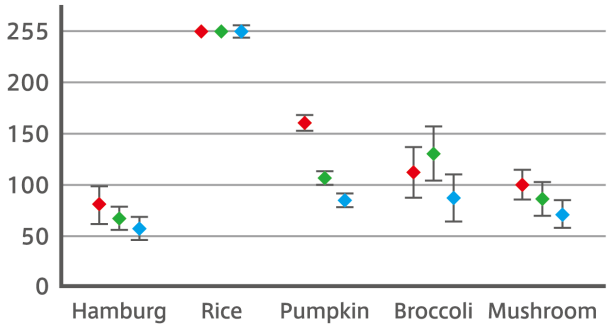


図 7. 各食材の RGB の平均値と標準偏差

次に，現在の腹ペコパンダで採用している，各食材を 5 色【赤・白・黄・緑・黒】にマッピングする方法を適用したところ，米【白】，カボチャ【黄】は 10 回中 10 回，ブロッコリー【緑】は 9 回認識に成功した．一方，シイタケ【黒】は 10 回中 5 回，ハンバーグ【赤】は 10 回中 4 回しか認識出来ず，大半が相互に誤認識していた．現在の食材色判別手法は単純な閾値で判定しているため，今後は色空間上の距離や機械学習などを利用して認識率を改良したい．

7 まとめと今後の課題

本論文では，センサ内蔵フォークを用いてユーザの食行動を検出するセンシングフォークの設計と実装について述べた．また，子供の食事中の「偏食」「注意散漫」といった問題行動を改善することを目指した「腹ペコパンダ」ゲームも実装した．さらに，フォークの基礎的な性能を検証するために評価実験を行い，食動作検出と食材の色による分類について，改良の余地はあるものの，一定の性能があることを確認した．

今後は，実際に子供を対象とした，家庭内での評価実験を計画している．さらに，食事中の更なる問題行動を検出したり，食材の摂食順序を記録したりすることで，新たな食行動改善アプリケーションの可能性を探る．

謝辞

本研究は，JSPS 科研費 25・10205 と科学技術振興機構さきがけプログラムの助成を受けた．

参考文献

[1] Hapifork. http://www.hapilabs.com/.

[2] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2(3).

[3] A. Grimes and R. Harper. Celebratory technology: new directions for food research in HCI. In *Proceedings of CHI '08*.

[4] A. Kadomura, C.-Y. Li, Y.-C. Chen, H.-H. Chu, K. Tsukada, and I. Siio. Sensing fork and persuasive game for improving eating behavior. In *Proceedings of Ubicomp '13*.

[5] A. Kadomura, K. Tsukada, and I. Siio. EducaTableware: computer-augmented tableware to enhance the eating experiences. In *Extended Abstracts on CHI '13*.

[6] J.-L. Lo, T.-Y. Lin, H.-H. Chu, H.-C. Chou, J.-H. Chen, J. Y.-J. Hsu, and P. Huang. Playful tray: adopting Ubicomp and persuasive techniques into play-based occupational therapy for reducing poor eating behavior in young children. In *Proceedings of UbiComp '07*.

[7] A. Mansour, M. Barve, S. Bhat, and E. Y.-L. Do. MunchCrunch: a game to learn healthy-eating heuristics. In *Proceedings of IDC '09*.

[8] M. Mori, K. Kurihara, K. Tsukada, and I. Siio. Dining Presenter: Augmented Reality system for a dining tabletop. In *Proceedings of UbiComp '09*.

[9] A. G. Parker, I. McClendon, C. Grevet, V. Ayo, W. Chung, V. Johnson, and E. D. Mynatt. I am what i eat: identity & critical thinking in an online health forum for kids. In *Proceedings of CHI '13*.

[10] 杉本恵子．「生活習慣病は、『食材 5 色バランス健康法』で予防・改善を」．In *Kellogg's Update*, 第 82 巻, 2005.

[11] 鈴木秀子．子どもから家庭へつなぐ食育〜保護者の「学び」からの検討〜．会津大学短期大学部研究年報, 第 67 巻, 2010.

鍵盤上への演奏補助情報投影機能をもつ ピアノ学習支援システムにおける熟達化プロセスに関する調査

竹川 佳成　　椿本 弥生　　田柳 恵美子　　平田 圭二 *

概要.　楽器の演奏技術の向上には多大な時間や労力を必要とするため，敷居の高さに利用を断念したり，習熟効率の低さから挫折してしまう演奏者が多い．この問題を解決するために，筆者らの研究グループは，鍵盤上部に設置したプロジェクタを用いて鍵盤上や鍵盤の周囲に打鍵位置情報など演奏補助情報を投影するピアノ学習支援システムを構築してきた．評価実験から比較手法である「光る鍵盤」の学習方法と比較して課題曲を効率的に学習できることが明らかになったが,「課題曲を効率的に学習できただけか，あるいは，ピアノ演奏そのものの演奏技術が向上したのか」「必要とする提示情報は熟達度に応じて変化するのか」といった提案する学習支援システムを使用した場合における熟達化プロセスは十分調査できていなかった．そこで本研究では，この問題を解決するために，ピアノ学習支援システムの熟達化プロセスの詳細な調査を目的とする．評価実験では，提案する学習支援システムを使いながら 5 日間かけて 6 人の被験者に 1 日 30 分間課題曲を練習してもらい，実験日ごとの熟達度を調査するためにシステムの補助がない状態で課題曲を演奏してもらった．練習中に記録した打鍵データや視線データなどをもとに熟達化プロセスを検証した．

1　はじめに

ピアノ演奏では，譜読み，指示されている鍵への正確な打鍵，適切な運指（指使い），リズム感覚，打鍵の強弱，テンポなど，さまざまな技術が求められ，それらの修得には長期間の基礎的な練習を必要とする．ピアノ演奏には多大な時間と労力を必要とするため，敷居の高さに利用を断念したり，習熟効率の低さから挫折してしまう演奏者が後を絶たない．特に初心者にとって，譜面上の音符および運指を見て，音符から鍵盤上の打鍵位置をイメージし，指示された運指で弾くという一連のプロセスは最初に立ちはだかる難関で，このプロセスに対する労力や精神的負荷の軽減が楽器演奏を楽しめ長続きさせる秘訣であるといえる．演奏初期段階（ピアノ初心者が初見の楽曲に対して運指や打鍵位置を覚えるために練習している段階）における敷居を下げるために，筆者らの研究グループは，鍵盤上部に設置したプロジェクタを用いて鍵盤上や鍵盤の周囲に打鍵位置情報など演奏補助情報を投影するピアノ学習支援システムを構築してきた [1]．評価実験から比較手法である「光る鍵盤の学習方法 [2, 3]」と比較して効果的に学習できることが明らかになった．しかし,「課題曲を効率的に学習できただけか，あるいは，ピアノ演奏そのものの演奏技術が向上したのか」「必要とする提示情報は熟達度に応じて変化するのか」といった提案する学習支援システムを使用した場合における熟達化プロセスは十分調査できていなかった．

そこで本研究では，この問題を解決するために，ピアノ学習支援システムの熟達化プロセスの詳細な調査を目的とする．

評価実験では，提案システムを使いながら 5 日間かけて 6 人の被験者に 1 日 30 分間課題曲を練習してもらい，実験日ごとの熟達度を調査するためにシステムを利用しない状況で練習後に課題曲を演奏してもらった．また，5 日間かけて練習してもらう課題曲とは異なる曲を実験初日と 5 日目に演奏してもらい，提案システムを使って練習したときのピアノ演奏技術の熟達度を調査した．さらに，打鍵タイミングや打鍵位置といった演奏データだけでなく，視線追跡装置を用いて鍵盤上や鍵盤の周囲のどこを見ていたか計測した視線データにより，熟達化プロセスにおける重要情報の変遷を分析する．

以下，2 章で関連研究について説明し，3 章で実験計画について述べる．4 章で評価実験について述べ，最後に 5 章で本研究のまとめを行う．

2　関連研究

これまでピアノ学習の支援につながる試みはいくつか行われている．蓄積した演奏データから演奏者の苦手な奏法を割り出し集中的にトレーニングするシステム [4] や，演奏を自動的に評価しアドバイス文や誤りを譜面上に提示 [5] するシステムがある．これらは，打鍵ミス，打鍵の強さなどを主に打鍵情報から評価している．Piano Tutor[6] は演奏追従認識による自動譜めくり機能や，ビデオや音声による模範演奏の提示や，演奏者の演奏データを解析し改善点をテキストなどで指示する機能などをもつ．先生と生徒のレッスン支援 [7] として，音量の変化やテンポ，スタッカートやレガートといったアーティキュレーションの具合等を示すシステムが提案されている．打鍵すべき鍵，運指，手本映像を表示するキーボードやソフトウェア [1, 2, 3] がある．これらはいずれも打鍵情報から演奏を評価し学習目的に必要な情報を提示している．本研究では，視線データと打鍵データを用いて演奏を評価している点で異なる．

学習者の視線の動きを観測する研究 [8] はいくつか行われている．特に，川崎ら [9] はピアノ学習者の視線の動きを観測し，プロフェッショナルのピアニストと初心者との視線の動きの違いを解析している．本研究は初心者を対象としており，対象者という観点では川崎らの研究と類似する部分があるが，本研究のように鍵盤上にアノテーションが投影されるピ

Copyright is held by the author(s).

* 公立はこだて未来大学

アノではなく，川崎らは一般的なアコースティックピアノを利用しているため，実験環境は大きく異なる．また，川崎らは楽譜上の視線追跡だけであったが，本研究では楽譜および鍵盤の視線を追跡しているという点でも異なる．

3 実験計画

筆者らの研究グループが提案するピアノ学習支援システムは，ピアノ初心者を対象としており，五線譜やシステムが生成する補助情報を活用しながら学習者はある楽曲を一から練習し，できるだけ速く習熟し，最終的にシステムの補助なしで演奏できるようになることをめざしている．このために，鍵盤上に次に演奏する打鍵位置や運指番号，譜読みを補助するために鍵盤の上部に提示した五線譜の音符とそれに対応する鍵を結ぶ線など多彩な情報を提示している [1]．1 章で述べたように，本研究では提案するピアノ学習支援システムを使用して学習したときにおける熟達化プロセスを調査するために詳細な評価実験を行う．

3.1 実験方針

実験方針として以下があげられる．

ピアノ演奏技能の獲得　提案するピアノ学習支援システムは，30 分間の評価実験から比較手法である「光る鍵盤の学習方法」と比較して打鍵ミス数が減り，提案するピアノ学習支援システムを使用しながら練習すれば効率的に課題曲を弾けるようになる．しかし，「新規の楽曲を提案システムの補助なしで演奏できるようになる」といったピアノ演奏の汎用的な技能を修得できたかは検証されていない．

そこで，本研究では提案する学習支援システムを使って練習する前と，提案システムを使った後で，新規の楽曲をそれぞれ演奏してもらう．その演奏精度を計測することで，提案システムが学習者のピアノ演奏技能の向上に貢献するか調査する．

提示情報の重要度の変遷　提案する学習支援システムは演奏を補助するさまざまな情報を提示しているが，例えば，「未熟な段階では打鍵位置が明確にわからないため手元の打鍵位置情報を頼りに演奏するが，熟達するにつれて打鍵位置情報に頼らず楽譜を見る」など熟達度によって情報の重要度は変わってくると思われる．

そこで，より効率的な学習支援システムの提案をめざして，熟達するにつれ必要とされる提示情報の変遷を明らかにする．具体的には，提案システムを使いながら 5 日間かけて 6 人の被験者に 1 日 30 分間課題曲を練習してもらい，実験日ごとの熟達度を調査するためにシステムを利用しないで課題曲を演奏してもらう到達度テストをうけてもらう．また，打鍵タイミングや打鍵位置といった演奏データだけでなく，視線追跡装置を用いて鍵盤上や鍵盤の周囲のどこを見ていたか視線データを計測する．これらのデータおよび実験中の観測から，熟達化プロセスにおける重要情報の変遷を定量的・定性的に分析する．なお，実験結果を明確にするために，打鍵位置をシステムの補助なしでできるだけ速く習熟できることを提案システムの目的とし，この目的に特に有効な情報であった「鍵盤上への打鍵位置および運指番号」・「楽譜」・「楽譜上に提示する現在の演奏位置」のみ演奏補助情報として提示する．

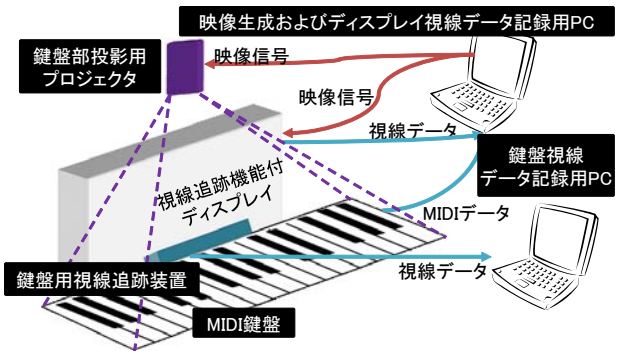


図 1. システム構成

3.2 実験システム

実験で使用した学習支援システムのシステム構成を図 1 に示す．鍵盤上部に設置したプロジェクタを用いて鍵盤上に演奏補助情報を提示する．また，演奏者の前方に視線追跡機能付ディスプレイを配置し，プロジェクタと同様に演奏補助情報を提示している．システムは，MIDI データ（打鍵位置や打鍵強度）を入力とする．さらに，鍵盤部の視線を追跡するために鍵盤上部に視線追跡装置を設置し，演奏の様子を記録するためにビデオカメラを設置した．

映像生成およびディスプレイ視線データ記録用の PC として SONY 社の VGN-SR94VS を使用し，鍵盤部視線データ記録用の PC として SONY 社の VPCSA を使用した．また，MIDI 鍵盤として CASIO 社の PriviA PX-110 を使用した．鍵盤部の視線追跡装置として Tobii 社の X1 ライトを使用し，ディスプレイ部の視線追跡装置として Tobii 社の T60 を使用した．プロジェクタとして BenQ 社の MP776ST を使用した．プロジェクタの鍵盤投影領域は 6 オクターブ（72 鍵）で，プロジェクタの映像がよく見えるように黒鍵を白く塗った．PC 上のソフトウェアの開発は，Windows 7 上で Microsoft 社の Visual C++ 2010 を用いて行った．

3.3 提示コンテンツ

図 2 を用いてシステムが提供したコンテンツについて説明する．図中の番号は，以下の箇条書き番号に対応している．

1. 既存の紙媒体の楽譜と同様の楽譜を提示する．
2. 現在の演奏位置を提示する．これにより，学習者は現在どこを演奏しているか直観的に理解できる．正しい鍵を弾いたときのみ演奏位置は進むようになる．
3. 次に演奏する鍵上に色付枠を提示する．運指情報は，運指番号（親指から小指にかけて 1 から 5 の番号がそれぞれ割り当てられている）ごとに対応している輪郭の色や，鍵上に運指番号を提示することで示す．これにより学習

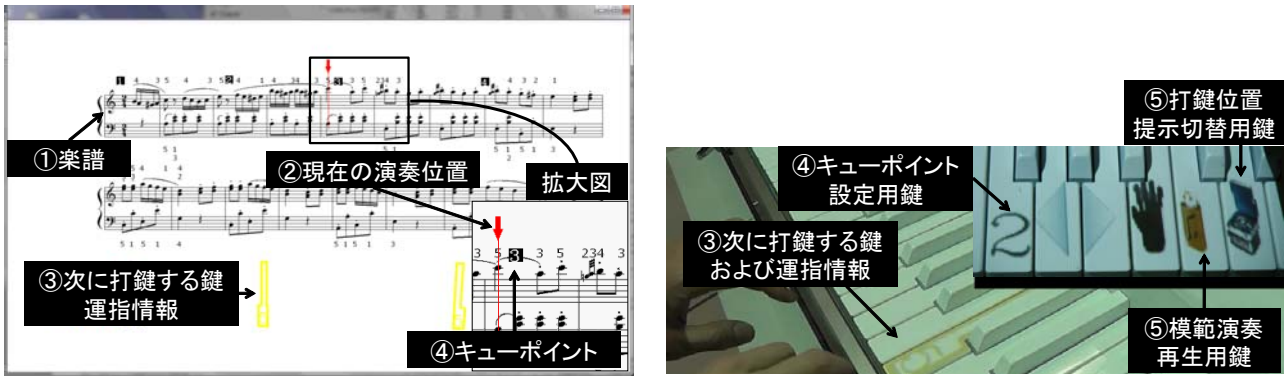


図 **2.** 提示コンテンツ

者は容易に打鍵位置や運指を把握できる．また，ディスプレイにも鍵盤と同様の打鍵位置情報や運指情報を提示する．ディスプレイには鍵の枠しか提示されないため，鍵盤上に提示された情報と比較して得られる情報は少なく，直観性に欠ける．

4. 楽譜上に表示されている番号付きの黒塗りの四角形は，現在位置を変更するキューポイントである．これは，学習者が集中的に練習したい場合や，途中から演奏したい場合に有効である．キューポイントを切り替えるアイコンを，演奏で使用しない鍵の鍵盤上に投影し，その鍵を打鍵することでユーザが選択的に利用できるようにする．これをキューポイント設定機能と呼ぶ．

5. 提案システムは打鍵位置および運指情報を提示するかどうかを切り替える機能（打鍵位置提示切替機能）をもち，この機能を操作するアイコンを演奏で使用しない鍵の鍵盤上に投影する．打鍵位置提示切替は，割り当てられた鍵を押すごとにトグル式で切り替わる．また，提案システムは模範演奏を再生する機能（模範演奏再生機能）をもち，同様に演奏で使用しない鍵の鍵盤上に模範演奏再生機能に対応するアイコンを投影する．学習者が模範演奏再生機能に割り当てられた鍵を押下すると模範演奏の再生が始まり，再生中にその鍵を再度押下すると模範演奏を途中で止められる．なお，本論文では，キューポイント設定機能・打鍵位置提示切替機能・模範演奏再生機能をまとめて付加機能と呼ぶ．

4 評価実験

評価実験では，演奏初期段階（ピアノ初心者が初見の楽曲に対して運指や打鍵位置を覚えるために練習している段階）における提案システムを用いた際のピアノ演奏に関する熟達化プロセスを，打鍵データや視線データをもとに分析する．

4.1 実験の手順

実験の手順を以下に示す．

被験者 実験に参加した被験者は 6 名で，五線譜がほとんど読めない鍵盤経験歴のない 20 代から 40 代の成人である．なお，各被験者にはあらかじめ楽譜上に書かれている音符の意味や，各種機能の使い方を説明した．

課題曲 熟達化プロセスを評価するために，W. A. Mozart のトルコ行進曲（最初から 17 小節目まで両手）を演奏してもらった．また，ピアノ演奏技術の熟達度を評価するために，J. S. Bach のメヌエット（BWV Anh. 114）の最初から 8 小節までを両手で演奏してもらった．

実験方法 実験では，「トルコ行進曲を 30 分かけて練習し，到達度テストとして通し演奏（最初から最後まで一通り演奏すること）を行う」という試行を 1 日 1 回行った．これを連続 5 日間かけて繰り返した．

ピアノ演奏技能の上達を検証するために，「初日のトルコ行進曲練習前」および「5 日目のトルコ行進曲到達度テスト後」にメヌエットの到達度テストを行った．練習中および到達度テスト中は，視線計測装置が生成する視線データ，MIDI 鍵盤が生成する打鍵データをシステムに記録し，演奏中の様子をビデオカメラで記録した．なお，メヌエットの到達度テストは初日と 5 日目の両方に行ったが，被験者は初日に行ったメヌエットの演奏を覚えておらず，5 日目のメヌエットの到達度テストは，初日のメヌエットの到達度テストに影響はない．

到達度テストでは，テスト中で計測している打鍵データをもとに打鍵ミス数を計測した．通し演奏時は，両手法において前面にある楽譜のみ（現在の演奏位置も提示しない）提示した．また，誤打鍵（間違えて打鍵した場合），未打鍵（打鍵しない場合），余打鍵（余分に打鍵した場合）を打鍵ミスとみなした．

被験者への指示 30 分間の練習では「自然なテンポで譜面をみながらミスなく弾けることを意識して，機能を自由に使ってもらって 30 分間練習してください．また，この後，到達度テストを行います．到達度テストは打鍵位置の情報などシステムからの補助情報がない状態で最初から最後まで弾いてもらいます．実験中に質問があれば何でも聞いてください」と指示した．また，到達度テストでは「今から到達度テストを行います．最初から最後まで模範演奏にでき

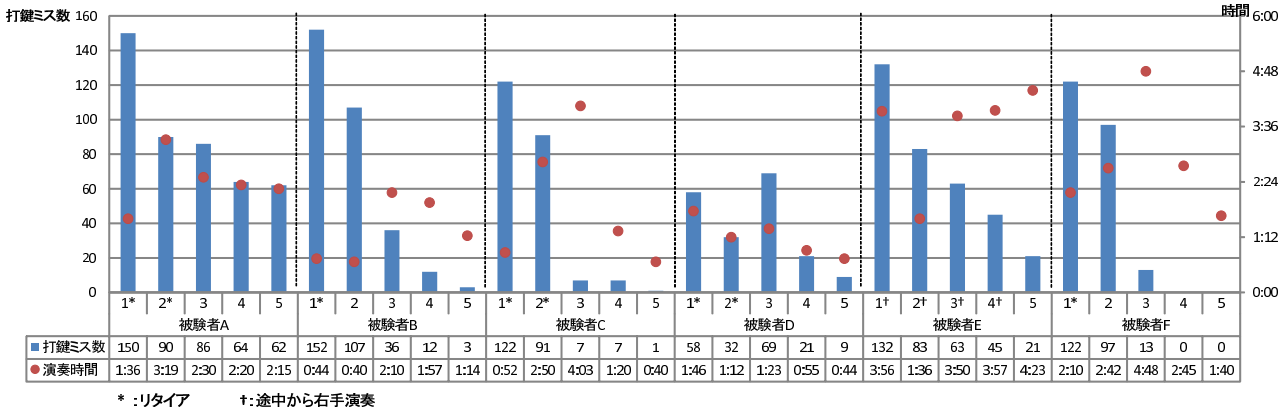


図 3. トルコ行進曲の到達度テスト - 打鍵ミス数および演奏時間

るだけ近いテンポでミスなく弾いてください．制限時間は 5 分間です．わからないところがあれば飛ばしてもらってもかまいませんし，これ以上演奏できなければ言ってください．たとえ間違っても弾き直しをしないようにしてください」と指示した．なお，難しすぎて練習を放棄した被験者はいなかった．

4.2 ピアノ演奏技能の獲得 - 実験結果と考察

トルコ行進曲の到達度テストにおける打鍵ミス数および演奏時間（課題曲を「最初から最後まで」あるいは「最初からリタイア」するまでにかかった時間）を図 3 に，メヌエットの到達度テストにおける打鍵ミス数および演奏時間を図 4 に示す．

全体的に，トルコ行進曲の到達度テストにおける打鍵ミス数は実験が進むにつれて減少し，途中リタイアした場合を除き演奏時間も減っている．また，メヌエットにおいては，打鍵ミス数が減少し演奏時間が短くなった被験者がいた一方，上達が見られなかった被験者もいた．この原因として被験者ごとの学習方略の違いが考えられる．

提案する学習支援システムは，打鍵位置情報を視覚的に提示することで，打鍵位置をできるだけ速く学習できることをめざしている．また，打鍵位置提示切替機能，模範演奏再生機能，キューポイント指定機能といった付加機能を使うことで学習の促進をねらっている．打鍵位置情報はデフォルトで提示されるが，これらの付加機能は選択的に利用される．したがって，提案する学習支援システムは「提示された打鍵位置情報を活用し打鍵位置を学習し，学習効率を高めるために能動的に付加機能を活用する」という学習方略を理想としている．

被験者の学習方略は，「学習支援システムが提供する打鍵位置情報を利用し打鍵位置を学習するものの，付加機能をほとんど使わない学習方略（打鍵位置情報依存型）」「学習支援システムが提供する打鍵位置情報を利用し打鍵位置を学習すると同時に，付加機能を積極的に利用し熟達度に応じて学習方法を柔軟に構築する学習方略（システム方略適合型）」「学習支援システムが提供する打鍵位置情報に頼らず，自身の過去の成功体験をもとに練習する学習方略（システム方略不適合型）」の 3 つの大きく分類できる．以降，各学習方略について考察する．

打鍵位置情報依存型　打鍵位置情報依存型の被験者 A は，他の被験者と比較して最も熟達度が低かった被験者である．被験者 A の付加機能を使用回数は，他の方略の被験者と比べて少なく，打鍵位置情報を利用しながら練習していた．実験 3 日目以降に打鍵位置提示切替機能を使用する様子が観測されるが，30 分間の訓練中 2 分〜3 分と短期間に，あるフレーズに対して打鍵位置情報を見ずに弾けるか確認をしただけである．被験者 A は実験が進むにつれてトルコ行進曲の到達度テストの打鍵ミス数は少なくなっている一方，5 日目のメヌエットの打鍵ミス数は初日と比べて増えている．被験者 A は「PC 用キーボードのタッチタイピングを修得するときに，タッチタイピング練習ソフトウェアで繰り返し練習をして習熟していった過去の成功体験があり，タッチタイピングの練習と類似している実験の学習方法は，学習しやすかった」とコメントしている．また，「タッチタイピングでは，ミスなくタイピングできた単語もできなかった単語も区別なく均等に練習したため，今回の実験においても，間違えやすいところを集中的に訓練するといったことはやらなかった」とコメントしている．したがって，被験者 A は学習支援システムの学習方法に肯定的でストレスなく学習できていたが，学習支援システムを十分使い切れていたとはいえない．

一方，被験者 B は被験者 A より付加機能を使っておらず，システムが提示する打鍵位置を常に使いながら練習していた．被験者 B も，繰り返し練習により新たな技能を習得した過去の成功体験があり，このような学習方略にいたった．被験者 B は被験者 A より受動的な学習方略であったにも関わらず，被験者 A と異なり実験が進むにつれて着実に熟達している．被験者 A の熟達が伸び悩んだ要因の調査は今後の課題である．

システム方略適合型　被験者 C や被験者 D は打鍵位置情報をもとに打鍵位置を学習すると同時に，付加機能を使って難しい箇所を集中的に練習したり，打鍵位置情報を使わずに修得できているか確認しな

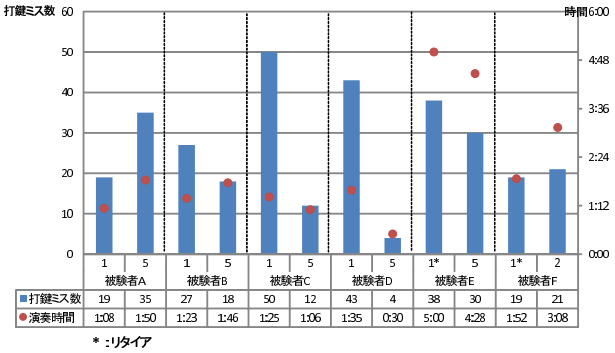


図 4. メヌエットの到達度テスト - 打鍵ミス数および演奏時間

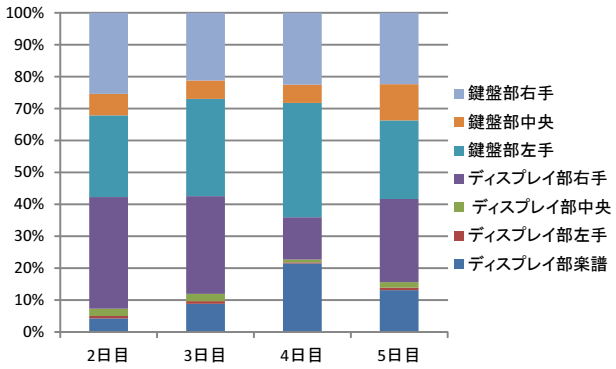


図 5. トルコ行進曲の練習中 - 実験日ごとの視線分布

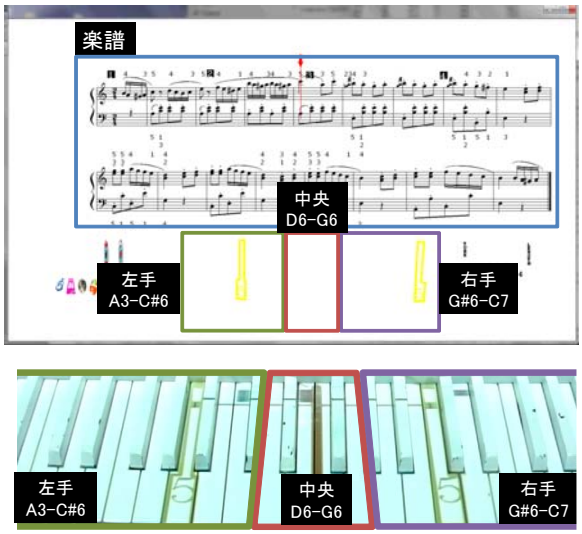


図 6. ディスプレイ部と鍵盤部における視線領域

がら練習していた．特に被験者 C は，打鍵位置提示切替機能やキューポイント選択機能を何度も使用し，集中的に部分練習していた．実験が進むにつれて着実に到達度テストの打鍵ミス数が少なくなり，途中リタイアを除き課題曲を弾き通すためにかかった時間も短くなり演奏テンポが速くなっていった．メヌエットにおいても実験初日と実験 5 日目では打鍵ミス数は減少し，演奏テンポは速くなった．これにより，提案する学習支援システムは課題曲を効果的に学習できるだけでなく，ピアノ演奏技術の向上にも貢献できたことがわかる．メヌエットの到達度テストにおいて，システム方略適合型被験者における打鍵ミス数増加率（「実験 5 日目の打鍵ミス数」を「初日の打鍵ミス数」で割った値）が最も小さく，「システム方略適合型被験者における打鍵ミス数増加率」および「システム方略不適合型被験者における打鍵ミス数増加率」間で有意水準 5%で有意差 (p 値=0.047) が観測された．その他の組み合わせにおいて有意差は観測されなかった．

システム方略不適合型　システム方略不適合型の被験者 E や被験者 F の特徴として，実験が進むにつれてトルコ行進曲の到達度テストの打鍵ミス数は減少する一方，メヌエットの到達度テストにおける打鍵ミス数は，実験初日と実験 5 日目と比べて変化は小さい．

被験者 E は，提示された打鍵位置情報にはできるだけ頼らず，模範演奏を聴きこんだり，楽譜をじっくり見ながら打鍵位置を模索していた．被験者 C は，途中まで模範演奏を聴いて練習するという訓練を何度もやっていたため模範演奏再生回数は増えたが，被験者 E は模範演奏を最初から最後までじっくり聴いていた．被験者 E は声楽を学んでいた経験があり，「声楽の新曲に取り組むときその曲を繰り返し聴いて覚える」という成功体験をもつ．今回の実験においても，自身の成功体験を適用し聴覚中心の学習方略で取り組んでおり，提案する学習支援システムの学習方略に適さなかった．

被験者 F も被験者 E と同様に，できるだけ打鍵位置情報を使用しないで訓練していた．被験者 F は，「トルコ行進曲の到達度テストでどうすれば打鍵ミスなく弾けるようになるか」を中心に考えるゴール指向型の学習方略で取り組んでおり，到達度テストでは打鍵位置情報が提示されないことから，打鍵位置情報を見ずに演奏できるように練習していた．このため，到達度テストにおけるトルコ行進曲の打鍵ミス数は減少しており，被験者 A〜被験者 F の中で最初に打鍵ミス数が 0 になったが，楽譜を中心に打鍵位置を考えながら弾いていたため演奏テンポは他の被験者と比べると遅い．また，トルコ行進曲用に組み立てられた学習方略はトルコ行進曲の到達度テストにおいて効果的であったが，5 日間の練習後に演奏したメヌエットの到達度テストでは，5 日間の練習で得られたスキルを活かせず，練習前に演奏したメヌエットの到達度テストと変わらない結果となった．これにより，ピアノ演奏においては楽譜や鍵盤をよく理解する認知的側面の熟達も重要であるが，楽譜上の音符から打鍵位置を正しく演奏できるようになるといった身体行動の熟達も重要であるといえる．これらがバランス良く熟達することでピアノ演奏技術そのものも向上すると思われる．

4.3　提示情報の重要度の変遷 - 実験結果と考察

「熟達度」と「提示情報の重要度の変遷」を分析するために，30 分間のトルコ行進曲の練習における

視線の分布を調査した．上述したように学習方略の違いが見られたため被験者ごとに視線分布を分析することが望ましいが，紙面の制約のため，被験者 B に着目して詳細に検討する．被験者 B を選定した理由は，実験日ごとの練習においてシステムが提供する付加機能を使用せず，提示される打鍵位置をもとに打鍵位置を覚える訓練淡々と繰り返しており，どの打鍵に対してどういう視線であるか特定しやすいためである．

図 5 に実験日ごとの視線分布を示す．実験で得られた視線データを図 6 に示す視線領域にマッピングすることで図 5 の視線分布を求めた．ディスプレイには打鍵位置だけを示す仮想的な鍵盤が提示されており，ディスプレイ上の仮想鍵盤と実際の鍵盤を演奏する手を基準に右手・左手・中央に領域をわけた．トルコ行進曲は右手で弾く領域と左手で弾く領域が明確にわかれており，右手と左手の境界線は E6（E は音名をおよび 6 はオクターブ数を意味する）である．視線認識精度の問題から，右手と左手の境界付近は明確に右手左手どちらの領域を見ているか特定できないため，中央という領域を設けた．また，実験日が 2 日目から始まっている理由はシステムの不具合により初日の鍵盤部の視線がとれていなかったためである．

図 5 に示すように，熟達度が低い 2 日目ではディスプレイ部楽譜を見る頻度は少なく，鍵盤あるいはディスプレイ部右手を中心に見ている．4 日目になると楽譜を見る頻度はあがり，ディスプレイ部右手や鍵盤部右手を見る頻度は少なくなった．これにより熟達するにつれて，鍵盤に提示された打鍵位置情報や，ディスプレイ上に提示された打鍵位置情報は不要となり，楽譜情報が重要となっていることがわかる．また，右手の打鍵位置に関してはディスプレイ部右手を見ており，左手の打鍵位置に関しては鍵盤上の打鍵位置を直接見ている．被験者 B は「余裕があれば演奏中に楽譜を見つつ練習したいと思っており，できるだけ視線移動を少なくするために，右手の演奏はディスプレイを見ていた」とコメントしている．また，被験者 B は「右手は利き手であり，左手と比べて鍵間の距離が鍵盤を直接見ないでも予測できた」と述べており，右手は左手と比べて鍵盤のブラインドタッチ能力が高く，このような視線の動きになったと考えられる．5 日目は，4 日目とくらべて被験者の調子が悪かったと思われる．このため，5 日目の視線は，ミスを避けるために楽譜を見る頻度が低くなった．しかし，5 日目におけるトルコ行進曲の到達度テストは 4 日目の同楽曲の到達度テストより成績が良くなっている．5 日目の 30 分間の練習においても着実に熟達度が向上したことがわかる．

5 まとめ

本研究では，学習者の視線に注目し，ピアノ学習支援システムの熟達化プロセスの詳細な調査を行った．成人した 8 名の鍵盤楽器初心者に，学習支援システムを使いながら 1 日 30 分間トルコ行進曲を 5 日間かけて練習してもらい，実験日ごとにシステムの補助を使わずに到達度テストを行ってもらった．

実験結果より，成人学習者は過去の成功体験から学習方略をもっており [10]，いずれの被験者も自身の学習方略を基準にシステムを利用していた．自身の学習方略とシステムの学習方略が適合した被験者は，効率的に習熟でき，ピアノ演奏そのものの演奏技術の獲得までいたった．学習方略が部分的に適合した被験者は学習効果にばらつきが生じ，システムの学習方略に適合しなかった被験者はピアノ演奏技術の獲得まではいたらなかった．

さらに，1 名の被験者に対して解析した実験日ごとの視線データの結果から熟達するにつれて被験者が必要とする情報は打鍵位置情報から楽譜情報に遷移していることが明らかになった．

今後は，音符に着目した提示情報の重要度の変遷の調査，他の被験者を対象とした実験データの解析，視線データを考慮した熟達度の算出，熟達度に応じて補助情報を適応的に提示する学習支援システムの構築などがあげられる．

参考文献

[1] 竹川佳成, 寺田 努, 塚本昌彦：運指認識技術を活用したピアノ演奏学習支援システムの構築, 情報処理学会論文誌, Vol. 52, No. 2, pp. 917–927 (2011 年).

[2] CASIO：光ナビゲーションキーボード：http://casio.jp/emi/key_lighting/.

[3] ヤマハ株式会社: 光る鍵盤 EZ-J210：http://www.yamaha.co.jp/product/piano-keyboard/ez-j210/index.html.

[4] 大島千佳, 井ノ上直己：不得手要素を克服させるピアノ学習支援システムにむけて, 情報処理学会研究報告 (音楽情報科学研究会 2007-MUS-71), Vol. 2007, No. 81, pp. 185–190 (2007 年).

[5] 森田慎也, 江村伯夫, 三浦雅展, 秋永晴子, 柳田益造：演奏特徴の強調およびアドバイス文呈示によるピアノ基礎演奏の独習支援, 日本音響学会平成 20 年度秋季研究発表会, pp. 933–934 (2008 年).

[6] R. B. Dannenberg, M. Sanchez, A. Joseph, P. Capell, R. Joseph, and R. Saul: A Computer-Based Multi-Media Tutor for Beginning Piano Students, Journal of New Music Research, 19 (2-3), pp. 155–173, 1990.

[7] S. Smoliar, J. Waterworth, and P. Kellock: pianoFORTE: A System for Piano Education Beyond Notation Literacy, Proceedings of the Third ACM International Conference on Multimedia, pp. 457–465 (1995).

[8] Wikipedia, Eye movement in music reading, http://en.wikipedia.org/wiki/Eye_movement_in_music_reading

[9] 川崎智子：ピアノ演奏時における読譜の為の眼球運動, 三重大学教育学部研究紀要教育科学, 第 33 巻, pp.49-66 (1982 年).

[10] W. W. Lee, D. L. Owens, 清水康敬：インストラクショナルデザイン入門–マルチメディアにおける教育設計, 東京電機大学出版局, p.38.

LumoSpheres: 複数浮遊物体の実時間追跡にもとづく映像投影

山口 裕明　　小池 英樹*

概要. 空中を浮遊する大量の粒子を実時間追跡し, 映像を投影することで, 体積型の立体ディスプレイを実現する構想を提案し, その基礎段階として, 運動している少数の球に実時間追跡にもとづいた映像の投影を行った. 構築したシステムの基本的な流れは, モーションキャプチャを用いて球の三次元位置を計測し, プロジェクタを用いて球に映像を投影するという処理であるが, これらを処理している間も球は空中を移動しているため, その時間分だけ正しい位置に映像が投影されない. この問題に対して, 統計的に位置を推定するカルマンフィルタと運動モデルを適用し, 軌道予測を行うことで安定した映像の投影を実現した. 現段階の応用例として, トスジャグリングの球に映像を投影する DigitalJuggling を実装した. そして, パーティクル型立体ディスプレイのプロトタイプを実装し, 実現するために必要となる要素を議論した.

1 はじめに

プロジェクションマッピングは, 現実の物体の形状に沿ってプロジェクタから映像を投影する技法で, 近年イベントやメディアアートとして盛んに用いられている. 投影対象は机上, 壁, 床, 建築物, 造形物などの様々な形状であるため, その形状に合わせた映像の制作や位置合わせが重要である. そのため, プロジェクションマッピングの投影対象は, 静止している物体や回転などの周期運動をする物体が使用されることが多く, 自由な運動や高速な運動をする物体には一般的に用いられてこなかった.

一方, ビジョンセンサの性能やコンピュータを用いた物体認識の技術は年々進歩している. 物体を追跡するモーションキャプチャシステムは, カメラ数十台を同時に使用することができ, 数百 fps というフレームレートを用いて, 物体の三次元座標を実時間で計測することが可能である.

このようなプロジェクションマッピングの技術と物体認識の技術を合わせて, 運動している物体に対して映像を投影しようというのが本研究の目標である. 静止している物体へプロジェクションマッピングを適用すると, 投影範囲は物体の表面に限定されるが, 運動している物体へ適用すると (物体の表面であることは同様であるが), 物体を動かす範囲だけ投影範囲を空間的に広げることができる. つまり, 映像の表現範囲を平面から空間へ拡張することができる.

このような平面から空間へ映像を拡張するという考えから, 図 1 のようなパーティクル型立体ディスプレイを提案する. パーティクル型立体ディスプレイは, 大量の粒子を空中に飛散させ, それらをビジョンセンサで計測し, それぞれの粒子にプロジェクタから任意の色を投影することを繰り返し行うことで実現する体積型の裸眼立体ディスプレイである. 立体像が結像される領域を囲うように複数台のプロジェクタを設置すれば, 360° どの方向からでも立体像を見ることができる. また, 粒子が質量の小さい物質であれば, 立体像が結像される空間に直接的なインタラクションが可能である.

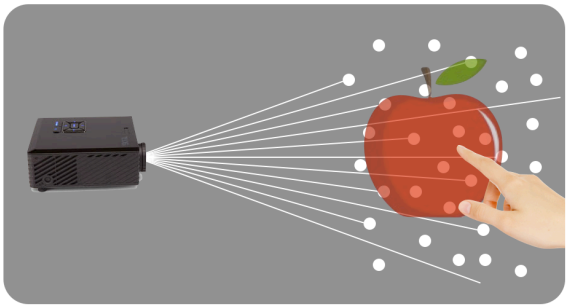


図 1. パーティクル型立体ディスプレイの概念図. 立体像に直接的なインタラクションができる裸眼立体ディスプレイである.

パーティクル型立体ディスプレイは, 空中に飛散している球の位置の計測と, その物体へ映像を投影するというのが基本的な処理となっている. しかし, それらの処理中も球は空中を移動しているため, 計測した球の位置にそのまま映像を投影すると, 装置の処理時間やデータの転送時間だけ投影位置がずれる.

本研究ではパーティクル型立体ディスプレイの基礎段階として, そのディスプレイの最小要素 (平面ディスプレイであれば pixel に該当) である一つの球を用いて, 遅延対策の考案と実装, その評価を行った. 具体的には, 放物線運動を用いた軌道予測を行うが, 単純に運動モデルを用いて軌道予測をすると映像が乱れてしまう現象が発生するため, 統計的に

Copyright is held by the author(s).

* Hiroaki Yamaguchi and Hideki Koike, 電気通信大学大学院 情報システム学研究科 情報メディアシステム学専攻

位置を推定するカルマンフィルタと運動モデルを合わせた軌道予測を行い, 安定した映像の投影を実現した.

現段階の応用として, トスジャグリングの球に映像を投影する DigitalJuggling をアプリケーションとして示す. そして, パーティクル型立体ディスプレイのプロトタイプを実装し, 実現するために必要な要素について探求する.

2 関連研究

Kimura らの “Laser produced 3D display in the air” はレーザの特性を利用した発光を空間中に行うことで, 体積的なディスプレイをつくっている [1]. Butler らの “Vermeer” は特殊なスクリーンを回転させることで体積的な投影対象の空間をつくり, そこへプロジェクタから映像を投影をしている [2]. これらのディスプレイは裸眼で 360° どの方向からも立体像を見ることができるが, レーザや回転機構により立体像が表れる空間が危険であるため, 立体像に対する直接的なインタラクションは工夫が必要である.

FogScreen は空中に霧の層を発生させ, プロジェクタからその層に映像を投影することで映像を結像させている [3]. Nakamura らの “bubble cosmos” は煙を充満させたシャボン玉に映像を投影する [4]. Matoba らの “SplashDisplay” は空中に拭きあげた発泡ビーズに映像を投影する立体的な表現のできるディスプレイである [5]. Barnum らの “A multi-layered display with water drops” は水滴の落下とプロジェクタからの映像を同期させるディスプレイである [6]. 霧ではなく水滴を用いることで, 時分割に水滴を落とす制御が可能となり, 4 層の体積的なディスプレイとなっている. 本研究では, 制御や計測のできる対象として球を用い, 投げ上げや自由落下をしている球に対して映像を投影することで, 層ではなく体積的な映像表現を目指す.

プロジェクションマッピング [7] は机上, 壁, 床, 造形物など, 物体の形状に合わせてプロジェクタから映像を投影する技術であり, 1970 年頃からテーマパークで実用化され, 近年盛んにイベントやメディアアートとして用いられる. Parker の “lumarica” は複数のひもに映像を投影する立体ディスプレイである [8]. Okumura らの “Lumipen” は高速視線制御ユニットと高速ビジョンセンサを用いて, 運動している球に即時的な映像の投影を実現している [9]. 高速な光軸制御により, 球に張り付いたような投影ができるが, 投影対象は高速視線制御ユニットに対して単数の球であり, 奥行き情報を得る場合も工夫が必要である. 本研究では, 複数の球に対して奥行き情報を考慮した映像を投影することで, 体積的な映像の投影範囲を確保する.



図 2. LumoSpheres: 運動している複数の球に映像を投影するシステム.

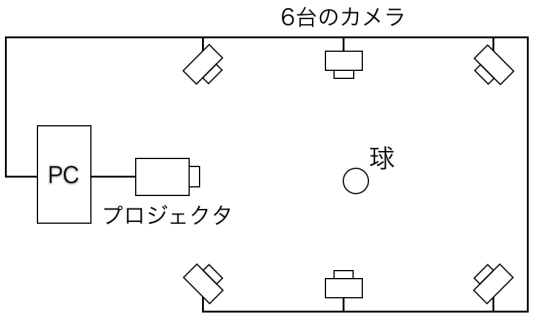


図 3. システム構成 (俯瞰図).6 台のカメラで球を撮影し, PC で処理をした後, プロジェクタから球に映像を投影する.

3 LumoSpheres

運動している複数の球に映像を投影するシステムを「LumoSpheres」と呼ぶ. 図 2 は三個の球に対して動作させている様子である. このシステムは現実空間で運動している球に, 任意の映像を投影することができる. 平面的なプロジェクションマッピングの空間的な拡張, スポーツの映像的な拡張, ビデオゲームなどのエンタテイメントなどに応用が可能である. 図では, 1.5m^3 の空間の中で三個の球にシステムを動作させているが, カメラからみた球の重なり (オクルージョン) が生じない限り三個以上も可能であり, 動作範囲は数十 m^3 まで拡張可能である.

3.1 システム構成

LumoSpheres のシステム構成を図 3 に示す. 映像の投影対象となる球, その球を動かす空間を囲むように配置した 6 台のカメラ (NaturalPoint OptiTrack S250e), 球の位置計測, 追跡を行うソフトウェア (NaturalPoint TrackingTools Ver.2.5), 予測などの処理をする PC(Lenovo ThinkPad T530, Intel Core(TM) i7 2.9 GHz, 8 GB RAM, Windows7 Service Pack 1), 映像を投影をするプロジェクタ (BenQ W1070, 60Hz, 1280 × 720pixel) で構成される. カメラ, PC, ソフトウェア, 予測プログラ

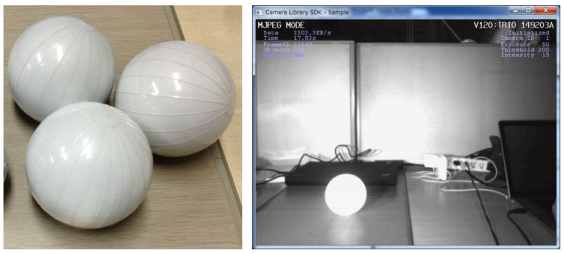


図 4. 再帰反射シートを貼り付けた球 (左) とその反射性能 (右). カメラから照射された赤外線光が再帰反射の性質により球から強く反射される.

ムの間のデータ転送は UDP 通信で行われる.

モーションキャプチャの一般的な用途である, フルボディモーションキャプチャとして用いる場合は, 直径 1cm 程度の再帰反射性マーカを身体全体に装着するが, 本研究では, 任意の大きさの球に再帰反射シートを貼ることで, マーカの代わりとしている. 本システムに用いる再帰反射シートを貼り付けた直径 10cm の球と, その反射性能を図 4 に示す. シートは舟形多円錐図 (地球儀の展開図) に切り, 球に貼り付けている. 一つの球は 180g の重さがあるため, 空中を運動する際の空気抵抗は無視できる.

3.2 プロジェクタとカメラのキャリブレーション

現実空間の球に合わせて適切に映像を投影するためには, モーションキャプチャから取得できる三次元座標とプロジェクタから投影する二次元映像の位置合わせが必要である.

プロジェクタはピンホールカメラと同様のモデルとして考える事ができるため, モーションキャプチャから取得できる三次元座標を (X, Y, Z), プロジェクタから投影する二次元映像の座標を (u, v) とすると, 3×4 の投影変換行列を用いて式 (1) のように表すことができる (式は同次座標系で表している)[10].

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \sim \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

この式は, キャリブレーションにより投影変換行列を求めることができれば, モーショキャプチャ座標 (X, Y, Z) が与えられた時に, プロジェクタから投影する二次元座標のどの位置に球の映像を描けばよいかを示している. 投影変換行列は 12 個の要素からなるため, 座標 (X, Y, Z, u, v) を 6 組以上与えることで求めることができる. 本システムでは, 精度を上げるため座標 (X, Y, Z, u, v) を 12 組与え, 最小二乗法により投影変換行列を求めている. また, 投影変換行列を用いてプロジェクタの外部パラメータを求めることで, モーションキャプチャ座標 (X, Y, Z) からプロジェクタからみた球の三次元位置座標に変換することができる. この変換をすることにより, プロジェクタからみた球までの三次元的な距離がわかり, 映像を描く際のサイズを求めることができる.

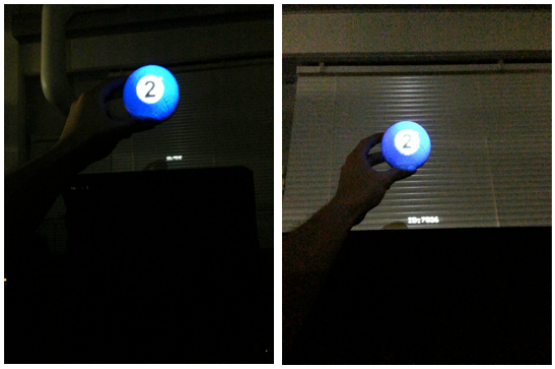


図 5. キャリブレーションによる位置合わせ.(左) と (右) は位置が 1m ほど離れているが, 球の形状に沿った映像が投影される.

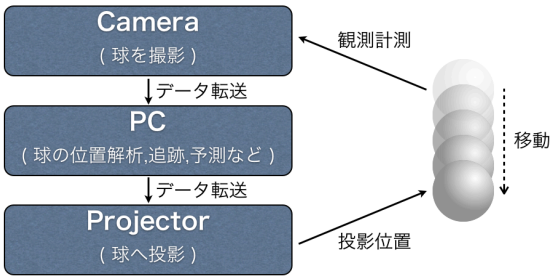


図 6. システムの流れで発生する遅延. 球をカメラで撮影してから, プロジェクタから投影するまでに遅延が発生するため, 軌道予測をする.

キャリブレーション後に映像を投影した様子を図 5 に示す. 現実空間で異なる位置に球を移動しても, 球に合わせた映像が適切に投影される.

3.3 位置推定と運動モデルによる軌道予測

図 6 はシステムの流れを示している. 球をカメラで撮影してから, その球にプロジェクタから映像を投影する過程には, 各装置の処理時間とデータ転送時間が発生する. これらの装置の処理時間とデータ転送時間を合計した時間を, 本研究では遅延時間という. 遅延時間が発生している間も球は運動しているため, 単に取得した三次元位置に映像を投影しても, 球に正しく投影されない. そこで, 運動している球に合うように映像を投影するために, カルマンフィルタを用いて球の状態を統計的に推定し, 運動モデルを適用することで遅延の対策を行う.

カルマンフィルタは, 事前の観測値から現在や事後の確率を最大化するシステムの状態モデルを構築することができる [11]. つまり, 次々と得られる球

の座標データを用いて球の状態 (位置, 速度, 加速度) を統計的に推定することができる. 球の位置ベクトルを式 (2) と表す.

$$p_k = (X_k, Y_k, Z_k)^{\mathrm{T}} \tag{2}$$

このとき, 球の状態を位置, 速度 (一つのドット), 加速度 (二つのドット) で表すと, 状態ベクトルは式 (3) と表す.

$$x_k = (p_k^{\mathrm{T}}, \dot{p}_k^{\mathrm{T}}, \ddot{p}_k^{\mathrm{T}})^{\mathrm{T}} \tag{3}$$

球の運動は等速度運動と仮定した場合の状態モデルを式 (4) に示す (w はシステムノイズ, I は 3 x 3 の単位行列, Δt はデータの取得周期).

$$\begin{bmatrix} p_k \\ \dot{p}_k \\ \ddot{p}_k \end{bmatrix} = \begin{bmatrix} I & \Delta tI & \frac{1}{2}(\Delta t)^2 I \\ 0 & I & \Delta tI \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} p_{k-1} \\ \dot{p}_{k-1} \\ \ddot{p}_{k-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ w \end{bmatrix} \tag{4}$$

観測マトリクスは式 (5) とし球の位置のみを観測する.

$$H = \begin{bmatrix} I & 0 & 0 \end{bmatrix} \tag{5}$$

これらの式を用いてカルマンフィルタの予測フェーズと修正フェーズを繰り返し行うことで, 球の状態を推定する.

カルマンフィルタにより, 現在の状態を推定したら, システム全体の遅延時間を ΔT として, 式 (6) の運動モデルにより, 球の軌道予測をする.

$$p_{k+1} = \frac{1}{2}\ddot{p}_k \Delta T^2 + \dot{p}_k \Delta T + p_k \tag{6}$$

4 投影精度の評価実験

4.1 評価の尺度

球にどの程度の映像が正しく投影されたかを, システム外部からカメラを用いて評価する. 評価の尺度として, カメラ画像からみた球全体の面積 (半球) に対する, 投影映像の面積の割合を投影精度と定義する. 例えば, 投影された面積が球全体の面積の半分ならば, 投影精度は 50 %となる.

4.2 評価のアルゴリズム

評価はカメラのグレースケール画像を以下の手順で処理することにより行う.

1. 球全体の面積を求める:
   Hough 変換を用いて円抽出を行い, 円の面積 (pixel) を算出する.
2. 投影部分の面積を求める:
   1 で円抽出をした領域をしきい値を用いて 2 値化し, 投影部分の面積 (pixel) を算出する.
3. 投影精度を求める:
   (投影部分の面積 / 球全体の面積) を求めることにより, 投影精度を算出する.

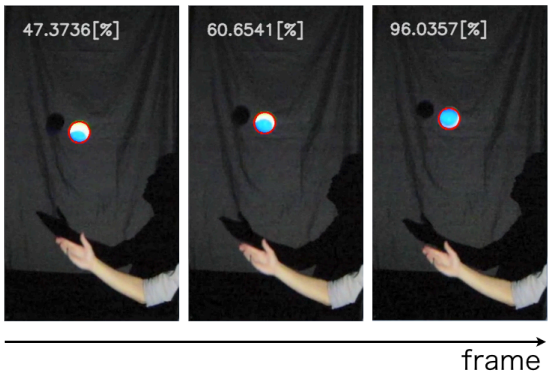


図 7. 投げ上げている最中の数フレームの投影精度. 球全体の面積と投影部分の面積の割合により投影精度が算出される.

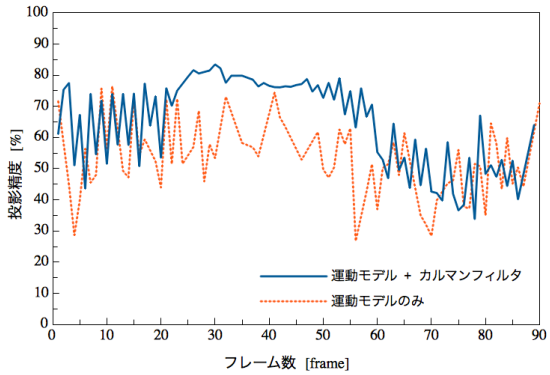


図 8. 投影精度 - フレーム数. カルマンフィルタにより位置推定をし, 運動モデルを適用することで投影精度は安定する.

しきい値は球面積の半分に映像が投影されるように設定し, 球部分のヒストグラムから大津の二値化 [12] を用いて事前に決定する.

図 7 は投げ上げ中のある連続したフレームの円抽出と投影精度を表している.

4.3 評価実験の結果

実験は, 120fps のフレームレートで撮影できるカメラを使用し, 球を投げ上げてからつかむまでの一連の運動を記録した. システム全体の遅延時間は 56ms と設定し単に運動モデルを適用した場合と, カルマンフィルタと運動モデルを合わせた場合の軌道予測を比較する. 遅延時間は球の個数により数 ms 変化するが, 今回は定数として任意に決定した.

図 8 は式 (6) の運動モデルを単に適用した場合と, カルマンフィルタで位置を推定した後に運動モデルを適用した場合の投影精度 - フレーム数のグラフである. カルマンフィルタと運動モデルを合わせた軌道予測では, 投げ上げの最高点 (35frame 前後) で球の速度が遅くなるため, 投影精度は 80%以上となる

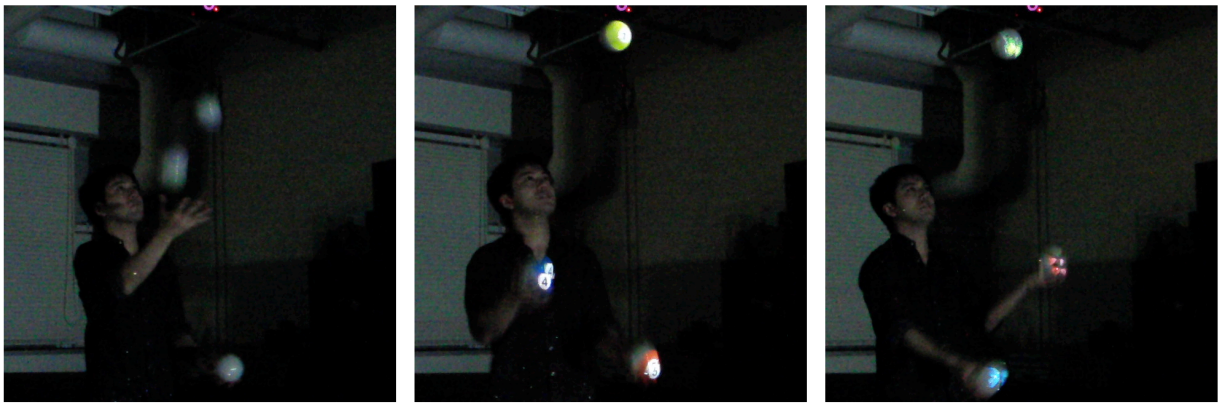


図 9. DigitalJuggling: 本システムをトスジャグリングに応用した例. 映像の投影なし (左), 球状の映像を投影 (中), 非球状の映像を投影 (右). トスジャグリング中も動的に映像を切り替えることができる.

が, 速度が速い投げ上げ直後とつかむ直前では投影精度がおよそ 40 %から 75%となる. 単に運動モデルを適用した軌道予測では, 球がどの位置であろうが, 投影精度はおよそ 30 %から 75%となる. カルマンフィルタは過去のデータを統計的に計算するため, 計測時に発生するノイズの影響に強く, 単純に運動モデルを用いた場合と比較して安定した投影となった. 球の速度が速い時に投影精度が振動するのは, 球は時間的に連続して動いているにも関わらず, プロジェクタが 60 Hz で動作しているのが原因である.

5 アプリケーション

本システムは, プロジェクションマッピングやディスプレイの空間的な拡張, スポーツの映像的な拡張, ビデオゲームなどのエンタテイメントなど, コンテンツ次第で様々な応用が考えられる. ここでは応用例として, LumoSpheres をトスジャグリングへ応用した DigitalJuggling, パーティクル型立体ディスプレイのプロトタイプを示す.

5.1 DigitalJuggling

図 9 に DigitalJuggling の様子を示す. システムをトスジャグリングに応用した例である. 映像を投影しない場合は通常のジャグリング球であるが, 投影する映像次第で様々なテクスチャの球でトスジャグリングをしているように見える. テクスチャは, ビリヤードの球のような球状でも, 蝶のような非球状でもよい. 通常のジャグリング球を用いた場合と異なり, 動的なテクスチャの変化が可能である.

5.2 パーティクル型立体ディスプレイ

図 10 にパーティクル型立体ディスプレイのプロトタイプを示す. 小球は直径 10cm の球と同様に, 直径 4cm のピンポン球に再帰反射シートを貼り付けている. 12 個の小球は, 人の手によって空中に飛散させ, ランダムに色を割り振られた映像が投影されている.



図 10. パーティクル型立体ディスプレイのプロトタイプ.12 個の小球に割り当てた色を投影している.

6 議論: システムの現状と利点, 問題点

LumoSpheres システムの構築を通して, 運動する少数の物体に対して映像を投影する手法を探求した. ここでは, パーティクル型立体ディスプレイを実現する上で必要な要素を議論する.

6.1 プロジェクタのフレームレート

本システムは球の運動速度が速くなるにつれて, 投影映像のちらつきが生じる. これは, 使用したプロジェクタの映像更新フレームレートが 60 Hz であることが原因である. 現実空間の球の運動は連続的であるのに対し, プロジェクタの映像更新は不連続である. 映像が更新された直後は球に映像が合っていても, 次の更新までに球が移動するため映像がちらついてみえる. これを解決するためには, 単純にプロジェクタのフレームレートを上げる, または, Okumura らの研究 [9] のように高速に光軸制御をする装置を用いることで, プロジェクタからの映像

の更新を連続的に近づける必要がある.

6.2 物体のオクルージョン

本システムは少数の球であれば映像を投影できるが, 多数の球になると正確な位置に映像が投影されない. これはカメラの画像で球同士の重なりが生じると, カメラ同士でステレオマッチングができなくなり, 位置を正しく計測できなくなるのが原因である. これを解決するためには, 単純にカメラの台数を増やす, または, 単一の球のみではなく複数の球の相互関係を考慮した予測を行うことで解決できる. これは, 球同士の衝突により生じる急激な運動の変化に対しても有効であると考えられる.

6.3 遅延対策としての予測

装置の処理時間やデータの転送時間が生じない理想的な装置であれば, 軌道予測をする必要はなく, 計測した球の位置に瞬間的に映像を投影すればよい. しかし, 今後もコンピュータやビジョンセンサの進歩はあるものの, 遅延がなくなることはない. そのような意味で, Okumura らの研究 [9] のようにセンサビジョン部分のハードウェアを工夫して高速にするというアプローチは重要であるが, コンピュータ部分のソフトウェアを工夫した軌道予測によるアプローチも意義のあるものであると考える.

6.4 想定される構成と利点

以上を踏まえて, パーティクル型立体ディスプレイの想定される構成は, 連続してパーティクルを空中に飛散させる装置, その装置の周囲を囲むように配置した複数のカメラとプロジェクタ, そしてパーティクル全体をシミュレーションするプログラムが必要である. 本構想の最大の特徴は, 直接的なインタラクションが可能な裸眼立体ディスプレイであるが, その他の利点として, 球に貼り付けた再帰反射シートの性質を利用して, 同じディスプレイ空間で複数人が異なる映像を見ることができる. 再帰反射の性質により複数台のプロジェクタから投影された映像が混合しないことを利用する. これは, 空中で色を発光させるような構造 [1] では実現できない.

7 結論と今後の方向性

運動している複数の球に実時間追跡にもとづき映像を投影することで体積型の立体ディスプレイとなる, パーティクル型ディスプレイを提案し, その基礎段階として, 運動している少数の球に実時間追跡にもとづいた映像を投影するシステム, LumoSpheres を構築した. そして, LumoSpheres システムの応用例として DigitalJuggling を実装し, パーティクル型立体ディスプレイのプロタイプを実装した.

今回の研究では, パーティクル型ディスプレイの基礎段階として, 球一つ一つに映像を合わせて投影する手法を探求した結果, 議論で述べた通り, 様々な問題が浮かび上がってきた. 今後はこれらの問題点を改善しつつ, 多くの物体へ安定した映像が投影できるような手法を更に探求していきたい.

参考文献

[1] Hidei Kimura, Taro Uchiyama, and Hiroyuki Yoshikawa. Laser produced 3D display in the air. In *ACM SIGGRAPH 2006 Emerging technologies*, July-August, 2006.

[2] Alex B., Otmar H., Shahram I., Steve H., David M., David K., and Danny K.. Vermeer: direct interaction with a 360° viewable 3D display. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, October, 2011.

[3] FogScreen. http://www.fogscreen.com/.

[4] Masahiro Nakamura, Go Inaba, Jun Tamaoki, Kazuhito Shiratori and Junichi Hoshino. Mounting and application of bubble display system. In *Proceedings of the 2006 ACM SIGCHI international conference*, June, 2006.

[5] Yasushi Matoba, Taro Tokui, Ryo Sato, Toshiki Sato, Hideki Koike. SplashDisplay: volumetric projection using projectile beads. In *Proceedings of the 2012 Virtual Reality International Conference*, March, 2012.

[6] Peter C. Barnum, Srinivasa G. Narasimhan, and Takeo Kanade. A multi-layered display with water drops. In *ACM Transactions on Graphics (TOG)*, v.29 n.4, July, 2010.

[7] Mine, M., Rose, D., Yang, B., Jeroen van Baar, and Grundhfer, A. Projection-Based Augmented Reality in Disney Theme Parks. IEEE Computer 45, 32-40, July, 2012.

[8] Matthew Parker. Lumarca. In *ACM SIGGRAPH ASIA 2009*, December, 2009.

[9] Kohei Okumura, Hiromasa Oku, and Masatoshi Ishikawa. Lumipen: Projection-Based Mixed Reality for Dynamic Objects. In *Proceedings of the 2012 IEEE International Conference on Multimedia and Expo*, p.699-704, July, 2012.

[10] Richard Hartley, and Andrew Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press, 2003.

[11] Greg Welch, and Gary Bishop. An Introduction to the Kalman Filter. University of North Carolina at Chapel Hill, 1995.

[12] Nobuyuki Otsu. A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-9, No. 1, pp. 62-66, 1979.

画像照合に基づく多メディア間の話題追跡システム

豊田 正史　伊藤 正彦 *　**Cai-Zhi Zhu**　佐藤 真一　宇野 毅明 †　喜連川 優 ‡

概要.　本論文では，Web と放送映像との間の画像の対応関係に基づき話題の追跡を行える新たなシステムを提案する．社会事象を解析する際には，話題がどのメディアから始まったかを同定し，その話題がどのようにメディア間に広がっていったかを分析する必要がある．Web と放送映像との対応を取るため，Web 上のテキストと放送の字幕情報を用いることも可能であるが，字幕の量は十分ではなく，字幕が付けられていない番組も多数存在するため，映像情報の対応関係を解析する事が不可欠となる．提案するシステムでは，ブログ等の Web メディアに投稿された大量の画像から大きな話題になり多数出現した類似画像を抽出し，これが放送映像を起源とするものかどうかを放送映像アーカイブと照合して，話題の時差や露出の差を調査できる手法を実現した．本システムを用いて，大規模なブログアーカイブと放送映像アーカイブを分析した事例を幾つか示すことで本システムの有用性を示す．

1　はじめに

Web はその出現から約 20 年の間に，ニュースサイト，掲示板，ブログ，ソーシャルメディア，コンテンツ共有サイトなど多数のメディアの集合体として進化を続けてきた．現在は，これら Web 上のメディアとテレビや新聞などの従来型マスメディアを包含する多メディアが人々の社会活動を反映し，かつ影響を与えていると言える．こうした状況において，社会事象を分析する際には話題がどのメディアから出現したかを同定し，その話題がどのように多メディアの間に広がっていったかを分析する必要がある．現在においても Web メディアはマスメディア，特にテレビ放送から受ける影響が大きく，Web メディアとテレビ放送の間の話題追跡は重要な課題である．

従来，Web メディアとテレビ放送での話題分析においては，Web 上のブログやソーシャルメディアなどの記事および番組表や字幕など，テキスト情報を用いたものが主であった．しかし，これらのメディアでは積極的に映像・画像を用いることで，文章のみでは伝えることが困難なイベントや事件事故などの様子を伝えており，また字幕の量は十分ではなく，字幕が付けられていない番組も多く存在するため，映像情報の対応関係を解析する事が不可欠となる．

Web 上で話題になった画像が，放送映像中では同じように現れているのか，また，その出現に前後関係はあるのか，話題や画像ごとに出現の仕方に違いや類似性があるのか，逆に特徴のある画像群は具体的にどのような話題を扱っているのか等を調査することにより，以下のような社会事象や多メディア間で及ぼされる影響の分析が可能になる．

- ニュースが各メディアにおいてどのような画像で扱われていたか，メディアにより現れる画像に違いがあるかを調査する．
- イベントや集会などの模様や賑わいを視覚的に把握することで，メディアによる扱われ方の差があるかどうかを調査する．
- 商品画像の出現数の変化から，人気の度合いやデザインの変化を把握する．

本論文では，Web メディアに投稿された大量の画像から大きな話題になり多数出現した類似画像を抽出し，これが放送映像を起源とするものかどうかを放送映像アーカイブを照合して抽出し可視化することで，話題の時差や露出の差を調査できる手法を提案し，それを実現するシステムを実装する．本システムを，大規模なブログアーカイブと放送映像アーカイブに適用し，実際に観測された社会事象の事例を示すことで本システムの有用性を示す．

以下，第 2 節では関連研究について述べ，第 3 節で提案システムの詳細を説明する．第 4 節において本システムを用いた実験結果を示し，第 5 節で本論文をまとめる．

2　関連研究

従来，Web メディアおよびテレビ放送における話題分析においては，単独のメディアを用いたものが多く，複数メディアを扱ったものでも Web 上のブログやソーシャルメディアなどの記事，および番組表や字幕などテキスト情報を用いたものが主であり，大規模な画像照合を用いて多メディア間の情報追跡を行えているものは筆者の知る限りでは存在しない．

Copyright is held by the author(s).

* 東京大学生産技術研究所
† 国立情報学研究所
‡ 東京大学生産技術研究所・国立情報学研究所

Adar らは，検索エンジンの検索クエリとテレビ番組情報の出現時差の解析を行っており [1]，Yang らは，ブログ，新聞，テレビ等の複数メディアにおける話題の時系列パタンを抽出しているが [7]，これらは基本的にテキストデータに現れる語やフレーズの出現頻度を用いたもので画像情報は用いられていない．Luo らは，大規模なニュース画像アーカイブの可視化，および話題分析を行っているが [5]，複数のメディアにまたがる分析までは行えていない．

3 多メディア間話題追跡システム

本論文で提案するシステムは，大規模な Web メディアのアーカイブを用いてユーザの指定した調査対象に関連する記事および画像を抽出し，さらに放送映像のアーカイブを用いて画像照合を行い，抽出した Web 画像に対応する放送映像での露出を抽出することで，これら画像の頻度，時間推移，および出現時差を調査可能にするものである．本システムは，Web メディアアーカイブとして，ブログやソーシャルメディアのように，各記事に対して書き手と書かれた時間が紐付けされた文書アーカイブを想定しており，全文検索インデックスによって記事が検索可能となっているものとする．また，放送映像アーカイブに関しては，画像を入力としてその画像に類似する映像が放送されたショットを検索可能とする検索インデックスを実現して用いている．

ユーザは，調査対象に関する複数の検索語を入力としてシステムに与え，システムは，以下のような手法で，画像抽出，分析，可視化を行う．

1. Web メディアアーカイブを検索し，与えられた検索語を含む記事の集合を取得する．各記事はそれが書かれた時間，書き手，および記事本文を含む．
2. 取得した記事の本文に含まれている画像をすべて取得し，各画像の周辺のテキストを抽出する．
3. Web メディア上で話題となり数多くコピーされた画像を抽出するため，ほぼ同一とみなせる画像のペアを列挙しクラスタリングを行う．この結果得られる *Web* 画像集合は単独の画像が多数出現したことを表し，画像に基づく話題を構成する要素となる．
4. 抽出した各 Web 画像集合を検索クエリとして用いて，放送映像アーカイブを検索し，類似画像が放送されたショットの集合を抽出する．抽出された各ショットの集合を，便宜上 Web 画像集合に対応する放送画像集合と呼ぶ．
5. Web 画像集合を，それらの周辺テキストの類似度を用いてクラスタリングすることで，高頻度で出現する画像の類似集合である *Web* 画像トピックを抽出する．さらに Web 画像トピックに含まれる Web 画像集合群に対応する放送画像集合を集約することで，放送映像トピックを抽出する．こうして得られた Web 画像トピックとそれに対応する放送映像トピックの時間推移を 3 次元空間に可視化し，メディア間におけるトピックの推移およびその差を閲覧探索可能とする．

以下では，ほぼ同一画像のクラスタリング，放送映像検索の実現方法，トピックの 3 次元可視化手法についてより詳細に説明する．

3.1 ほぼ同一画像のクラスタリング

Web 上の画像は，オリジナルの画像に対し拡大，縮小，回転や，切り取り，一部変更などの加工を経て公開されるものが多く，ほぼ同一の画像をクラスタリングするための類似度算出にはこうした加工に対し頑健な特徴量を用いる必要がある．本手法ではこうした状況に適した局所特徴量の 1 つである SIFT 特徴量 [4] を用いている．SIFT を用いた画像類似度算出では，各画像から抽出された多数の SIFT 特徴点の類似度を算出し，画像間でマッチする特徴点の数をカウントする必要がある．画像数は通常多数に及ぶため類似する SIFT 特徴点のペアを高速に列挙する手法を開発した．本列挙手法は，宇野による大規模部分文字列マッチングアルゴリズム [6] を基にノイズとなる特徴点の除去処理を加えて改良を施したものである．

各画像からは，サイズの正規化を行った上で，最大 100 個の SIFT 特徴点を抽出する [1]．SIFT 特徴点は，128 次元の数値ベクトルで表わされるが，各値を中央値以上か，それ以下かの 2 値に量子化し，128 ビットのベクトルに変換する．さらにこの数値ベクトルは 8 個の値からなる 16 のブロックから構成されることを考慮し，各 8 ビットを 1 文字とみなすことで書く特徴点を 16 文字の文字列と考える．宇野の類似文字列列挙手法は，大量の固定長文字列から，ハミング距離が一定以下の文字列ペアを高速に列挙するもので，本手法ではほぼ同一の画像を抽出するため，ハミング距離 1 以下の類似特徴点のペアをマッチしたものと見做して列挙し，画像間でマッチする特徴点の数を，画像ペアをキーとしたハッシュ表を用いて算出する．2 つの画像の類似度は，画像間でマッチする特徴点の数で表わされる．

SIFT 特徴点を用いた類似画像抽出においては，非常に多数の特徴点とマッチする特徴点が，ノイズとして類似度の精度に悪影響を与えることが分かっ

[1] ここでは特徴点の最大数を経験的に 100 個に定めたため，これが最適であるとは限らない．しかし最大数をこれ以上大きくすると，ブログ用の縮小画像や携帯電話などで撮影した小さな画像に関して特徴点の数が多くなりすぎノイズが増すことが実験から分かっている．

ており，本手法では，類似ペア列挙の途中段階で，閾値以上の特徴点とマッチする特徴点をフィルタする処理を加えることで，ノイズに対して頑健な類似画像抽出を実現している．ノイズとなる特徴点の数は非常に少ないため，計算量を大きく削減するものではないが，大量の類似特徴点のペアを出力した後フィルタする必要がなく直接類似画像ペアを出力できるため，中間出力を削減することができる．本手法は，4 百万画像から抽出した 4 億の SIFT 特徴点の集合を用いて類似画像ペア (閾値 1000 の場合，約 1.35 億ペア) を列挙するのに，10GB のメモリと 20 コアを用いて 2 時間程度しかかからないスケーラビリティを持っている．

類似する画像のペアが列挙できると，画像をノード，類似度をエッジとしたグラフを作ることが可能となる．本手法では，このグラフに対して，スタークラスタリング [2] という簡易なクラスタリング手法を適用することで，画像集合を抽出している．スタークラスタリングは，次数の一番高いノードを中心とし，閾値以上の類似度を持つ隣接ノードを，中心と同じクラスタとして出力するという処理を繰り返す単純なものである．本手法では，あるノードについて後からより類似度の高い中心が見つかった場合，元の中心からその中心へ所属を切り替えるという簡易な改良を施して用いている．スタークラスタリングの計算時間はノード数とエッジ数の和に比例しており，大規模なデータをインタラクティブにクラスタリングするのに適している．また，類似度の低いノードをあえてクラスタに所属させることは無いため，ノイズの多いデータのクラスタリングにも適している．こうした理由から本手法ではスタークラスタリングを用いているが，他のクラスタリング手法を用いても良い．

3.2 放送映像の検索手法

放送映像アーカイブ中のショットを任意の画像により問い合わせ可能な，問い合わせ画像中と同一の物体を含む映像を検索可能なシステムを [8] の手法に基づき構築した．まずオフライン処理として，放送映像アーカイブを構成する各映像クリップから，非常に密にフレーム画像を切り出し（例えば 3 フレームごとに 1 フレーム），各フレーム画像から密に SIFT 等の局所特徴量を抽出する（たとえば各フレーム画像から数百点，SIFT もしくは color SIFT を使用）．局所特徴量は階層的 k-means により，100 万コードブックというきわめて疎な離散化を行い，各クリップにつき一つの Bag of Feature (BoF) ベクトルを得る．次いで，検索時のオンライン処理では，与えられた問い合わせ画像から，同様に BoF ベクトルを得る．これとデータベース中の BoF ベクトルとの間で L1 距離を求め，ランキングリストを得ることにより，検索を実現している．検索は転置索引により効果的に高速化可能である．

この方法で実際に検索を行うと，大規模なデータベースが対象となった場合，検索結果にノイズがきわめて多くなってしまうことが判明した．そこで，幾何整合性判定によりリランキングを行う技術 [3] を導入して，精度の向上を図っている．

3.3 多メディア間のトピック可視化・閲覧手法

本システムでは，Web メディアおよび放送映像から抽出されたトピックの時系列推移を可視化することで，多メディア間の影響分析を視覚的に可能にする新たな可視化手法を提案する．

図 1 に本可視化手法の概要を示す．平面上に，時間軸とトピックの軸を持ち，各時間軸上に Web メディアアーカイブから抽出された Web 画像集合，および放送映像アーカイブから抽出した対応する放送画像集合をヒストグラムとして可視化している．画像は単位時間ごと（例えば 1 日ごと）に集約され，その時間に出現した画像を高さ方向に積み上げることで，画像による棒グラフから構成されるヒストグラムとして表現される．

画像の出現の時間差を認識しやすくするため，Web メディアと放送映像の画像ヒストグラムは，図 1 のように並列して配置することもでき，また，図 3 のように上下にヒストグラムを反転させて配置することも可能である．

上記の手法で抽出された Web 画像集合は，さらにその周辺テキストを用いて意味的にトピックにまとめることも可能である．図 1 では，各 Web 画像集合に含まれる画像の周辺テキストをまとめて 1 文書とみなし，tf-idf で重み付けした単語頻度ベクトルを算出し，コサイン類似度を用いて上記のスタークラスタリングを行い，*Web* 画像トピックにまとめている．Web 画像集合のクラスタリングに関しては，スタークラスタリングの閾値をインタラクティブに変更しながらクラスタリング結果を閲覧するシステムを用意し，適切な閾値を選択可能としている (図 2)．こうして得られた Web 画像トピックに対し，放送映像側には各 Web 画像集合に対応する放送画像集合をまとめて放送画像トピックとして表示している．

ユーザは，Web 画像数およびショット数の時系列に対する交差相関係数，コサイン類似度等のレンジをパラレルコーディネートを用いたダイナミッククエリにより対話的に指定することで，様々な時間差をもつ画像クラスタを抽出し，容易に詳細を探索することが可能である．交差相関係数を用いると，指定したメディアで先行して（もしくは遅れて）現れた画像集合を抽出することができ，放送先行もしくは Web メディア先行で流行した話題を抽出することが可能になる．時系列間のコサイン類似度が高い画像集合を抽出することで出現タイミングが似てい

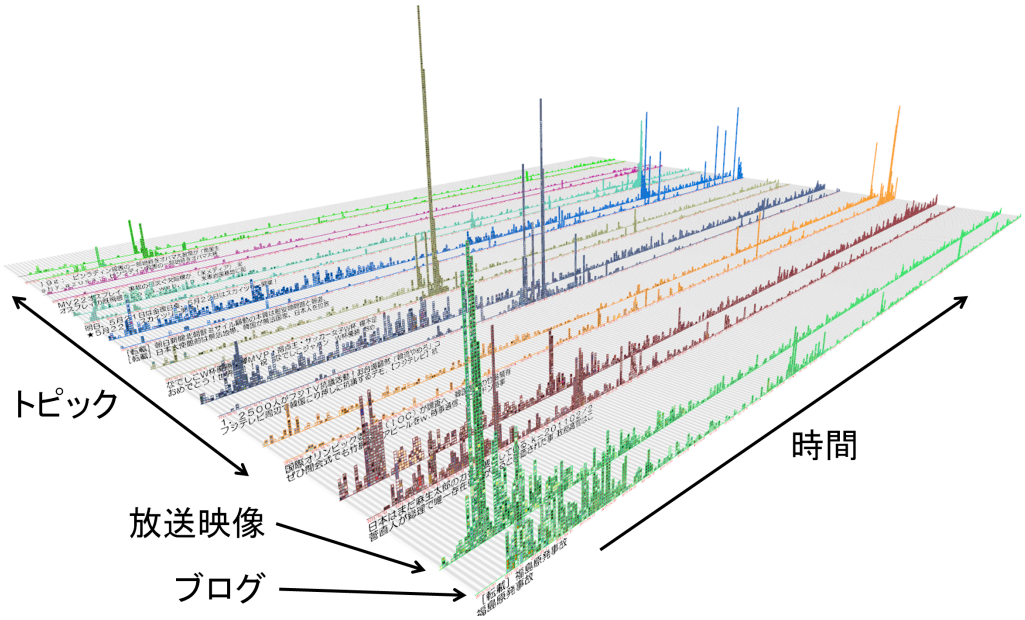


図 1. 複数トピックの推移の概要を可視化

る話題を抽出でき，類似度が低い画像集合からは，放送もしくは Web メディアのみで流行した話題を抽出することができる．

4 大規模ブログアーカイブおよび大規模放送映像アーカイブを用いた実験

4.1 データセット

前節のシステムを実装し，我々が東京大学生産技術研究所において収集している大規模ブログアーカイブおよび大規模放送映像アーカイブを用いた実験を行った．Web メディアとしては，我々が 2006 年から収集している延べ 200 万ブログを対象とした RSS/Atom フィードのアーカイブを用いた．特に，社会分析のターゲットとして 2011 年 3 月から 2012 年 9 月に起きた，東日本大震災およびそれに引き続く原発事故，サッカー女子ワールドカップ，北朝鮮ミサイル問題，ロンドンオリンピックなどを想定して，検索用のキーワードを約 90 個程度選出し，アーカイブから抽出したブログ記事に含まれる約 400 万の画像を実験データとして用いた．

また，国立情報学研究所においてアーカイブした，2011 年 3 月より 2012 年 9 月までに放送された全チャンネルのニュース番組，合計約 6000 時間分を対象として検索インデックスを構築した．検索インデックス構築対象の 6000 時間分のニュースに対し，連続フレーム間の色ヒストグラム差分に基づくショット分割を行い，約 300 万ショットを検出した．これが検索対象の単位となる．さらに，各フレーム単位での詳細な物体検索を実現するため，1 秒おきに約 2000 万キーフレームを抽出した．局所特徴量数は約 200 億となる．こうして構築したデータベースに対し，ブログから想定したターゲットに関連するキーワードに基づいて抽出した 4000 の Web 画像集合 (46000 画像) を問い合わせとして照合を行い，相互の影響解析につながる根拠となるような画像が数多く検出されたことを確認した．

4.2 事例

本節では上述のデータセットを用い，提案システムを用いて抽出した多メディア間の話題追跡の事例を示す．

図 1 は，震災，原発，オリンピック等の検索語でブログアーカイブから抽出した Web 画像トピック，さらに放送映像アーカイブから抽出した対応する放送画像トピック，それぞれをトピック毎，一日毎に集約し時間軸上に可視化している．全体像を俯瞰（図 (a)）することで，トピック間の類似性や，ブログ・放送映像間で同時にバーストしている期間などを観測できる．さらに，時系列に対する交差相関係数，コサイン類似度等から様々な反応差をもつ画像集合を対話的に抽出し詳細を探索することが可能である．

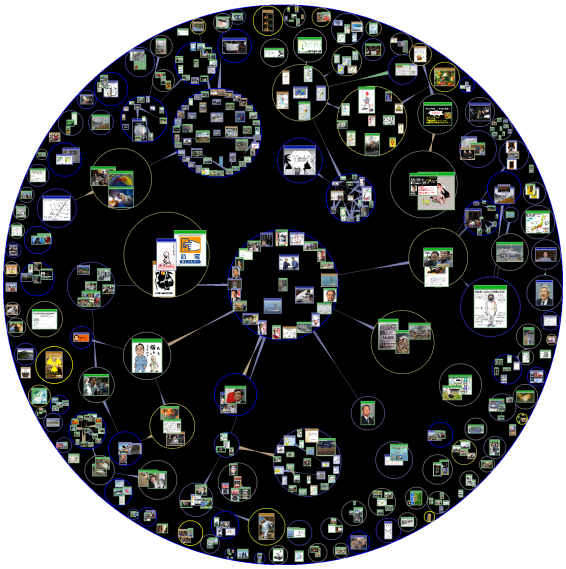


図 2. Web 画像集合のクラスタリング結果の可視化

この中から，東日本大震災の際の福島第一原発事故に関するトピックに着目し，拡大したものが図 3 である．このトピックに含まれる画像の多くは，事故のあった建屋を撮影したもので，画像ヒストグラムの形状，及び写真の出現順序を見ることで，これらがニュース番組において繰り返し使用され，その後ブログ上でも広まったことが見て取れる．同種の写真は，放送のみならず新聞等のメディアにおいても使用されており，この場合においてはマスメディアが先行して写真を掲示し，それがブログに広まったと解釈することができる．

本システムのダイナミッククエリ機能を用いると，このトピックの中からブログが先行して用いた画像があるかどうかを調べることができる．図 4 は，交差相関係数を用いてブログが先行している画像クラスタを抽出したものである．この画像は，MIT の研究者によるブログにおいて福島第一原発事故解説に用いられた原子炉の透視図である．本記事が日本のブログにおいて翻訳され話題となり，その後 2 日遅れてニュース番組で用いられるようになったことが分かる．番組作成者が実際にブログを見てこの画像を発見したかどうかまでは不明であるが，これは画像の照合を用いることで初めて発見できる映像素材の原典追跡の事例となっている．メディア社会学と呼ばれる学問分野においては，映像の原典を探ることが重要なテーマの一つとなっており，本手法はそうした調査に有用である．

テレビとブログとでは，同じ話題に対する反応の大きさや勢いが異なり，それは国際問題，政治問題において顕著であることが今回の実験結果から分かっている．図 5 はその一例を示したものである．2012 年のオリンピックにおいて，試合の終了時にある選

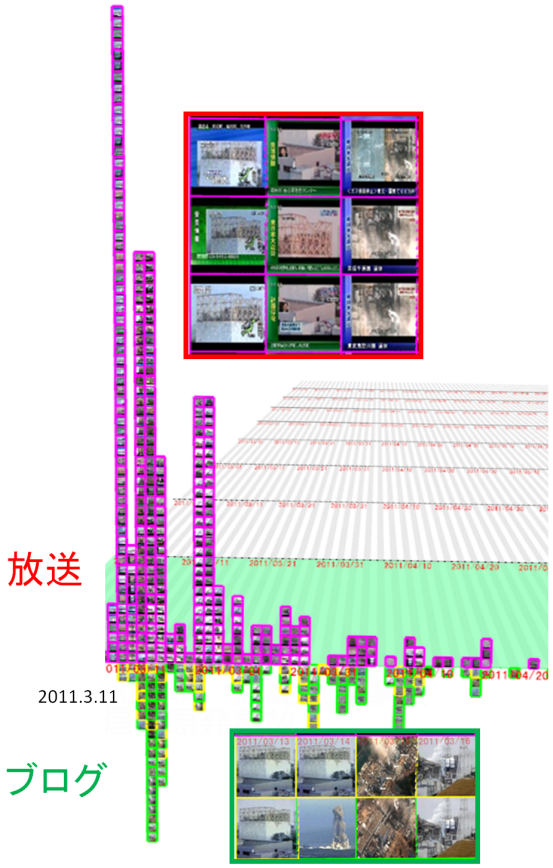


図 3. 福島第一原発事故に関するトピックのメディア比較

手が政治的なアピールをしたことが問題となった．そのニュースは最初テレビで放送されたが，その画像の露出は少なかった．しかし，このニュースは多くのブロガーを刺激しブログ上ではこの画像が大量に使われることとなった．この大きな反応を受けて，テレビのニュースでもこの問題が大きく扱われるようになった．こうした，テレビとブログが相互に反応して話題が大きくなるケースは，センシティブな国際問題や政治問題において良く見られるパターンである．

5 まとめ

本論文では，ブログ等の Web メディアに投稿された大量の画像から大きな話題になり多数出現した類似画像を抽出し，これが放送映像を起源とするものかどうかを放送映像アーカイブを照合して，話題の時差や露出の差を調査できるシステムを実現した．さらに大規模なブログアーカイブと，放送映像アーカイブを用いて実験を行い，様々な社会現象に関する多メディア間の情報追跡を行えることを示した．

本システムを用いることで，過去に起きた様々なイベントに対しての Web メディアと放送映像での反応の違いを学ぶことが可能であり，得られた知見



図 4. ブログがテレビ放送に先行した画像の例

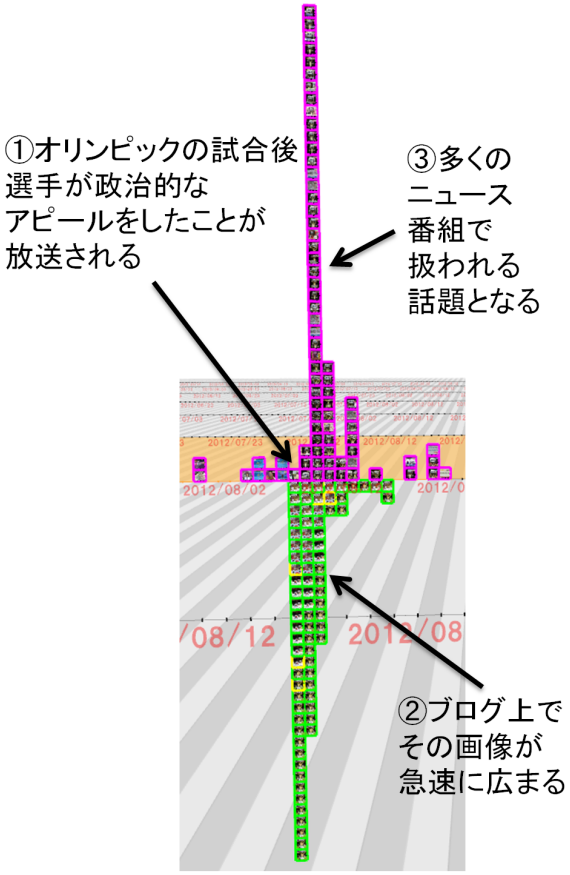


図 5. テレビとブログが相互に影響を与えた例

は今後同種のイベントが起きた際にどのような反応が各メディアで起こるかを予想する際の手がかりとなり得る．また，一方のメディアで話題になっているのに，もう一方のメディアでは大きな話題にならないケースも，幾つか本システムで観測されており，その原因を探ることで企業や組織の広報活動に活用することも可能である．ここで示した事例は，本実験で得られた知見のごく一部であるが，本システムを用いることで多メディアに関わる様々な興味深い社会事象を捉えられる可能性を示せていると信じる．

現状では，Web 画像集合と対応する放送画像集合が算出された後はある程度インタラクティブに閲覧・探索が行えるが，そのための事前計算には依然としてコストが高い．画像の特徴量抽出や放送映像データベースの索引を事前に作成しておくことはできるものの，画像類似度の算出や，対応画像の検索はインタラクティブに行えるほどではないため，一層の高速化が必要である．

謝辞

本研究は，文部科学省「多メディアＷｅｂ解析基盤の構築及び社会分析ソフトウェアの開発」，および JSPS 科研費 25280111 の助成を受けたものです．

参考文献

[1] E. Adar, D. S. Weld, B. N. Bershad, and S. D. Gribble. Why We Search: Visualizing and Predicting User Behavior. In *Proceedings of WWW'07*, pp. 161–170, 2007.

[2] J. A. Aslam, E. Pelekhov, and D. Rus. The Star Clustering Algorithm for Static and Dynamic Information Organization. *Journal of Graph Algorithm and Applications*, 8(1):95–129, 2004.

[3] K. Lebeda, J. Matas, and O. Chum. Fixing the Locally Optimized RANSAC. In *Proceedings of BMVC*, 2012.

[4] D. G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[5] H. Luo, J. Fan, J. Yang, W. Ribarsky, and S. Satoh. Analyzing Large-Scale News Video Databases to Support Knowledge Visualization and Intuitive Retrieval. In *Proceedings of VAST'07*, pp. 107–114, 2007.

[6] T. Uno. Multi-sorting algorithm for finding pairs of similar short substrings from large-scale string data. *Knowledge and Information Systems*, 25(2):229–251, 2010.

[7] J. Yang and J. Leskovec. Patterns of Temporal Variation in Online Media. In *Proceedings of WSDM'11*, 2011.

[8] C. Z. Zhu and S. Satoh. Large Vocabulary Quantization for Searching Instances from Videos. In *Proceedings of ICMR'12*, pp. 52:1–52:8, 2012.

静止画内物体への変形指示による動画検索

川手 裕太　　岡部 誠　　尾内 理紀夫　　平野 廣美 *

概要. 我々は動画から視聴したいシーンを素早く，容易に検索するためのユーザインタフェースを提案する．既存の動画共有サイトで使用されている検索エンジンはテキストを使用したものであるため，視聴したいシーンの様子までは検索することができないといった問題点が存在する．我々のシステムでは，ユーザが動画内の一時停止フレーム等の静止画を入力し，静止画内の物体の変形を矢印 2 本のスケッチによって指示をすることができる．その後，システムはユーザが矢印で指定した変形に類似したシーンを検索し，ユーザは目的のシーンを視聴することができる．我々の手法では，ユーザが視聴したい目的のシーンを検索するために 2 つのステップを採用している．第 1 ステップでは標準的な画像検索技術を使用し，入力された静止内の物体と同じ外観のフレームを検索する．第 2 ステップでは前ステップで選択したフレームから自動的に前後フレームへと早送りと巻戻しを行うことで，目的のシーンを検索する．早送りと巻戻しを行い，ユーザが 2 本の矢印で指定した変形が発見された場合，早送りと巻戻しを停止させ，目的のシーンを出力する．この 2 つのステップを行うことにより，静止画内物体が横向きにもかかわらず，外観の異なる正面向きの物体が映るシーンを検索するといったことを可能にした．我々はこの手法が F1 カーと馬，飛行機の 3 種類の物体において有効であることを確認した．

1　はじめに

動画共有サイトの普及により，インターネット上で視聴することができる動画の数は日に日に増加している．ユーザはそれらの動画から視聴したいシーンを迅速かつ容易に検索したいと考えている．ただし，YouTube[1] やニコニコ動画 [2] などの動画共有サイトの動画検索エンジンは，テキストでしか検索することができない．その検索エンジンでは，各動画のタイトルや割り当てられているテキストタグの情報に基づいて検索が行われている．従って，これらの検索エンジンで，検索項目として視聴したい物体の様子に関したテキストを入力しても検索できない．例えば,「左上に飛んで行く飛行機」というテキストに対して，YouTube やニコニコ動画ではいずれも，図 1 のような左上方向へ飛行する飛行機のシーンを取得することはできない．さらに，そのようなテキストを理解できる動画検索エンジンがあったとしても，ユーザが物体の様子を言葉で正確に記述することは困難であり，面倒である．

このような問題を解決するために，我々はユーザが迅速かつ容易に視聴したい物体の様子を指定することができる新しいユーザインターフェースを提案する．ユーザインターフェースでユーザが検索を行う際の流れは以下の通りである．

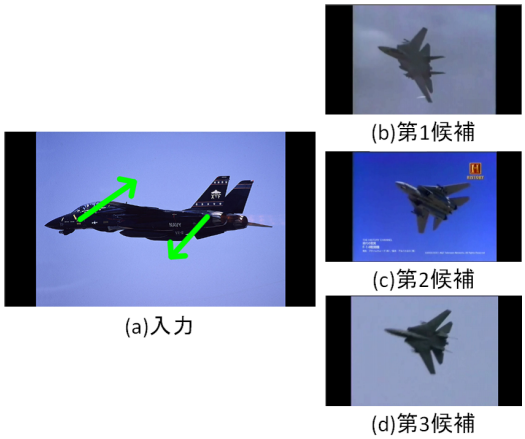


図 1. 提案システム

1. 視聴したい物体の映る動画の一時停止のフレームまたは静止画を入力静止画とする．
2. スケッチインターフェース (図 1(a)) にて，2 本の緑矢印の描画で入力静止画内の物体の変形を指定する．
3. システムが動画データベースから検索結果の候補のシーンを表示する (図 1(b,c,d))．

ユーザは 2 本の緑矢印で入力静止画内の物体の変形を指示することで視聴したいシーンをインタラクティブに検索することができる．また，この 2 本の緑矢印は 2 次元の入力であるが，横向きの入力静止画から正面向きのシーンを検索するといった，3 次元で変形させたようなシーンも検索することが可能

Copyright is held by the author(s).

* Yuta Kawate, 電気通信大学, Makoto Okabe, 電気通信大学/JST CREST, Rikio Onai, 電気通信大学, Hiromi Hirano, 楽天株式会社/電気通信大学

[1] http://www.youtube.com/

[2] http://www.nicovideo.jp/

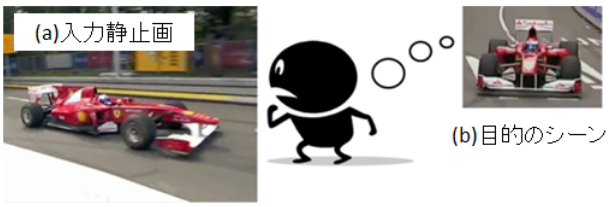


図 2. モチベーション

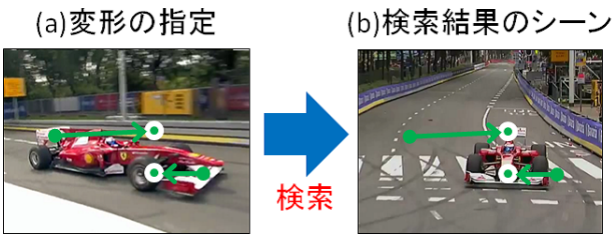


図 3. 提案するユーザインタフェース

である．実験において，この方法が F1 カー，馬，飛行機といった複数の種類の物体に有効であることを確認した．

2 関連研究

多くの動画からインタラクティブな検索を可能にする外観ベースのインタフェースの提案はなされている．bag-of-features を使用し，比較的低次元のベクトルとして各動画のフレームを表すことにより，入力した静止画から静止画や動画の検索が提案された [3, 15]．多くの動画や静止画から構成された 3 次元のシーンをユーザがインタラクティブに動きまわる事ができるシステムが提案された [1, 5, 16, 17]．これらは structure-from-motion を介してシーンやカメラの位置を 3 次元再構成を行なっている．しかし，これらの方法は，建物などの静止物体にしか適用されておらず，移動したり，変形したりする物体に適用することは困難であり，非常に計算時間がかかる．

動画内の物体を操作することで，連続した単一の動画内のシーンのインタラクティブなナビゲーションを可能にするシステムが提案された [4, 7, 8, 9, 10]．ユーザは，単一動画内の物体をドラッグによってインタラクティブな姿勢の操作ができ，ナビゲートできる．これらは，3 次元再構成ではなく，2 次元の画像処理技術が行われているので，比較的計算時間が少ない．しかし，我々のシステムは単一の動画だけではなく，複数の別々の動画から物体を検索することができる．

3 ユーザインタフェース

我々の提案するユーザインタフェースは，ある物体の動画の視聴時に別のシーンが見たいといった，物体の様子を検索するシステムである．例えば，図 2(a) のような F1 カーが左から右に走行するシーンの視聴している時，図 2(b) のように正面向きに走行する F1 カーのシーンを視聴したいという状況を考える．ユーザはそのようなシーンを視聴するために，早送りや巻き戻しを行うか，シークバーを操作する．もし，今視聴している動画にそのようなシーンが含まれない場合，ユーザが他に所持している F1 カーの動画を一つ一つ見る必要がある．それでも発見できなかった場合，YouTube 等の動画共有サイトで見つけ出さなければならない．その作業は非常に面倒である．

我々の提案するユーザインタフェースは目的のシーンを素早く，容易に検索が可能である．ユーザが入力するものは図 3(a) のように，視聴したい F1 カーが写っている静止画と 2 本の緑矢印だけである．このユーザインタフェースにおいてユーザは，「視聴したい F1 カーが現在見ている F1 カーの回転したバージョンである」という考えをもとに 2 本の緑矢印を入力する．この考えをもとにして，3 次元右向きの F1 カーを正面向きへ回転させることを考えた場合，正面向きの F1 カーを 2 次元の静止画上で見ると，図 2(b) のように F1 カーの前方は下部，後方は上部に配置される．このように F1 カーを回転させたと想定して 2 本の緑矢印で変形を指定する．ここでは，ユーザは図 3(a) のように F1 カーの前方を下方向へ，後方を上方向へと矢印を入力する．我々のシステムはその 2 本の緑矢印で指定した変形に類似した目的のシーン (図 3(b)) を検索する．

4 アルゴリズムの概要

動画や静止画検索の技術において，今視聴しているシーン (図 2(a)) を入力として，目的のシーン (図 2(b)) を直接探し出すことは困難である．例えば，SIFT 特徴量 [11] を動画から抽出し，今視聴しているシーンと目的のシーンのマッチングの計算を行う．マッチングに F1 カーを使用した場合の結果は図 4 で示す通りである．

右向きと左向きの F1 カーといった外観が類似している左のケースでは正確なマッチングが多く抽出されている．一方で，右向きと正面向きの F1 カーといった右のケースでは不正確なマッチングが抽出されているか，マッチングの数が少ない．ここで興味深いことに，図 4(a) と (b) のフレームは図 5 で示す通り同じシーンに存在している．

図 6(b) のような入力静止画 (図 6(a)) と外観が類似したフレームを探し出し，そのフレームから動画を巻き戻すことで図 6(c) のような目的のシーン (図 6(d)) を探し出すことができる．巻き戻しで図 5(d)

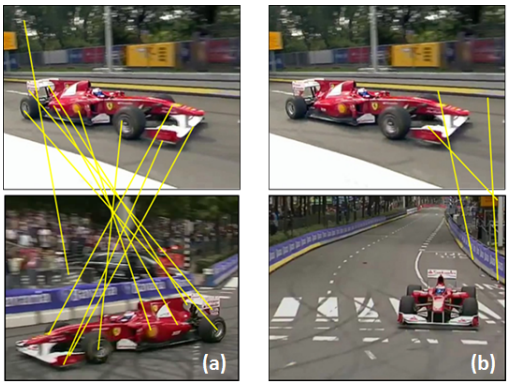


図 4. マッチングの例

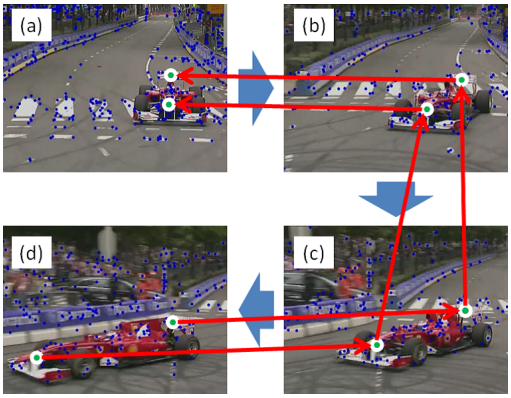


図 5. F1 カーが正面から左向きになる動画

と (a) を関連付ける方法として，図 5 の緑色の点のような物体上に配置された点を追跡する．

検索アルゴリズムは 2 ステップある．

1. フレーム検索ステップ

   SIFT 特徴量を使用したマッチングにより，入力静止画内の物体と類似したフレームを検索する．

2. フレーム追跡ステップ

   類似フレームから早送り・巻き戻しを行い，ユーザの入力矢印で指定された変形に類似したフレームを出力する．

2 ステップについて 4.1 と 4.2 で記述する．

4.1 フレーム検索

入力静止画と SIFT 特徴量を使用したマッチングを行った結果，マッチング数が最多のフレームを類似したフレームとして選択する．ただし，SIFT 特徴量をそのままマッチングに使用する方法では計算時間がかかる．計算時間を減らすために，Visual Words[15] を使用したマッチングを行う．この計算時間の比較は 5.1 で記述する．

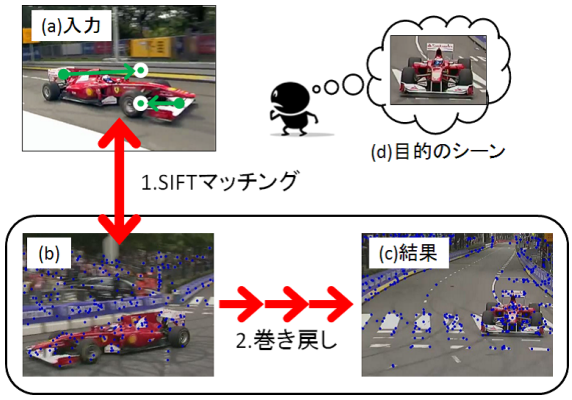


図 6. アルゴリズムの概要

Visual Words を使用したマッチングを行うために以下の前準備を行う．

1. 検索対象の動画の全フレームから SIFT 特徴量を抽出する．
2. 抽出した特徴量をランダムに選択し，クラスタリングを行う．クラスタリングには K-means 法を使用する．
3. それぞれのクラスタの中心のデータ (Visual Words) を辞書と定義する．
4. 辞書を使用し，残りの SIFT 特徴量についてどのクラスタに分類されるかを特定する．

この前準備で求めた動画の各フレームのクラスタの特定が行われた SIFT 特徴量は，ユーザインタフェースのシステム起動時に読み込まれる．ユーザインタフェースで静止画が入力された時，入力静止画と動画の各フレームとのマッチングは以下の流れで行う．

1. 入力静止画から SIFT 特徴量を抽出する．
2. 抽出した特徴量を辞書を使用し，クラスタの特定を行う．
3. 動画内フレームから抽出された各特徴量とクラスタの番号が一致する場合にマッチングとする．

マッチングは多く抽出されるが，その中には外観と比較すると不正確に見えるものが存在する．そのようなノイズを減らすために，RANSAC[6] を使用する．

Visual Words を使用し，入力静止画と外観が類似したフレームをマッチングを行った結果の例が図 7 である．図 7 下のフレームを含む動画において，マッチング数が最多であるので，このフレームをその動画での類似したフレームとして選択する．ただし，Visual Words を使用したことにより SIFT 特徴量をそのまま使用した場合よりも精度が低下している．よって，人間の見た目には類似していないにもかかわらず，マッチング数が動画内で最多のフレームが存在することがある．この問題に対して，マッ



図 7. Visual Words を使用したマッチングの例

チングの数が多い順から複数の候補のフレームを類似したフレームとして使用する．

4.2 フレーム追跡

4.1 で選択した類似フレームから前後フレームを探索し，ユーザが矢印で指定した変形にマッチングしたフレームを出力する．前後フレームの探索には図 5 のようにフレーム上に配置された点を追跡する．追跡点を抽出する処理は予め行い，システム実行時に追跡点データを読み込むようにする．物体の追跡には particle video[13] を使用したアルゴリズムと，SIFT 特徴量のマッチングを使用したアルゴリズムの 2 種類の方法を行った．前者は精度は高いが計算時間が非常にかかるため，実験では SIFT 特徴量のマッチングを使用する．SIFT 特徴量を使用した物体の追跡は多く行われている [2, 14] が，簡略化した以下のような流れで行う．

1. 動画内の各フレームで SIFT 特徴量を抽出する．
2. $n-1$ と n フレーム，n と $n+1$ フレームの SIFT 特徴量のマッチングをそれぞれ計算する．
3. $n-1$ と $n+1$ フレームとのマッチングが存在する n フレームの特徴点を追跡点とする．

そのアルゴリズムを使用し，追跡は以下のように行う．

1. 4.1 のマッチングの結果を使用し，入力矢印の始点と対応した類似フレーム上の点 (対応点) を求める．
2. 対応点と近傍の追跡点を求める．
3. フレームを移動させ，入力矢印の終点と追跡点が最近似のフレームを求める．

システムは，各動画の最近似フレーム中，上位 3 フレームを検索結果として出力する．

表 1. 各マッチングの計算時間 (秒)

	SIFT 特徴量 [12]	Visual Words
特徴量抽出	1.23	1.23
クラスタ特定	-	1.84
マッチング	91.11	1.78
合計	92.34	4.85

5 実験

すべての実験は CPU Intel i7-3930k 3.20GHz，メモリ 16GB のパソコンで行った．用意した全ての動画の大きさは 640×360 である．

5.1 計算時間比較

4.1 で記述したフレーム検索ステップにて，SIFT 特徴量をそのまま使用してマッチングを行った場合 [12] と，Visual Words を使用したマッチングを行った場合の計算時間は表 1 である．ただし，動画の SIFT 特徴量は予め抽出しており，それぞれのフレーム数は 761 である．また，ノイズを減らすために RANSAC[6] を使用している．

表 1 の結果から，Visual Words を使用したマッチングは，SIFT 特徴量をそのまま使用した場合より，マッチングを高速に行うことが分かった．

5.2 動画検索

我々は YouTube 上にある F1 カー，馬，飛行機の 3 種類の物体の動画でシステムの実験を行った．3 種類の物体の動画をダウンロードした後，1 シーンごとに動画を分割した．各物体ごとに検索対象は 20 本の動画である．それらの動画の全フレームから SIFT 特徴量を抽出し，Visual Words マッチングで使用する辞書とそれぞれの Visual Words データ，SIFT 追跡データを作成した．各物体の動画には 20 種類のシーンが存在する．

図 8 の 1 列目は静止画と矢印を入力したユーザインタフェースの画面，右の 3 列は各入力に対するシステムの検索結果である．実験に使用した動画のフレーム数と検索時間は表 2 に示す通りである．ここで，特徴量の抽出は 1 枚の入力静止画に対して 1 回だけ行えば良いので，2 回目以降の検索は抽出したデータを読み込むようにしているため，1 回目より高速になる．(クラスタ数　F1 カー 2500，馬・飛行機 500)

図 8 の 1 行目の (a) は F1 カーの動画を検索した結果である．この行では，ユーザが正面向きに移動する F1 カーのシーンを検索しようと，右向きの F1 カーと 2 本の緑矢印を入力した．それを想定して，1 本目の矢印は F1 カーの前方を指定して下に向けて入力する．続けて，2 本目の矢印は F1 カーの後方



図 8. 検索結果 (a) 正面向きに移動する F1 カーのシーン (b) 左向きに移動する F1 カーのシーン (c) 右向きに移動する馬のシーン (d) 正面向きに移動する馬のシーン (e) 右向きに飛行する飛行機のシーン

表 2. 検索時間

	F1 カー	馬	飛行機
フレーム数	3941	7018	3741
特徴量抽出 (秒)	7.08	7.04	1.41
特定 (秒)	3.29	2.86	1.64
追跡 (秒)	0.97	2.29	0.36
合計 (秒)	11.34	12.19	3.41

を指定し上に向けて入力する．検索結果は入力静止画とユーザが指定した矢印に近似した上位 3 フレームである．上位 3 フレームの中に，ユーザが検索しようとした正面向きに移動する F1 カーのシーンが含まれている．以降の行でもユーザが視聴したいと思ったシーンが上位 3 フレームの中に含まれている．

また，図 1 も飛行機について検索した結果である．これが我々の技術の限界を示している．ユーザは図 9 のような左上向きの飛行機のシーンを発見しようと 2 本の緑矢印を描画した．検索された結果は全てで左上を向いており，ユーザの指定が満たされている．



図 9. 左上へ飛行する飛行機

しかし，飛行機はその前方と後方を軸として回転しているシーンばかりである．ユーザが図 9 のように飛行機が回転していないシーンを望む場合であっても，我々のシステムを使用して，飛行機が回転しているシーンを排除することは困難である．

6 まとめ

本論文では，静止画内の物体に 2 本の矢印を描画することでシーンの検索ができる新しいユーザインタフェースを提案した．実験において，F1 カー，馬，飛行機といった 3 種類の物体において検索できることをを確認した．

今後として，解決すべき課題がある．まず，実験においてユーザが検索しようとしたシーンとは違う意図しないシーンが存在したという問題がある．SIFT 特徴量をそのまま使用し，Particle Video を使用して検索を行った場合 [12] では意図しないシーンはあまり出てこなかった．このような精度低下の原因は，使用するアルゴリズムを変更したからである．しかしながら，フレーム検索では 17 倍，フレーム追跡では 1.5 倍以上高速になったため，より多くの動画を以前よりも時間をかけること無く検索対象にすることができる．より多くの動画を使用することで，シーンの種類が増え，ユーザの意図に近い動画が増えることにより，意図しないシーンが上位に来なくなるのではないかと考える．

次に，色情報を使用していないという点がある．SIFT 特徴量のマッチングや追跡において，色情報を使用していない．そのため，例えば，F1 カーの動画を大量に用意した場合，形は同じだが色が異なる F1 カーが検索結果として表示される可能性がある．この場合のシステムの対応について，色の違う車を許容して検索結果を出すのか，色の違いを許さず色情報を使用する方法を検討するのか等を考える必要がある．

3 点目に，検索する対象が剛体で特徴的な部分が存在しないとうまく働かない点がある．F1 カーや飛行機といった乗り物では色や模様が異なることがあるが，形状についてあまり変わらない．また，タイヤやロゴ，翼等といったように特徴的な部分が存在する．馬は足を動かし形状が変化するが，馬衝や鞍といった馬具を付けられてこれが特徴的な部分となっている．一方で，小動物や魚を検索対象にして実験を行ったが，うまく機能しなかった．その原因は，動画内で形状が変化し，特徴的な部分が存在しないからである．形状が変化すると，SIFT 特徴量を使用したマッチングがうまく機能しない．また，特徴的な部分がないと正確なマッチングも判定するのが困難である．そのような物体に対しても検索可能にするためには，使用する特徴量について検討する必要がある．

参考文献

[1] S. Agarwal, N. Snavely, I. Simon, S. Seitz, R. Szeliski: Building rome in a day. In: ICCV 2009. (2009) 72-79

[2] S. Battiato, G. Gallo, G. Puglisi, and S. Scellato: Sift features tracking for video stabilization. In Proc. of International Conference on Image Analysis and Processing (2007) 825-830

[3] R. Datta, D. Joshi, J. Li, J.Z. Wang: Image retrieval: Ideas, influences, and trends of the new age. ACM Comput. Surv. 40(2) (2008) 5:1-5:60

[4] P. Dragicevic, G. Ramos, J. Bibliowitcz, D. Nowrouzezahrai, R. Balakrishnan, K. Singh: Video browsing by direct manipulation. In: Proc. of CHI ’08. (2008) 237-246

[5] J. M. Frahm, M. Pollefeys, S. Lazebnik, C. Zach, D. Gallup, B. Clipp, R. Raguram, , C. Wu, T. Johnson: Fast robust large-scale mapping from video and internet photo collections. ISPRS Journal of Photogrammetry and Remote Sensing 65(6) (2010) 538-549

[6] M. A. Fischler, R. C. Bolles: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. In: Comm. of the ACM 24(6) (1981) 381-395.

[7] A. Girgensohn, D. Kimber, J. Vaughan, T. Yang, F. Shipman, T. Turner, E. Rieffel, L. Wilcox, F. Chen, T. Dunnigan: Dots: support for effective video surveillance. In: Proc. of ACM Multimedia. (2007) 423-432

[8] D. B. Goldman, C. Gonterman, B. Curless, D. Salesin, S. M. Seitz: Video object annotation, navigation, and composition. In: Proc. UIST 2008. (2008) 3-12

[9] T. Karrer, M. Weiss, E. Lee, J. Borchers: Dragon: a direct manipulation interface for frame-accurate in-scene video navigation. In: Proc. of CHI ’08. (2008) 247-250

[10] D. Kimber, T. Dunnigan, A. Girgensohn, F. Shipman, T. Turner, T. Yang: Trailblazing: Video playback control by direct object manipulation. In: IEEE International Conference on Multimedia and Expo. (2007) 1015-1018

[11] D. G. Lowe: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision 60 (2004) 91-110

[12] M. Okabe, Y. Kawate, K. Anjyo, R. Onai: Video Retrieval based on User-Specified Appearance and Application to Animation Synthesis. In: Proc. of MMM (2013)

[13] P. Sand, S. Teller: Particle video: Long-range motion estimation using point trajectories. In: Proc. of CVPR ’06. (2006) 2195-2202

[14] S. N. Sinha, J. M. Frahm, M. Pollefeys, Y. Genc: Gpu-based video feature tracking and matching. In: Workshop on Edge Computing Using New Commodity Architectures (2006)

[15] J. Sivic, A. Zisserman: Video google: a text retrieval approach to object matching in videos. In: ICCV. (2003) 1470-1477

[16] N. Snavely, S. M. Seitz, R. Szeliski: Photo tourism: exploring photo collections in 3d. In: ACM SIGGRAPH 2006 Papers. (2006) 835-846

[17] J. Tompkin, K. Kim, J. Kautz, C. Theobalt: Videoscapes: Exploring sparse, unstructured video collections. In: ACM Transactions on Graphics (Proc. of SIGGRAPH). (2012)

料理画像をアニメーションすることによる魅力的な料理動画生成システム

崎山 翔平　　岡部 誠　　尾内 理紀夫　　平野 廣美*

概要. 本研究は，1 枚の料理画像をアニメーションすることで，より魅力的な料理動画を作成するためのシステムを提案する．泡や湯気，飛沫，具材の揺れなど，料理をより魅力的に見せる要素は多々ある．これらの要素は動きの中に多く現れるため静止画よりも動画の方が表現しやすいが，そうした要素をうまく捉えた動画を撮影することは専門家でも困難である．例えば湯気や泡を発生させるには，実際に火を点火するという方法が挙げられるが，こうした撮影を行うと料理の状態が撮影中に変化してしまうため，撮影のタイミングは極めてシビアであり，状態の変わってしまった料理は二度と撮影に用いることができない．そこで我々は，点火などといった料理の状態を変化させてしまう行為を行わず，料理を見栄えよくセッティングしただけの状態で撮影された静止画に，あとから湯気や泡などの動画を合成することでこれを解決する．こうした動画の合成を行うことのできる既存のソフトウェアは存在するが，非常にパラメータが多く，そのソフトウェアに対する専門的な知識が必要なうえに，数秒の動画の作成にも膨大な作業時間が必要となる．そこで本提案では泡や具材の揺れのパラメータ設定を行うためのアルゴリズムを独自に考案することで，ユーザのわずかな入力を基に，短時間で魅力的な料理動画の作成を可能とするシステムを実現した．

1　はじめに

現在，料理店の WEB メニューや電子メニュー，ショッピングサイトの商品イメージには静止画が多く使用されているが，これらの料理静止画を動画に置き換えると，美味しさの表現の幅がさらに広がり，コンテンツ全体をよりリッチにすることができる．料理は「活き」がうまく表現されているとより魅力的に見えるとされており，新鮮な野菜を滴る滴，グツグツと食材を揺らす泡，パチパチと鉄板から弾ける飛沫など，料理の活きを表す要素は数多く存在するが，これらの多くは動きの中に現れるため，静止画よりも動画の方がより豊かな表現が可能である．しかし，料理の活きを思い通りに表現して動画を撮影することは大変難しい．

ここで，料理撮影の専門家が，撮影が困難であるとする「鍋料理」を例に説明する．鍋料理をより魅力的に撮影しようとした場合，鍋に火をかけてグツグツと煮立っている状態を撮影することが最も一般的であるが，実際に鍋に点火してしまうと，肉には火が通りきってしまい黒く変色し，白菜などの菜っ葉は萎れ元の形状が維持できず，多くの具材はやがてスープに沈んでいってしまう (図 1-a)．状態が変化してしまった具材は二度と撮影には使用できないため，撮影タイミングは極めてシビアであり，専門家ですら魅力的な動画を撮影することは困難である．

一方で，図 1-b のような静止画の撮影は可能である．この静止画は鍋に冷たい水を張り，具材を綺麗に敷き詰めて，火をつけずに撮影したものである．火をつけていないため，肉には火が通りきることなく明るい色を保っており，春菊などの菜っ葉は萎れず新鮮さを保ち，全ての具材はスープに沈むことなく綺麗に並んでいる．撮影中に具材の状態が変化することがないため，ゆっくりと時間をかけて構図を決めて撮影できるうえ，何度も撮り直しが可能である．しかし，図 1-b には泡や湯気，具材の揺れなどといった要素が全く存在していないため，料理の活きが感じられない．



図 1. (a) 実際に鍋に火を点火して撮影された動画の一部. (b) 点火せず冷水に具材を並べただけの状態で撮影された静止画.

ここで，図 1-a のように具材の新鮮さを欠いてしまうことなく，図 1-b に，図 1-a にあったような泡や湯気などといった要素を加えることができれば，魅力的な料理動画の作成が可能である．

そこで我々は図 1-b のように見栄えよく撮影された 1 枚の静止画に，あとから泡や湯気，具材の揺れなどを合成することで，より魅力的な料理動画を作成する手法を提案する．こうした動画作成は Adobe After Effects[1] などの既存ソフトウェアでも可能であるが，ソフトウェアに対する専門的な知識を必

Copyright is held by the author(s).

* Shohei Sakiyama, 電気通信大学, Makoto Okabe, 電気通信大学/JST CREST, Rikio Onai, 電気通信大学, Hiromi Hirano, 楽天株式会社/電気通信大学

要とするうえに，わずか数秒間の動画を作成するために数時間の作業時間を要するという問題がある．そこで本提案では，魅力的な動画作成において特に表現が困難とされる「泡」や「具材の揺れ」を中心に，静止画 1 枚とユーザの少ない入力を基に，短時間で魅力的な料理動画を生成するシステムを開発した．また，実際にシステムを使用して料理動画を作成し，自然で魅力的な動画が作成できることを確認した．

2 関連研究

静止画をアニメーションする既存研究は数多く存在する [6, 5]．これらは画像の中に入っていくような 3 次元的な視点の変更やキャラクタアニメーションに便利であるが，料理動画の作成には対応していない．Chuang らは静止画に確率的な動きを与え，水面の揺らぎ等を扱った [4]．水面の揺らぎは泡の表現に似ているが，我々が行う泡の配置では，泡領域同士の関係やユーザが入力した情報によってそれぞれの領域にパラメータの異なった泡を配置するため，この手法を使用することはできない．岡部らは静止画に流体の流れる向きをスケッチすることで流体アニメーションの生成を行った [10]．この手法は，流体動画の配置がユーザのスケッチ技術に依存してしまうため，今回扱うような料理静止画を入力したところで質の高いアニメーションを得られるという保証がない．

また，ビデオテクスチャ合成の研究も多く存在する [13, 12] が，これらで合成した動画の見た目や動きを編集するのは難しい．Bhat らは手書きスケッチに基づいて流体ビデオを編集した [3]．見た目の編集も可能だが画像アニメーションは扱っていない．

さらに，本研究では具材画像をその種類に応じて自動で歪ませることで，具材の揺れを表現する．画像を変形させる研究は以前から多く行われており [2]，五十嵐らはこれをインタラクティブに操作することを実現した [7]．その後もこの手法の拡張などが多く行われている [11, 8]．しかし，本研究のように料理画像に特化し，具材の種類に応じて自動で画像を歪ませ，具材の揺れを表現するといった手法は存在しない．

そこで本研究ではユーザの少ない入力を基に，具材の揺らしアニメーションの生成，各泡領域に対する泡の配置及びパラメータ設定等を行うためのアルゴリズムを独自に考案することで，短時間で料理動画を作成することができるシステムを目指す．

3 システム概要

本手法における動画作成の典型的な流れについて述べる．ユーザはまずマウスを使用して入力静止画 (図 2-a) に簡単なスケッチを行い，泡を配置する領域 (泡領域) と食材領域の指定を行う (図 2-b)．同じスケッチの方法を用いて，図 2-b で指定した食材領域を，具材ごとに分割する (図 2-c). 次にユーザは図 2-c で分割した具材それぞれについて,「変形パターン」と「形状」というもの指定する (図 2-d)．変形パターンには「変形，非変形」の 2 つの選択肢が存在し，その具材が春菊のように柔らかく形が変わりやすいものであれば「変形」，大根のように固く形が変わりにくいものであれば「非変形」を選択する．形状には「平面，球形，棒状，葉類，麺類」という 5 つの選択肢が存在し，その具材に対してユーザが最も適切であると思うものを 1 つ選択する．あとは”合成”ボタンを押すだけで, ユーザが指定した領域や具材の情報を基に，システムが自動で泡の位置やサイズ，速さ，具材の揺らし具合などを決定し，動画を合成する (図 2-e)．このとき，動画に付与する湯気や音の指定も行うことができる．

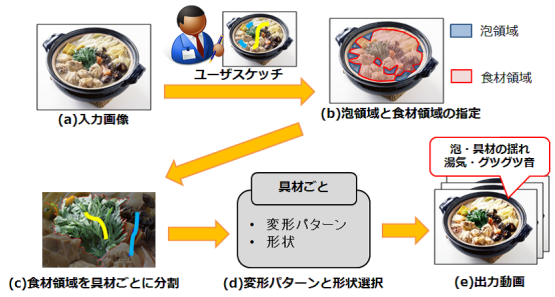


図 2. 動画作成手順.

4 ユーザインタフェース

ここでは作成したユーザインタフェースについて述べる．

4.1 泡領域と食材 (具材) 領域の指定

ユーザは簡単なスケッチで入力静止画に対し，泡領域と食材領域を指定を行う (図 2-a,b)．この指定には，スケッチベースの領域分割器を用いる [14]. 泡領域を指定する場合，ユーザは左クリックで泡領域を指定し，右クリックで鍋の縁や背景，食材部分などといった泡領域以外の領域を指定する (図 3-a). この指定は数少ないストロークで実現可能であるため，この操作は数秒で完了することができる．同様の指定を食材領域についても行う (図 3-b).

また，鍋料理のような煮込み料理の場合は，泡の発生と共に「具材が揺れ」が生じていた方がより自然である．そういった料理の場合，図 3-b で指定した食材領域を，さらに具材ごとに細かく分割する (図 3-c)．この操作も 1 つの具材に対しておおよそ 2 ストロークのみで指定を行うことができるため，全ての具材に対しこの指定を行っても，数秒から数分といった短時間で完了することができる．

具材の揺れが存在しない料理 (例えば，鉄板の上でジュージューと焼けるハンバーグには具材の揺れは存在しない) については図 3-c の操作は必要ない．

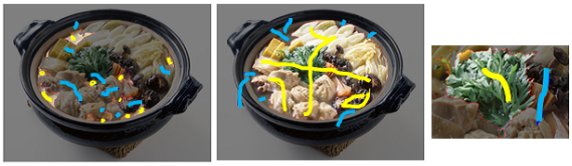


図 3. (a) 泡領域の指定．黄色のストロークが泡領域，青色のストロークが泡領域以外．(b) 食材領域の指定．(c) 具材の分割．

4.2 具材情報の入力

ユーザは分割した具材 (図 2-c) それぞれに対して，「変形パターン」と「形状」の指定を行う．変形パターンとは，具材の形が変わりやすいか否かを示すものであり，「変形，非変形」という 2 つの選択肢が存在する．春菊のようにひらひらと柔らかく形が変わりやすいものは「変形」，大根のように固く形が変わりにくいものは「非変形」をユーザが選択する．形状については，「平面，球形，棒状，葉類，麺類」という 5 つの選択肢が存在し (図 4)，この中から，その具材に対して最も適切であると思うものを 1 つ選択する．これらの情報は，具材の揺らしアニメーションを生成する際に使用する．



図 4. 具材に対する「形状」の分類例．

4.3 湯気や音の選択

作成する動画に湯気や音を加えることができる．この指定は”合成”ボタンを押す前に行う．湯気を加える場合は，いくつか用意された湯気の素材動画の中から好きなものを選択し，インタフェースに表示された料理画像をクリックすることで，湯気を貼りつける位置を決定する．音を追加する場合は，用意されたいくつかの音素材の中から好きなものを選択するだけでよい．

4.4 泡素材動画の選択

料理動画の合成に使用する泡素材動画を変更することができる．用意した泡素材動画には，水やソース，醤油など複数の液体の泡が存在する．例えば水とソースでは，粘性について大きな差があるため，泡の膨らみ方に違いがみられる．ユーザは自分の作成したい動画の感触により合った泡素材動画を選択することができる．

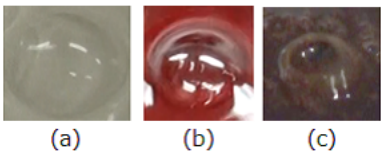


図 5. 泡素材動画．(a) 水 (b) 醤油 (c) ソース．

4.5 火元位置の指定

ユーザは火元の位置を指定することができる．火元とは画像中で最も強く加熱されている点であり，例えば鍋料理において，鍋の中心を火元として設定したとすると，鍋の中心部に最も熱が加わっているかのように鍋の中心部には激しい泡が発生し，鍋の縁の方では中心部と比較して穏やかな泡が発生する．

この火元の位置の指定は，料理動画視聴者の視線の誘導につながると考えている．一般に，人は動きの大きな部分に着目する．本システムでは火元がその部分に該当するため，動画視聴者に特に注目してほしい部分を火元に設定することで，視聴者の視線をそこへ誘導することができる．

4.6 動画の生成

ここまでの情報を入力すれば，あとは”合成”ボタンを押すだけで料理動画が生成される．ユーザは泡の位置や粘性，具材の揺らし方など，細かなパラメータ設定を一切行う必要がない．計算は数十秒で終了し，料理動画が出力される．ユーザは出来上がった料理動画を見て，泡の状態や具材の揺れ方などに関して気に入らない点があれば変形パターンや形状といった情報を変更して，料理動画を修正することができる．動画作成に必要な情報の入力のほとんど (領域分割以外) はワンクリックで行うことができるため，動画作成はすべて合わせてわずか数分で完了することができる．

5 アルゴリズム

ここでは，泡の配置や具材の揺らし方の決定，泡のレンダリング，音や湯気の追加で用いたアルゴリズムについて述べる．

5.1 泡の配置アルゴリズム

用意した泡データベースには水，醤油，ソースなど複数の液体それぞれについて，大きさや速度の異なるものが多数含まれている (図 5)．本システムではこのデータベースから適切な泡を選び出し，ユー

ザが図 2-b で指定した泡領域に対して配置する．この配置を行うために独自のアルゴリズムを考案した (表 1)．

表 1．泡の配置アルゴリズム．

	対象の泡領域		配置する泡	
	位置	広さ	速度	サイズ
手順 1 基本配置	火元付近	狭	極遅	小
		広	遅	小
	火元遠く	狭	遅	極小
		広	速	大
手順 2 重み付け	食材領域との境目及び 食材領域間の泡の速度をより遅く			

この泡配置アルゴリズムは我々が実際に鍋料理や鉄板料理などといった泡の発生しやすい料理の映像を複数観察し，手動で料理動画の試作を何度も試みた結果定義したものである．手順 1 では各泡領域に配置する泡の速度，サイズ，座標を決定し，手順 2 では手順 1 で配置した泡に対して重み付けを行う．

我々が調査したところ，発生する泡の特徴の違いには，その泡の領域の火元から距離と，その泡の領域の広さが大きく関係している．基本的に火元に近い部分では，発生から消失までの速度が遅く，縦横のサイズが小さい泡が発生しやすい．火元に近い部分の中でも特に，他の領域と比較して領域の広さが狭い場所では速度がより遅い泡が発生する傾向が見られる．一方，火元から離れた部分では領域の広さによって発生する泡の特徴が大きく異なる．狭い領域では速度が遅く，且つ火元付近で発生する泡などと比較してよりサイズの小さい泡が発生するのに対し，比較的広い領域では速度が速くサイズの大きな泡が発生する．手順 1 はこれらの特徴を考慮している．

手順 1 を本システムにおいて実現するために，各泡領域のピクセル数と，各泡領域の重心から火元までの距離を使用して，各泡領域に配置する泡のサイズ，速度のパラメータを決定する．泡領域 i のピクセル数を $a(i)$, 重心から火元までの距離を $d(i)$ としたとき，その領域に配置する泡のサイズ $S(i)$, 速度 $V(i)$ はそれぞれ次のように定義される．

$$S(i) \propto a(i) \cdot d(i) \tag{1}$$

$$V(i) \propto \frac{a(i)}{d(i)} \tag{2}$$

ここで，泡領域 i に対して，サイズ $S(i)$ の泡をグリッド状に並べる．これによりそれぞれの泡の座標が求まる．

さらに，食材との境目に存在する泡や複数の食材領域の間に挟まれて存在する泡は，その周囲の泡と比較して速度が遅くなる傾向が見られるため，手順 2 では，そのような泡に対して手順 1 で決定した速度を 0.7 倍し，これを重み付けとする．

5.2 具材の揺らしアニメーション生成

鍋のような煮込み料理では，グツグツと発生する泡と共に具材が揺れる．本システムではこの揺れのアニメーションを，ユーザが図 2-d で指定した「変形パターン」と「形状」に応じて自動生成する．

揺らし方については 2 パターン用意した．1. 具材画像を変形して歪ませることで揺れを表現するパターンと，2. 具材画像の座標を移動させることで揺れを表現するパターンである．図 2-d において，ユーザが「変形」を選択したい具材についてはパターン 1 の揺らし方,「非変形」を選択した具材についてはパターン 2 の揺らし方を適応する．

パターン 1: 画像変形による揺れの表現

春菊や青葉のような葉類が揺れるとき，具材全体が一様に揺れるのではなく，固い茎のある根本はあまり動かずに，柔らかい葉先だけが揺れていた方がより自然に見える．本提案では，リジット変形 [11] を用いて，春菊の根本のように形の変わりにくい部分を固定したうえで，葉先のように形の変わりやすい箇所を歪ませるように変形することで，具材の揺れを表現する．ここで，リジット変形を行う際，固定する点を固定制御点，歪ませる点を振動制御点という．

各具材に対する固定制御点と振動制御点の決定には，ユーザが図 2-d にて指定した「形状」の情報を使用する．我々は，用意した 5 種類の形状に対して，固定制御点と振動制御点をそれぞれどのように設定するかについて定めた (図 6)．



図 6．それぞれの形状に対する固定制御点 (中抜き円:赤) と振動制御点 (塗りつぶし円:黄色) の位置．

ユーザが入力した「形状」情報のみを使用して，システムが自動で図 6 のように制御点を設定するには，具材画像の傾きを考慮しなければならない．例えば図 6 の「平面」の例にある牛肉は，横軸に対して水平ではなく，左下から右上に向かって斜めに横たわっているため，この傾きに応じて制御点の座標を決定しなければ，変形が不自然なものになってしまう．そこで，本手法では主成分分析 [9] によってこの傾きを推定したうえで，制御点の座標を決定する (図 7)．決定された振動制御点は上下左右ランダ

ムに動かすことで具材が揺れているようなアニメーションを生成する．

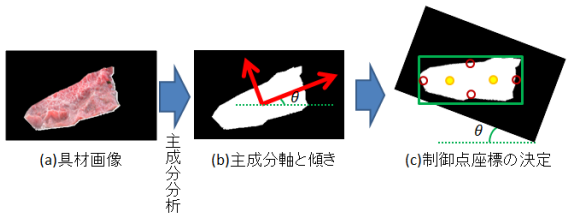


図 7. 主成分分析を利用した具材の傾き推定による制御点座標の決定．

パターン 2: 座標移動による揺れの表現

鶏肉やカットされた大根のような固めの食材が揺れるときは，パターン 1 のように具材の一部が揺れるよりも，具材全体が一様に揺れた方がより自然に見える．そこで，そのような具材に関しては具材画像全体を上下ランダムに座標移動し揺れを表現する．

5.3 レンダリング

泡素材動画と入力静止画の泡領域の色が異なっていても泡が自然に見えるように，泡素材動画の泡の動きだけを取り出して泡領域に合成する．これは図 8 のように HSL 色空間の輝度 (Lightness) チャンネルの動きのみを用いることで実現可能である．ただしこのとき，入力静止画の泡領域に存在したエッジが目立ってしまわないよう，泡領域に対してブラーをかけたうえで泡を合成する．

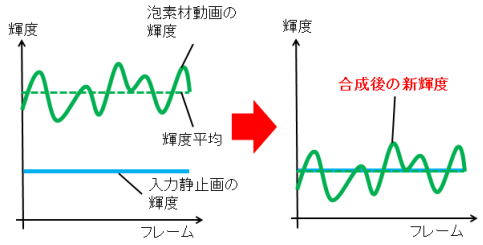


図 8. 輝度の揺らぎを付与するアルゴリズム．

5.4 湯気と音の合成

湯気は泡や具材の揺れと違い，細かいパラメータの設定をしなくても，湯気の素材動画を配置するだけで，動画に馴染んだ適切な湯気に見える場合がほとんどである．そのため，本システムではユーザが指定した位置に湯気の素材動画を貼りつけるだけで湯気の合成を完了する．音に関しても，料理の音はジュージューやパチパチ，グツグツといった白色性の高いものがほとんどであるため，音と映像の同期を考える必要はなく，ユーザが選択した音を動画に付与するだけで合成を完了する．

6 結果と考察

本システムを用いて 4 つの料理動画を作成した (図 9)．料理動画 a，b に使用した泡素材動画は図 5-a，料理動画 c，d に使用した泡素材動画は図 5-c である．



図 9. 実験に使用した静止画と動画作成に要した時間．

作成した料理動画 a-d について，動画全体や具材の揺れ方などの自然さを確認するためにユーザテストを行った．ユーザテストによって確認したのは Q1-Q4 の 4 項目である．

Q1 料理動画は静止画よりも美味しそうに見えたか．

Q2 料理動画は全体的に見て自然であったか．

Q3 泡は自然であったか．

Q4 具材の揺れは自然であったか．

4 つの料理動画 (図 9 a-d) に対し，Q1-Q4 それぞれについて，学生 7 人に 5 段階評価を行ってもらった (5:最高評価, 1:最低評価)．

ユーザテストの結果を図 10 に示した．なお，料理動画 d に関しては映像中に具材の揺れが含まれていないため，Q4 への回答は存在しない．

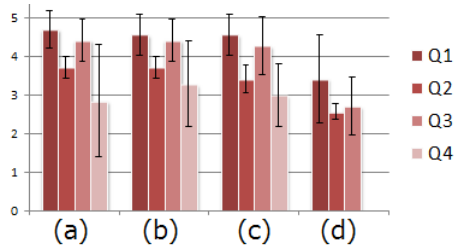


図 10. ユーザテストの結果．色棒は平均点，誤差棒は標準偏差．

ユーザテストの結果，システムで作成した料理動画 a-d すべてにおいて，作成に使用した静止画よりも美味しそうに見えるとの結果が得られた (図 10-Q1)．料理動画 d 以外は標準偏差も小さく，料理静止画をシステムによってより美味しそうな動画にすることができたことには，被験者のほぼ全員が同意したといえる．

動画全体の自然さ，泡の自然さ，具材の揺れの自然さなどを総合的に見て，最も評価が高かったのは料理動画 b であった．料理動画 b に使用した泡素材動画は図 5-a であるが，元画像の赤い色が合成した

泡にうまく反映されており，違和感をほとんど感じさせなかった．具材の揺れに関しても,「棒状」と指定したネギなどが非常に自然に揺れており，他の結果と比較して評価が高かった．一方，最も低い評価となったのは料理動画 d であった．料理動画 d は泡の配置も動画全体としても多少不自然であるという評価を受けた (図 10 d-Q2, Q3)．料理動画 d に対する被験者のコメントには「泡がぼやけて見える」という意見が多かった．これはレンダリング時，入力静止画の泡領域にブラーをかけてから泡の合成を行っていることが大きく影響したと考えられる．入力静止画の泡領域の色の差が激しいときや，指定した泡領域が大きい場合はこうした結果になりやすい．

料理動画 a, b, c のスコアを見ると，どれも同じような形をしている．泡の自然さ (Q3) のスコアの高さが，具材の揺れ (Q4) のスコアの低さに引っ張られた結果，動画全体としての自然さ (Q2) のスコアが下がっているという構図が見て取れるため，具材の揺れが今後改善されれば，動画全体としての自然さも向上することが期待される．

7 まとめと今後の課題

本研究では，ユーザの少ない入力を基に短時間でより魅力的な料理動画を作成するシステムを提案し，ユーザテストによって，その有用性を示した．ここでユーザテストに使用した静止画 (図 9 b, c) に注目したい．これらの静止画にはあらかじめ「料理画像に湯気を合成する」という加工が施されていた．料理画像に湯気を合成するのは，料理をより美味しく見せる方法としてメジャーなものであり，現在多くの料理店で使用されている静止画はこのように湯気が合成されているが，そういった加工を施すよりも (あるいは施した上で) 本システムによってその静止画を動画にした方が，料理をさらに美味しそうに見せることができる可能性が高いことをユーザテストの結果は示している (図 10 c, d-Q1)．つまり，現在の料理店のメニューなどが本システムにより動画に置き換われば，料理の美味しさの表現はその幅を広げることにつながると考えられる．

今後の課題としては，具材の揺らしアルゴリズムの改善などが挙げられる．現状として，具材の揺れにおける若干の不自然さが動画全体としての自然さを損なっているため，今後具材の揺れが改善されることで，いまと同じ労力で，より魅力的な動画の作成が可能になると考えている．さらに，今後はユーザインタフェースにも改良を加え，システムによって自動生成された動画をユーザがより自分の感覚にあった動画を直感的に微調整できるようにしたいと考えている．現在レンダリングに数十秒の時間を要しているが理論的には 2，3 秒程度で可能であるため，ユーザのスケッチを瞬時に結果に反映するなど，よりインタラクティブなシステムにしていきたい．

参考文献

[1] Adobe After Effects. http://www.adobe.com/jp/products/aftereffects.html.

[2] T. Beier and S. Neely. Feature-based image metamorphosis. In *ACM SIGGRAPH Computer Graphics*, Vol. 26, pp. 35–42. ACM, 1992.

[3] K. S. Bhat, S. M. Seitz, J. K. Hodgins, and P. K. Khosla. Flow-based video synthesis and editing. In *ACM Transactions on Graphics (TOG)*, Vol. 23, pp. 360–363. ACM, 2004.

[4] Y.-Y. Chuang, D. B. Goldman, K. C. Zheng, B. Curless, D. H. Salesin, and R. Szeliski. Animating pictures with stochastic motion textures. In *ACM Transactions on Graphics (TOG)*, Vol. 24, pp. 853–860. ACM, 2005.

[5] A. Hornung, E. Dekkers, and L. Kobbelt. Character animation from 2D pictures and 3D motion data. *ACM Transactions on Graphics (TOG)*, 26(1):1, 2007.

[6] Y. Horry, K.-I. Anjyo, and K. Arai. Tour into the picture: using a spidery mesh interface to make animation from a single image. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pp. 225–232. ACM Press/Addison-Wesley Publishing Co., 1997.

[7] T. Igarashi, T. Moscovich, and J. F. Hughes. As-rigid-as-possible shape manipulation. In *ACM Transactions on Graphics (TOG)*, Vol. 24, pp. 1134–1141. ACM, 2005.

[8] A. Jacobson, I. Baran, J. Popovic, and O. Sorkine. Bounded biharmonic weights for real-time deformation. *ACM Trans. Graph.*, 30(4):78, 2011.

[9] I. Jolliffe. *Principal component analysis*. Wiley Online Library, 2005.

[10] M. Okabe, K. Anjyo, T. Igarashi, and H.-P. Seidel. Animating pictures of fluid using video examples. In *Computer Graphics Forum*, Vol. 28, pp. 677–686. Wiley Online Library, 2009.

[11] S. Schaefer, T. McPhail, and J. Warren. Image deformation using moving least squares. In *ACM Transactions on Graphics (TOG)*, Vol. 25, pp. 533–540. ACM, 2006.

[12] A. Schödl, R. Szeliski, D. H. Salesin, and I. Essa. Video textures. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pp. 489–498. ACM Press/Addison-Wesley Publishing Co., 2000.

[13] L.-Y. Wei and M. Levoy. Fast texture synthesis using tree-structured vector quantization. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pp. 479–488. ACM Press/Addison-Wesley Publishing Co., 2000.

[14] 清野達也, 林貴宏, 尾内理紀夫, 三條正裕, 森正弥. 改良領域拡張法による高速画像切抜き手法の提案と評価. 情報処理学会論文誌, 12(50):3233–3249, 2009.

SyncBlink: 同期刺激によるナビゲーション方式の提案と評価

志摩 駿介　綾塚 祐二　寺田 努　塚本 昌彦*

概要. 近年，スマートフォンのアプリケーションや HMD (Head Mounted Display) などを用いた様々なナビゲーションシステムが開発，使用されている．ユーザを誘導する情報を提示するシステムにおいて課題となるのは，提示情報と現実世界とのマッチングである．例えば，画面上の地図に進むべきルートや曲がるべき交差点が明示されていても，それらが目の前のどの道，どの交差点に対応するかわからなければユーザは迷うことになる．そこで本研究では，パーソナルデバイスと環境側が連携し，同期した刺激をユーザに提示することでナビゲーションを行う方式 SyncBlink を提案する．環境側の各案内表示とユーザ側のデバイスが同期して刺激を発することにより，どの案内表示が目的のものであるかを容易に識別できる．また，複数のユーザに同じ，または異なった案内を同時に提示することも容易であり，人の多い環境でも利用できる．本論文では，ユーザへの刺激として視覚，聴覚，触覚刺激を提示するプロトタイプデバイスをそれぞれ実装した．評価結果から，提案方式がナビゲーション手法として有効であることを確認した．

1 はじめに

近年，カーナビゲーションやスマートフォン上の歩行者向けナビゲーションなど，様々な場面でナビゲーションシステムが使用されている．このような道案内を提示するシステムにおいて大きな問題となるのは，案内情報とユーザの目の前の現実世界とのマッチングである．例えば，画面上の地図に進むべきルートや曲がるべき交差点が明示されていても，それらが目の前のどの道，どの交差点に対応するかに自信がもてなければユーザは迷うことになる．また，GPS を用いたユーザの現在位置の確認は，実際の位置とずれた位置を示すことがあったり，電波の届かない地下では利用できない．

一方，現実世界に設置されている案内板などの環境側のナビゲーション設備は，適切に設置されていれば対応関係のわかりやすい案内となりうる．しかし，この案内は一般的に様々な目的をもつ不特定多数の人々に対して複数の案内情報が近接して提示されるため，あるユーザのための案内がどれであるかを見つけ出すのが難しい．一種類の案内しかない場合でも，それに気が付くのが困難な場合もある．これらは，確実にユーザ自身のために提示されるパーソナルデバイス上での案内とは対照的である．

そこで本研究では，パーソナルデバイスと環境側が連携し，同期した刺激をユーザに提示するすることにより，これらの欠点を解消したナビゲーション方式 SyncBlink を提案する．この方式はユーザに最低限の刺激を提示できればよく，またパーソナルデバイス側は視覚的な刺激だけでなく聴覚や触覚の刺激も活用できるため，シンプルで応用範囲が広い．

以降，2 章で関連研究について述べ，3 章ではシステム設計について述べる．4 章で評価実験を行い，最後に 5 章で本論文をまとめる．

2 関連研究

環境側と同期したパーソナルデバイスを用いたナビゲーションシステムはこれまでにも研究されている．澤野らによる AR (Augmented Reality) を用いたカーナビゲーションシステム [1] では，車載カメラからの映像を用いて道路端を抽出し，3 次元 CG を生成して実写画像とマッチングさせている．高田らによるウェアラブル AR システム [2] は，GPS とジャイロセンサによりユーザの位置と向きを認識し，ユーザが向いている方向にあるレストランの名前やメニューなどの情報を提示する．また，AR を用いたカーナビゲーションは実用化もされている（Pioneer サイバーナビ [3]，MapFan eye[4] 等）が，奥行き方向も含めた実景と提示情報のリアルタイムの表示位置合わせを高精度に行うのは (特に実景をそのまま見せる光学シースルー型の場合) 難しい．仮に完璧な位置合わせと表示ができたとしても，人間の奥行き方向の知覚の精度が限られるので，必ずしもそれが見やすい，判りやすい表示になるとは限らない．また，視界上の広い範囲を占める情報を提示するのは安全性の面での懸念が生じる．

デバイスがユーザに刺激を与えることでユーザを案内する方式も研究されている．塚田らによる Active Belt[5] は，8 方向に振動モータをつけたベルトを用いて，振動させる振動モータの位置で目的地の方向を提示し，振動の強さで目的地との距離を提示する．この他，水野らによる触覚ナビゲーショ

Copyright is held by the author(s).

* Shunsuke Shima and Masahiko Tsukamoto, 神戸大学大学院工学研究科, Yuji Ayatsuka, 株式会社電通国際情報サービス, Tsutomu Terada, 神戸大学大学院工学研究科/科学技術振興機構さきがけ

ン装置 [6] では，両耳裏付近に装着した扁平型偏芯振動モータにより触感覚にファントムセンセーションを生起し，位置情報の伝達を試みている．

環境側とパーソナルデバイスを連携させた研究では，松原らによる視覚障害者向け対話型情報提供システム [7] がある．このシステムでは点字ブロックに埋め込まれた位置情報を，ユーザが先端にアンテナのついた杖で読み込み，携帯端末装置に伝送することにより現在地を知ることができる．

3 SyncBlink

本章では，提案方式 SyncBlink の基本的なアイディアを説明してシステムの要求仕様を明確にした後，デバイスのプロトタイプを紹介する．

3.1 基本アイディア

1 章で述べたように，環境側の案内情報は，ユーザが案内情報の存在を知覚している場合はどれが必要な情報であるかを判別できる必要がある．また，そもそも案内情報がどこにあるのかをユーザが知覚していない場合も案内情報の存在がユーザに自然に伝わればわかりやすい案内となる．したがって例えば案内板の必要な部分を光らせ，点滅させるなどの方法を用いて，ユーザが必要とする案内だけを目立たせればこれらの目的は達成される．一方，この方法では複数ユーザに対応できず，多くの人が様々なタイミングで利用する案内情報提示には向かない．そこで,「自分に向けての情報であること」を，確実にユーザに伝えるために，ユーザのパーソナルデバイスを活用する．

筆者らは以前「画面上に複数の写真がタイル上に表示されており，それらがランダムに更新される場合，偶然同じタイミングで更新された部分は関連があるように感じられる」という知見を得ている [8]．これを応用すると，パーソナルデバイス上で，そのユーザが必要とする案内板の点滅と同期した点滅を提示すれば，必要な情報をユーザに認識させられるといえる．また，点滅による視覚的な刺激だけでなく，同期して鳴る音や振動による触覚刺激も同じように利用できると考えられる．この,「環境側の案内の必要な部分の点滅と，パーソナルデバイス上での同期した刺激」によるナビゲーション方式を SyncBlink と呼ぶ．この方式は，図 1 のように環境側が 1 つで同時に複数のユーザに同じまたは異なる案内を行う (一対多) ことも，環境側が複数で複数のユーザに同じまたは異なる案内を同時に行う (多対多) ことも可能である．

3.2 プロトタイプの設計と実装

提案方式を用いて常時自然な案内を行うために，装着するデバイスは邪魔になりにくいものである必要がある．また，提案方式は環境側デバイスとユーザデバイスが正確に同期する必要がある．これらの要件を考慮したシステム構成を図 2 に示す．この図のように，環境側が直接ユーザ側に同期信号を送る場合と，管理サーバが環境側とユーザ側両方に信号を送る場合が考えられる．また，図 3 に作成した，イヤホン，LED ，振動モータをそれぞれ用いたデバイスのプロトタイプと，それらを装着した様子を示す．イヤホンを用いたデバイスでは環境側との同期のタイミングでビープ音が鳴る．LED を用いたデバイスは眼鏡をかけヘアバンドで頭に装着し，環境側との同期のタイミングで LED が点灯する．振動モータを用いたデバイスはリストバンドで手首に装着し，環境側との同期のタイミングで振動モータが振動する．イヤホンを用いたデバイスは arduino nano，LED と振動モータのデバイスは arduino fio で制御し，XBee で無線通信を行っている．環境側が点滅の開始時に，無線通信で点滅の間隔（数値情報）を送信すると，デバイス側は受信した間隔の刺激を提示する．複数のデバイスに異なる情報を提示する場合は，提示したい表示の点滅開始時にそれぞれ点滅間隔を送信する．

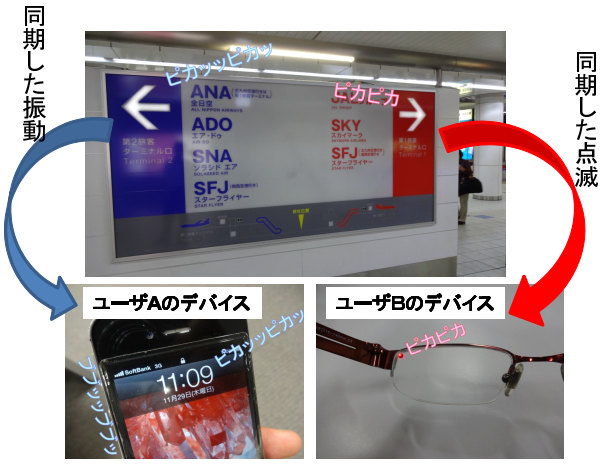


図 1. 案内板とデバイスの同期の例

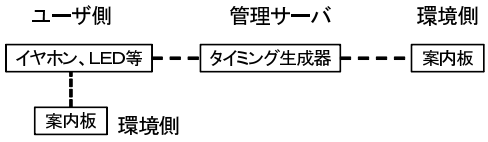


図 2. システム構成

4 評価実験

提案方式の有効性を評価した．まず，SyncBlink 方式を用いて提示情報の識別を行えるかを確認するため，PC 画面上での簡単な実験を行い，次に装着型デバイスを用いた場合の特性を評価する．最後に，実際に案内表示を環境内に設置し，ユーザを目的地



図 3. デバイスのプロトタイプ

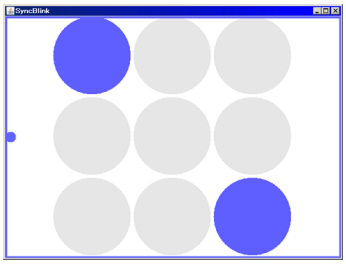


図 4. 実験用アプリケーション 1

に案内する実験を行う．

4.1 PC のみを用いた評価

まず，装着型デバイスは用いずに，視覚刺激と聴覚刺激を個別に，もしくは複合的に用いた場合に提示情報を識別できるかを確認するための実験を行った．実験では異なる周期で点滅する複数の円を PC 画面上に表示し (図 4)，ターゲットとなる円の点滅と同期して「ウインドウの枠を点滅させる (視覚刺激)」「ビープ音を鳴らす (聴覚刺激)」のいずれか，もしくは両方がある状態で，ユーザはターゲットを特定しクリックする．ウィンドウの枠を点滅させるのは，注視点以外への刺激を意図している．円はそれぞれが異なる 1〜4 秒程度の一定の周期で点滅し，2 つ以上の円が同時に点灯することもある．ターゲットをクリックすると円の周囲が赤く，ターゲットで

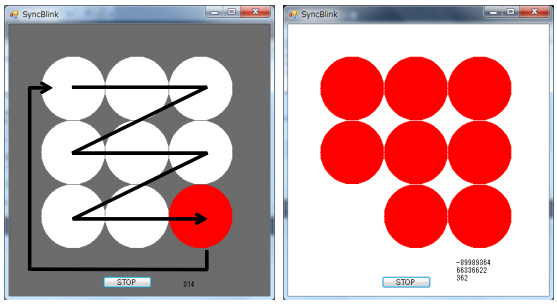


図 5. 実験用アプリケーション 2

ないものをクリックすると円の周囲がグレーになり，被験者は正否を認識できる．クリックすると，その正否に関わらず次のターゲットに進む．

このタスクを，a) 視覚的刺激のみ，b) 聴覚的刺激のみ，c) 視覚的刺激＋聴覚的刺激の三種類を各 10 回 (これを 1 試行とする) ずつ行わせた．開始前には，c の条件で説明及び 10 回の練習を行った．被験者は 9 人（男性 6 人，女性 3 人）で，a〜c の順序はランダム (ただし，c は常に最初か最後) に行った．被験者には，速さよりも正確性を優先するよう教示した．

表 1 が各被験者の各条件下での平均の回答時間，回答までの平均のターゲットの点滅回数，誤答の数である．被験者の番号は，誤答の少ない順にソートした．まず，全ての条件において誤答数が小さいことがわかる．被験者 9 は誤答が多いが，最大でも過半数には達していない．点滅数が 2 回以下になっている（すなわち，ターゲットを見定めたあとに再度点滅を待ち確認していない）試行では誤答が多い．9 名中 4 名は全ての試行を誤答なしで終了し，8 名は誤答なしで終えた試行がある．回答までの時間はターゲットの点滅周期の長さに依存するので，点滅回数で比較すると，全ての被験者において視覚的刺激のみの場合がもっとも点滅回数が多い．

これらの結果から，聴覚的刺激を用いた場合には劣るものの，視覚的刺激のみを用いても同期するターゲットの検出が可能であり，提案方式が有効に使えることが推察される．また，誤答の多かった被験者 8，9 は回答までの平均のターゲットの点滅回数が少ないことから，誤答は確認不足から起こることが示唆されるので，ユーザに対し確認を促すようなインタフェースの設計をした方がよいことがわかる．

4.2 ユーザデバイスを用いた評価

前節における PC 上での評価結果をふまえ，3 章で述べた装着型デバイスを用いて評価実験を行った．前節の実験と同様に，点滅する複数の円を PC 画面上に表示する．規則的，またはランダムな順序で円を点滅させ，図 3 に示したデバイスのいずれかを装着して，ターゲットとなる円の提示のタイミングとデバイスの刺激提示の同期の識別を確認した．規則

表 1. PC のみを用いた評価の結果

	視覚的刺激のみ			聴覚的刺激のみ			視覚的刺激 + 聴覚的刺激		
	時間 (s)	点滅数	誤答数	時間 (s)	点滅数	誤答数	時間 (s)	点滅数	誤答数
被験者 1	6.52	2.9	0	5.56	2.2	0	5.98	2.9	0
被験者 2	7.38	2.6	0	4.76	2.5	0	4.10	2.0	0
被験者 3	8.45	3.3	0	6.93	3.3	0	5.87	2.9	0
被験者 4	10.8	3.8	0	7.80	3.3	0	5.72	2.6	0
被験者 5	8.73	3.4	1	6.94	2.8	1	8.20	3.2	0
被験者 6	5.85	2.1	0	4.23	2.1	0	4.32	1.7	2
被験者 7	12.0	5.4	3	7.44	3.4	0	4.59	1.8	0
被験者 8	5.00	2.1	0	4.24	1.4	2	5.08	1.8	2
被験者 9	6.11	2.3	3	2.20	1.1	2	2.46	1.3	4

表 2. ユーザデバイスを用いた評価の結果

			イヤホン			LED			振動	
			0.3(秒)	0.2(秒)	0.1(秒)	0.3(秒)	0.2(秒)	0.1(秒)	0.3(秒)	0.2(秒)
間隔無	規則的		0.75	0.75	0.50	0.75	0.75	0.50	1.00	1.00
	ランダム	点灯	0.83	0.75	0.33	0.83	0.75	0.25	1.00	0.92
		消灯	0.92	0.75	0.58	0.83	0.75	0.50	1.00	0.58
間隔有	規則的		0.92	0.83	0.67	1.00	0.83	0.67	1.00	1.00
	ランダム	点灯	1.00	1.00	0.83	1.00	1.00	0.67	1.00	1.00
		消灯	1.00	1.00	1.00	1.00	1.00	0.92	1.00	1.00

的な順序で点滅する表示では，図 5 左の矢印の順序で点滅する．ランダムな順序で点滅する場合のみ，図 5 右のように提示方法を点灯ではなく消灯にした場合の実験も行った．これは，規則的な順序で点滅する場合は，次に点滅する円がわかっているため点灯と消灯で認識に差がないと判断したためである．規則的，ランダム共に，点滅と次の点滅の間に点滅時間と同じ長さの間隔がある場合とない場合（例えば点灯との同期なら，すべての円が消灯している時間がある場合とない場合）の 2 パターン行った．点灯時間は施行中一定で，0.3 秒，0.2 秒，0.1 秒とした．ただし，振動デバイスは 0.1 秒間隔では振動が起こらなかったため 0.1 秒は省略した．点滅順序が規則的な場合では，同期場所をランダムに決定した後 3 周期点滅させる動作 3 回を 1 試行とした．点滅の順序がランダムな場合では，同期場所の点滅周期が長いとその場所が予測できてしまい，同期回数が多いほど予測できる可能性が増すので，規則的な場合より同期回数を 1 回減らし，2 回同期するまで点滅を行う動作 3 回を 1 試行とした．このように，前節の実験と異なり正解が分かったらすぐに回答させるのではなく一定期間の点滅の後回答させるようにした理由は，前節の実験結果から確認不足による誤回答が多く，また実際の環境でも複数回の点灯期間が存在することが一般的であると考えたためである．被験者は 4 人（男性 3 人，女性 1 人）で，図 5 左の矢印の順に左上の円から 1〜9 の番号を割り当て，同期した円の番号を答えてもらった．また誤認識の傾向を調べるため，ランダムの場合は同期場所の点滅の前後に点滅した円の場所を記録した．

表 2 に実験の結果をまとめた．数値は被験者 4 人の平均正答率である．この結果から，0.3 秒と 0.2 秒の結果で比較すると，振動モータを用いたデバイスでの認識率が他の 2 つのデバイスでの認識率と比べやや高いが，各デバイス共にある程度高い正答率が出ているため，点滅間隔を長くすれば確実な認識が可能だと推察できる．また，点滅間隔が 0.1 秒の場合，振動以外のデバイスで点灯との同期よりも消灯との同期の方がわずかに認識率が高かった．このことから点灯との同期より消灯との同期の方が認識率が高いと推察できるが，点滅間隔を長くすれば点灯，消灯ともに確実な認識が可能である．また各デバイス共に点滅と点滅の間に間隔がある場合の方が間隔がない場合に比べ認識率が高かった．

この他，ランダムな順序の場合の平均認識率は，規則的な順序の場合に比べやや高かった．このとき，規則的に点滅させた場合の誤認識の 78.4% がターゲットの点滅の前に点滅した円との誤認識であった（ランダムの場合は 47.3% ）．このことより，規則的な順序で点滅させた場合，前後の点滅と誤解した場合のリカバリが難しくなることが分かる．ランダムな場合は，前後の点滅が毎回異なるため，そのミスが起こりにくかったと考えられる．よって，点滅の順序を規則的にする場合は，誤認識をなくすため点滅の間隔を長くする必要がある．

4.3 実際の案内による評価

環境側の表示と 3 つのプロトタイプを用いて，実際にユーザを案内できるかの実験を行った．前節の評価より，点灯との同期と消灯との同期では，消灯

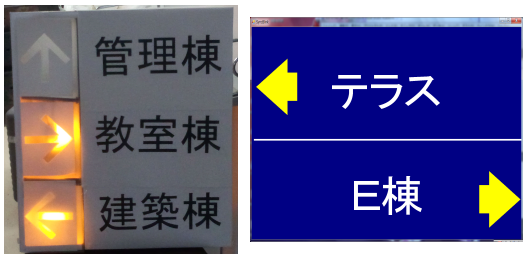


図 6. 案内表示



図 7. 案内実験の風景

との同期の方が認識率が高かったため，実験は消灯との同期で行った．また，ランダムな順序での点滅では，ユーザが表示を注視する必要があり，間隔の長い規則的な順序での点滅での認識は十分確実だと推察できるので，実験は規則的な順序での点滅で行った．点滅の間隔は，ユーザが容易に識別できかつ次の提示までの待ち時間が長いと感じないように，筆者の判断で 0.7 秒とした．実験は以下に示す 3 種類行ったが，環境側が 1 つの場合は図 6 左に示す案内版を用い，複数の場合は図 6 左の案内板と右の PC 画面上の表示とを併用した．また複数のユーザを同時に案内できるかを調べるため，各実験 2 人または 3 人 1 組で行い，同じまたは異なる案内を提示した．実験風景を図 7 に示す．この図では被験者は直進を提示されている．

円滑に案内できるかの評価

歩行中のユーザが立ち止ることなく円滑に案内された方向へ進めるかを調べるため，実験を行った．被験者に分岐を示す十字路から約 16 m 手前にスタートラインを設け，十字路の 1 m 手前の道端に直進または右左折を提示する案内板を設置し，歩行中の被験者が立ち止ることなく十字路を案内された方向に進めるかを調べた．被験者には始め進行方向とは反対を向いてもらい，システムを起動して最初の同期を感じてから，進行方向に進んでもらった．被験者は 12 人で各被験者それぞれ 3 つのデバイスで 1 回ずつ 3 回実験を行った．実験の結果，歩行速度には個人差があったものの，どのデバイスを用いても立ち止まった被験者はいなかった．

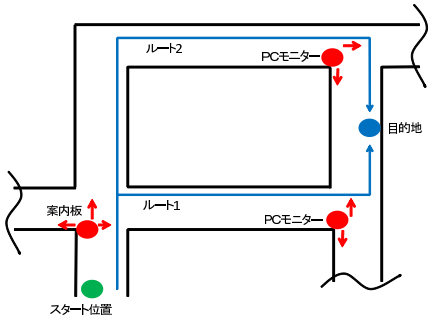


図 8. 目的地案内実験の簡易図

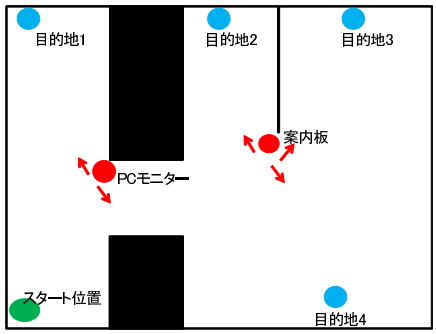


図 9. 複数表示実験の簡易図

目的地への案内の評価

実際にユーザを目的地へ案内できるかを調べるため，被験者に 2 つの分岐を進んでもらう実験を行った．図 8 のように環境側の表示は，案内板と PC モニターを用い，始めに案内板を用いて 3 方向の分岐を，次にモニターを用いて 2 方向の分岐を進んでもらった．スタート位置から目的地までには 2 通りのルートがあり，始めの分岐でどちらかのルートを提示し，2 つのルートそれぞれに次にある分岐にモニターを設置した．案内板と PC モニターの中間地点付近では，デバイスは両方からの信号を受け取り，意図した案内とは別の刺激を提示するが，案内板とモニターが同時に視界に入ることはなく，どちらかの表示がはっきり見える距離まで近づくともう一方からの提示は届かなかった．被験者は 10 人である．

実験の結果，すべての被験者を意図した場所へ案内できた．中間地点付近での両方からの刺激の提示について違和感を覚えた被験者もいたが，ほとんどの被験者は表示の見えていない地点での刺激なので気にならなかったと答えた．表示がはっきり見えるところでは片方からの刺激の提示しかないことから，間違った方向への案内は起こらなかった．また，この違和感を感じる問題は，次に示す実験の手法を用いることで解決できる．

複数の案内表示が同時に見える場合の評価

複数の案内表示が同時にユーザの視界に入る場合に，ユーザを正しく案内できるかを調べるため，短

い間隔で 2 つの分岐点がある環境に案内板と PC モニターを設置して被験者を意図した場所に案内できるかの実験を行った．始めに右または左方向への分岐があり，左方向には新たな分岐はなく，右方向へ進んだ先にはさらに 3 方向への分岐を設けた（図 9）．最初の 2 方向への分岐へ近づくと，次の 3 方向への分岐を示した案内版が見えるようにした．2 つの案内表示が同時に見えても正しい案内を行えるようにするため，それぞれの案内板の要素数 2 と 3 の公倍数 6 を 1 周期とした点滅を行わせることで，すべての表示の組み合わせに対応する．例えば案内する場所が図 9 の目的地 1 の場合，ユーザ側のデバイスは始めの案内の左の表示で常に刺激を提示するが，目的地 2，3，4 の場合，始めの案内が右を表示し，次の案内がユーザの目的地を指す組合せの場合にのみ刺激を提示する．このようにすることで，刺激提示間隔は延びるが，ユーザが間違えた情報を受け取ることがなくなる．

この実験の結果，案内場所が図 9 の目的地 2，3，4 の場合には最初の分岐での情報提示に多少時間はかかったが全ての被験者意を意図した目的地に案内できた．一方，視界内の案内版数および各案内板における選択肢数が増加すればその組合せも膨大になるため，刺激提示の周期が実用的なレベルに保てなくなる可能性がある．そのような場合，ユーザに最も近い案内表示からの信号とのみデバイスを同期させ，かつ後方の案内表示からの信号を受け取らない機構や，同時に見える複数の案内表示は点滅の周期を早くするといった工夫が必要となる．

4.4 考察

以上の実験より，提案方式は持ち運べるデバイスを用いて確実で円滑な案内ができることが確認できた．また，複数のユーザに同じまたは異なる案内が可能であることも確認できた．目的地への案内の評価において，被験者に同期のわかりやすかったデバイスについて聞いたところ，イヤホンを用いた音との同期が一番わかりやすいと答えた被験者が最も多く 6 人だった．またそのうち 5 人が，振動についても音と同じぐらいわかりやすかったと答えた．光が一番わかりやすかったと答えた被験者は 1 人だけで，それ以外の被験者は光が一番わかりにくかったと答えた．これは，案内板の表示が視覚刺激なのに対し，同じ視覚刺激を用いていることから，案内板の表示とデバイスの光提示両方を意識するのが難しいからだと推察でき，ユーザ側パーソナルデバイスに視覚刺激は不向きだと考えられる．一方，この結果から音声案内等に対して視覚刺激によりユーザをナビゲートするといったシステムが有効になる可能性がある．

このように，提案方式を用いて歩行中のユーザを円滑に目的地に案内できることを確認できたため，案内方式を用いると長時間装着型デバイスに注意を向けることなく提示を識別できる可能性が高い．したがって，パーソナルモビリティや自動車などの運転中の使用も可能だと考えられる．

5 まとめ

本研究では，パーソナルデバイスと環境側が連携し，同期した刺激をユーザに提示することでナビゲーションを行う SyncBlink を提案した．実験結果より，環境側の各案内表示を独立したタイミングで点滅させ，ユーザが必要とする案内に同期した刺激を端末から受け取ることにより，適切に案内が行えることを確認した．ユーザ端末側の刺激としては視覚や聴覚，触覚を活用でき，歩行中だけでなく，パーソナルモビリティの運転中等，さまざまな状況に対応することが可能だと考えられる．同じまたは別々の案内を複数のユーザに同時に行うことも容易であり，人の多い環境でも活用できる．

今後は，さまざまな形態での実装を進める．その上で実環境での実験を行い，有効性を検証する予定である．

参考文献

[1] 澤野弘明, 岡田 稔: 拡張現実感技術を用いた次世代カーナビのための道路端抽出法の検討, 情報処理学会研究報告 (グラフィクスと CAD 研究会), Vol. 2004, No. 121, pp. 103–108 (Nov. 2004).

[2] 高田大輔, 小川剛史, 清川 清, 竹村治雄: 身体動作に基づき提示情報を切り替えるコンテキストアウェアなウェアラブル AR システム, ヒューマンインタフェース学会論文誌, Vol. 12, No. 1, pp. 47–56 (Feb. 2010).

[3] CyberNavi: `http://pioneer.jp/carrozzeria/cybernavi/`.

[4] MapFan eye: `http://www.mapfan.com/iphone/ar`.

[5] 塚田浩二, 安村通晃: Active Belt: 触覚情報を用いたベルト型ナビゲーション機構, 情報処理学会論文誌, Vol. 44, No. 11, pp. 2649–2658 (Nov. 2003).

[6] 水野統太, 内田雅文, 井出英人: 振動刺激によるナビゲーション情報の認知評価, 電子情報通信学会技術研究報告 (ME とバイオサイバネティックス研究会), Vol. 106, No. 592, pp. 41–44 (Mar. 2007).

[7] 松原 広, 深澤紀子, 後藤浩一: 視覚障害者向け対話型情報提供システムの開発, 情報処理学会研究報告. ITS, [高度交通システム], Vol. 2001, No. 83, pp 75–82 (Sep. 2001).

[8] 綾塚祐二: 描き足していくことを強調した手書きコミュニケーションツール, 第 12 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS 2004), pp. 9–12 (Dec. 2004).

Traxion: 仮想力覚提示デバイス

暦本 純一*

概要. 仮想力覚を作り出す新しい触覚アクチュエータの構成について報告する．このデバイスは人間の振動刺激とその知覚の非線形な関係に基づき，振動波形を制御することで押したり引いたりといった「力」を感じさせる．既存研究と比較してアクチュエータのサイズや重量が非常にコンパクト (35.0 mm　5.0 mm　7.5 mm，5.2 g) で，力覚提示方向を迅速に変更することが可能である．振動アクチュエータへの入力信号と，実際にアクチュエーターに発生している振動や加速度をレーザ測距装置で計測し，仮想力覚の発生原因を解析した．評価者実験の結果，すべての実験参加者が仮想力覚を感じることを確認した．仮想力覚の強さを評価するために，現実の錘と仮想力覚を比較する評価実験を行い，このアクチュエータが約 30 g の仮想力覚を作り出していることを推定した．本研究は多様な場面でインタラクティブシステムに応用することが可能であると考えているが，本論文では歩行者ナビゲーションと，DLP(Digital Light Processing) プロジェクター光による可視光通信で空中に力覚場を作り出す応用例についても報告する．

1 はじめに

触覚はわれわれが生活する上でとても重要な感覚である．触覚として知覚することで，われわれは現実世界の物体の形状を把握し，物体の操作を精密に，正確に，そして素早く行うことができる．コンピュータなどの電子機器の操作においても，触覚フィードバックを利用している研究事例は多い．

ここで，触覚によるフィードバックを２つに分類する．ひとつめは力覚フィードバックと呼ぶもので，実際に「力」を発生させるものである．機構部品（ロッドなど）により操作部に力を与えているもの [7, 9]，ワイヤーにより力を与えるもの [5] などが提案されている．もうひとつは，触覚ディスプレイあるいは振動フィードバックとよばれるもので，力は発生させずに，フィードバックとして振動を利用するものである．たとえばタッチパネル上のボタンをタッチしたときに，クリック感を振動によって提示するものなどが提案されている [4, 11].

これらの二種類はそれぞれの目的・用途があり，また長所・短所がある．力覚フィードバックは機構部品やワイヤーなどで固定物と接続する必要があり，装置が大掛かりになってしまうので，モバイル用途などに適用することができない．一方，触覚ディスプレイでは本当の「力」を提示することができないので，現実での力覚を再現するような用途には向かない．また，フィードバックとして用いる場合も，利用者や操作部を特定の方向に導くために力の方向を利用することができない．

機構部品によらずに力覚を提示するいくつかの研究事例がある．2 ないし 3 個の偏心モーターによって慣性モーメントにより力覚を提示する装置を提案されている [14, 10, 3]．雨宮らは刺激の認知の非線形性にもとづき，回転体とクランク機構によって非対称的な加速度パターンを生成し，力覚が感じられることを示している [2, 16]．これらの研究により，実際に物理的な力を発生させなくても，知覚としての力覚の提示が可能なことがわかってきた．しかし，従来提案されている機構はいずれも回転体やクランク機構などの機械部品からなっており，全体のサイズや重量の点で改良の余地がある．たとえば雨宮らのデバイスは単方向の力覚提示に 70 mm　200 mm　48 mm の装置を要し，重量も 250 g であった [2]．これらを組み合わせて２次元ないし３次元の力覚提示装置を構成するためには，さらに重量が増すことが予測される．

回転体やクランク機構による力覚提示のもう一つの問題は反応性である．回転方向を急に変化させたり逆転することは難しいので，インタラクティブなシステムにおいて力覚の量を迅速に提示することができない．クランク機構は，構造上，回転方向を逆にしても非対称な加速度のパターンは逆転しないので，一次元かつ単方向への力覚提示しか行えない．回転体ではモーターの回転方向を変えることで力覚の提示方向の変化が可能であり、小型モーターによってアクチュエータの小型化も可能なことが報告されている [10]。しかしモーターの回転を急激に停止・開始することは依然困難である。

本論文では，これらの問題を解決し，従来よりも飛躍的に小型軽量化された仮想力覚提示デバイス，Traxion について報告する．Traxion の外形は 7.5 mm　35.0 mm　5.0 mm で重量は 5.2 g であり，双方向一次元の仮想力覚を創り出すことができる．また提示方向を変化させるのに要する時間も数ミリセカンド程度である．

2 Traxion の動作原理

Traxion はアルプス社のゲームパッド用振動アクチュエータ (Force Reactor [1]) の入力信号を制御す

Copyright is held by the author(s).

* 東京大学大学院情報学環・ソニーコンピュータサイエンス研究所

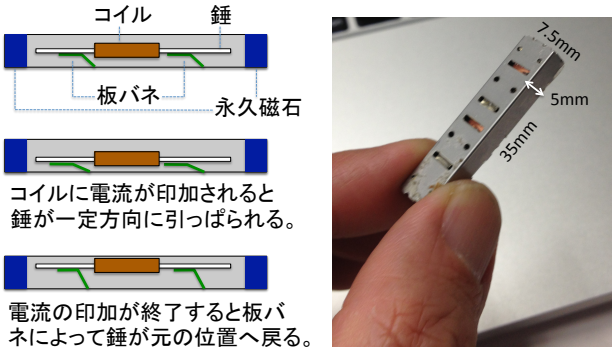


図 1. 振動子 (フォースリアクター) の構成

ることで仮想力覚を作り出している．Force Reactor は，コイルが板バネで支えられた振動体が格納されている．ケースの両端には永久磁石が設置されている (図 1)．この構造でコイルに電流を加えると磁界が発生し，コイルが一定方向に移動し，加速度を生み出す．印加を停止すると板バネによってコイルが元の位置に戻る．この構造をゲームパッドに組み込むと，様々な振動パターンを作り出すことができる．

実際にアクチュエーターに発生する加速度をレーザー測距装置 (キーエンス LK5000) によって測定した結果を図 2 に示す．指で保持した場合と，スポンジ，シリコン，ゴムの台上にアクチュエータを置き計測した場合を比較した．図では指で保持した場合と比較的加速度パターンが似ているスポンジ台での測定結果を示している．加速度計測周期は 50 μ s である．

図 2(1) は，コイルに直流電圧を印加したときのアクチュエーターに発生する加速度である，印加後コイルの移動にともない加速度が発生し，それが減衰していく様子がわかる．信号の印加／停止を周期的に繰り返すことで定常的な振動を作り出すことができる．図 2(2) は 125 Hz の正弦波を，(3) は同じく 125 Hz の矩形波（デューティー比 5:5) を加えた場合の加速度である．コイルに発生する磁界による運動と板バネによる減衰効果が組み合わさっており，周期運動ではあるものの，入力信号がそのまま単純には加速度変化に反映しているわけではないことが観測される．この (2) と (3) の場合，アクチュエーターを指で保持しても，振動は感じられるものの，力を感じることはなかった．

次に，アクチュエータにデューティー比の異なる矩型信号を印加してみると，仮想力覚が感じられることを発見した．図 2 (4) がそれで，周波数 125Hz でデューティー比 2:6(2 ms on,6 ms off) の信号を与えると，アクチュエータが一方向に引っぱられるような力を感じることがわかった．図からも明らかなように，このときアクチュエータが非対称な加速度運動をしていることがわかる．また，デューティー比を変化させることで，仮想力覚の方向も制御できることがわかった．図 2(4) では下向きの力を感じるが，(5) のデューティー比 6:2(6 ms on, 2 ms off) では逆に上向きの力を感じる．以上の場合でも，コイルに印加された信号が単純に振動に反映されているわけではなく，コイルと板バネの効果が複合していることを指摘しておく．

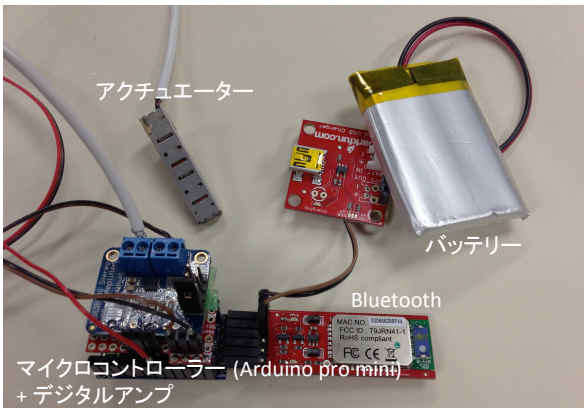


図 3. 試作システムの構成：マイクロコントローラー (Arduino) が Bluetooth 経由で命令を受信しアクチュエーター用の駆動波形を生成する．信号増幅にはオーディオ用デジタルアンプを利用している．

ここで，仮想力覚の発生原因について考察する．アクチュエーターは外部と機構部や糸などで支えられている訳ではないので，振動加速度を一定時間積分すればその値はゼロになり，現実には力は発生していない．したがって感じられる力は人間の知覚機構によるある種の錯覚であると推定される．仮想力覚の発生原理については Stevens の冪法則 (Stevens's power law) [13] と呼ばれる，刺激とそれに伴う感覚の強さの関係によるものが仮説の一つである．Stevens の冪法則は，刺激 I による感覚の強さ S が $S = kI^{\alpha}$ となるというものである（ここで k は比例定数，α は刺激の種類により決まる係数）．測定された加速度の積分値はゼロになるが，このようにして知覚された感覚 S の積分値はゼロにはならず，力として感じられるのではないかと推測している．仮想力覚の発生原因として考えられる他の要因は、刺激のマスキング効果である。強い刺激の直後の弱い刺激がマスクされるという現象である。図 2(4),(5) では大きな加速度の直後に（反対方向への）小さな加速度運動が連続しているので、この部分の知覚がマスキング効果によって抑制されれば、やはり力覚を感じることになる。

今回のデバイスでは，アクチュエーターをデューティー比の異なる矩型信号で駆動した場合に、コイルと板バネによる構成によりその駆動信号が変形された結果，力覚を感じる非対称な加速度を生み出していると推定される．

システム構成

本研究で利用しているシステム構成を図 3 に示す．駆動波形を発生させるマイクロコントローラー

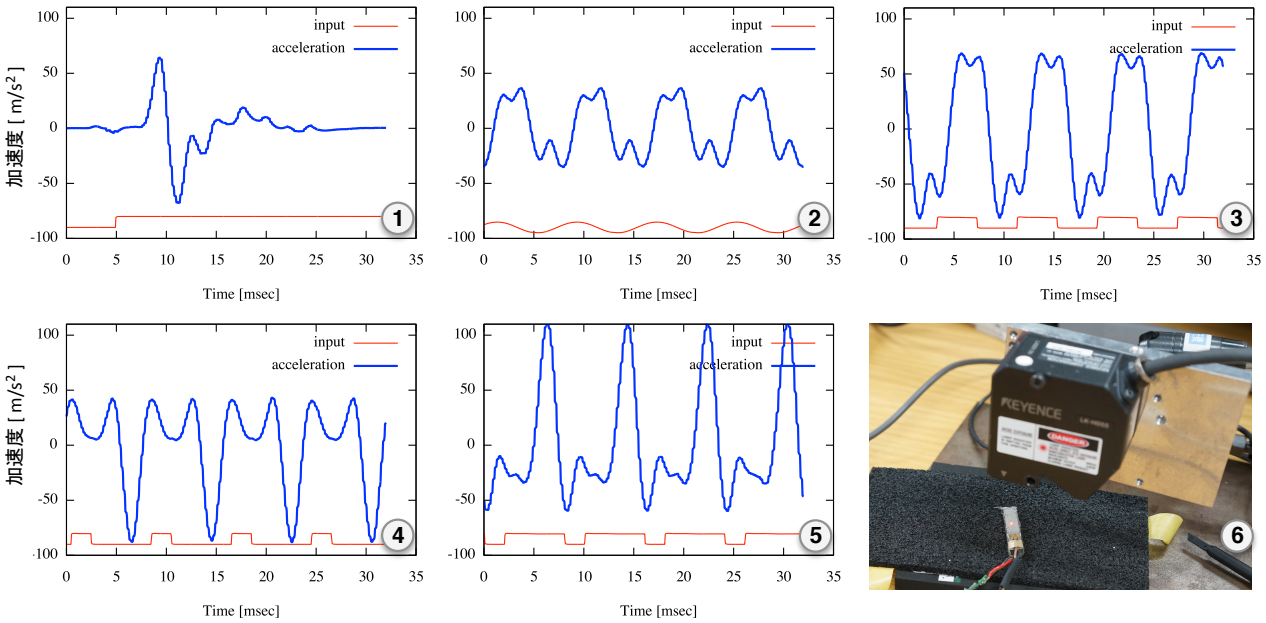


図 2. Traxion によって駆動される振動形状：デューティー比が異なる入力信号により異なる加速度が形成される．(1) 単一パルス印加と対応する加速度 (2) 正弦波 (3) 矩形波（デューティー比 1:1）(4) デューティー比 2:6 (5) デューティー比 6:2 (6) 加速度計測環境

(Arduino) が Bluetooth 経由で指令を受け取りアクチュエーターを駆動する．信号増幅のためにオーディオ用デジタルアンプを利用している．前章の実験により，仮想力覚を発生させるための波形は矩型で充分なことが分かったので，マイクロコントローラーの出力端子 (0V/5V を切り替え可能）を直接オーディオアンプの入力端子に接続し，DA 変化を省略することで構成の簡素化を実現している．

3 評価

3.1 実験 1: 一次元仮想力覚デバイス

Traxion の効果を確認するために利用者評価実験を行った．実験参加者は 10 名で，全員コンピュータサイエンス関係学科の大学院生である．参加者はアクチュエーターを指で保持し，仮想力覚が感じられたときにその方向を回答する．アクチュエーターを駆動する時間は 640，320, 160, 80, 40 ms で，正方向と逆方向の仮想力覚を作り出す 2 つの振動パターン (2 ms:6 ms と 6 ms:2 ms のデューティー比）との組み合わせを提示する．振動時間，振動パターンの組み合わせはランダムに選択されている．各参加者は，それぞれの振動時間に対して 5 回の刺激提示・回答のタスクを行う．回答はゲームコントローラーのスティックによって入力する．振動以外の情報を遮断するため，参加者はアイマスクとノイズキャンセルヘッドフォンを装着している（図 4:左).

評価結果を図 4(右) に示す．グラフに示されているように刺激時間が短縮されると若干正答率が落ちるものの，いずれも高い正答率を示しており，この

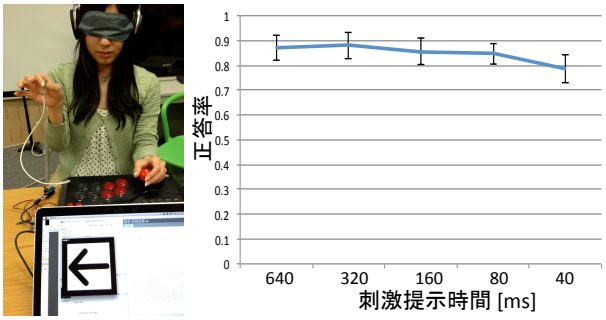


図 4. 実験 1: 一次元仮想力覚の評価実験環境および結果

仮想力覚が複数の実験参加者が同様に感ずる感覚であることが確認できた．

実験後のインタビューによると，多くの実験参加者が，指で保持しているアクチュエーターを動かしてみると仮想力覚をより明確に感じられるという意見が多く聴かれた．とくに力覚の方向にアクチュエーターを動かした時にこの効果は顕著だった．この効果に対する仮の説明は以下のようである．現実世界での牽引力，たとえばゴムやスプリングによるものは，牽引する方向に物体を動かすことで力が弱くなるものが多い．一方，仮想力覚の場合，移動させても感じられる力はとくに減らないので，それを「逆算して」元の力が強くなったと錯覚するのではないか，という可能性が考えられる．

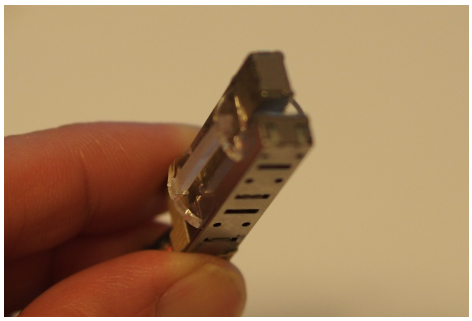


図 5. 2 次元仮想力覚デバイスへの拡張

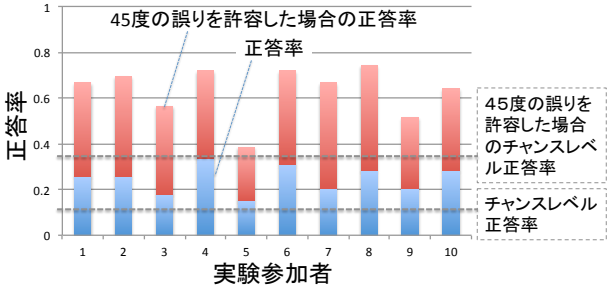


図 6. 実験 2: 二次元仮想力覚提示実験結果

3.2 実験 2: 二次元仮想力覚デバイス

次に，本デバイスを組み合わせることで多次元の仮想力覚が提示可能か検証した．図 5 のように，2 つのアクチュエーターを，それぞれの牽引方向が直角になるように接合した．このデバイスで，同時に加速力覚を提示した場合に，その合成方向に力覚が感じられるかを評価した．

実験 1 と同じ 10 名が評価実験に参加した．アクチュエータを指二本で保持し，力が感じられた方向を答える．この実験では 45 度ごとに 8 方向からいずれかを回答するようにした．刺激提示の時間は 640 ms で固定し，デューティー比は 2 ms:6 ms を用いた．縦横方向の仮想力覚はどちらか一つのアクチュエーターをその方向の力覚が提示されるように駆動し，斜め方向は両方のアクチュエーターを駆動した．参加者は実験 1 と同様，アイマスクとノイズキャンセルヘッドフォンを装着し，ゲームスティックによって回答した．

この実験では，図 6 で示すように正答率はチャンスレベルを上回るものの，正しく力覚が感じられているとは言いがたい状況だった．複数のアクチュエーターを同時に駆動しても，全体の振動が斜め方向を力覚を感じさせるような振動としては合成されないことが主な理由であると考えているが，複合アクチュエータでの実際の振動の解析を今後実施する予定である．

3.3 実験 3: 仮想力覚量の推定

では，発生している仮想力覚の強さはどのくらいだろうか．これを評価するために実験 3 を計画・実施

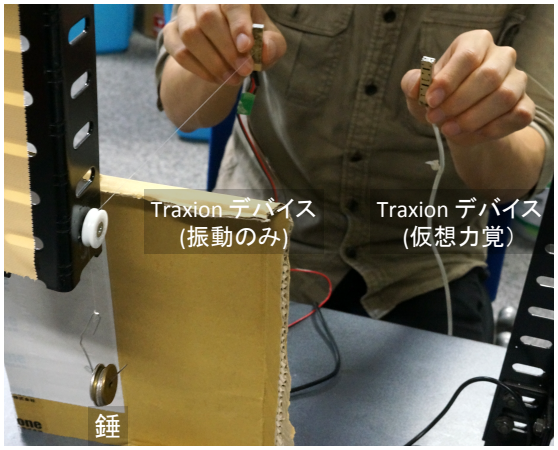


図 7. 実験 3:仮想力覚量の評価

した．実験参加者は，仮想力覚を発生するアクチュエータと，振動はしているが仮想力覚が発生しないように印加信号を調整したアクチュエーターを両手で保持する．力覚が発生していないアクチュエーターにテグスと滑車を介して錘を取り付け，牽引力が発生するようにしている．それぞれのアクチュエーターを交互に駆動し，錘の重さを変え，ちょうど同じ位の牽引力だと感じたときの重さを記録する．実験参加者には，錘の重さを感じるために，それぞれのアクチュエーターをゆっくりと前後に動かすことをお願いした．

それぞれの錘に対して，「錘の方が重い／軽い／同じ位」と回答してもらい，同じ位と答えた値を実験結果として記録した．実験 1,2 と同じ 10 名が実験に参加した結果，平均値は 29.8 g (0.292 N) 標準偏差は 8.5 g (0.083 N) だった．この実験により，発生している仮想力覚がおおむね 30 g 程度であると推定される．

ただし，この実験条件の場合，片方の手は錘の重量を常に支えているが，仮想力覚は振動しているときしか発生しない．したがってこの二つを単純に比較して正しいかどうかは疑問が残る．今後実験条件を変更するなどして仮想力覚の推定精度を上げていきたい．

4 応用

4.1 方向の提示

仮想力覚は，歩行者を誘導するために有効だと考える．たとえば指輪型のアクチュエータに仮想力覚を発生させて，誘導した方向へ牽引することができれば，視覚情報によらず空間ナビゲーションができるかもしれない．また，たとえば物の位置を何らかの手段で認識できるとすると，その物に向かって指先を牽引するデバイスなどが考えられる．また視覚障碍者のナビゲーション手段としても利用できるのではないかと考えている．

こういった位置誘導応用の可能性を検証するため

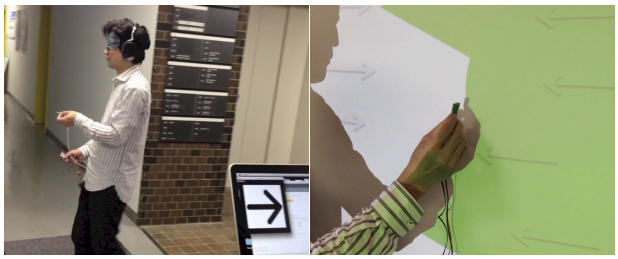


図 8. Traxion アプリケーション例: (左) 歩行者の誘導. (right) 仮想力覚つき入力装置. プロジェクターの照射によって力場を作り出す. DLP プロジェクターにより力の方向を光の点滅パターンとして照射する. 仮想力覚つき入力デバイスでその光をデコードし, 力覚に変換する.

に, 簡単な実験を行った. 利用者はアイマスクを装着し, 指で保持したアクチュエーターの牽引する方向に曲がるようにする. たとえば廊下の角に来た時に右に曲がるか左に曲がるかを仮想力覚によって支持する. 今回の実験では, 直進する場合には力覚は提示しなかった. また方向の支持はウィザードオブウォズ方式で, 実験参加者を目視して必要なナビゲーション支持をワイヤレスで与えるようにした (図 8(左)). 実験の結果, 視覚情報にたよらずとも歩行が可能で, 適切な場所で曲がれることが分かった.

4.2 仮想力覚つき入力デバイス

他の応用可能性として, 仮想力覚による入力デバイスが構想できる. もしスタイラスなどに仮想力覚を組み込むことができれば, アプリケーションによって牽引力や斥力を発生させることが出来る 8(右). 従来の力覚付き入力デバイスはスタイラス全体をロッドによって保持してたため作業範囲に制約があったが, 仮想力覚ではその制限がない.

図の例では, 作業空間に力の場を動的に作り出している. DLP プロジェクターの照射光で力場を定義する. DLP プロジェクターは, 色ごとに光の点滅パターンが異なるので, 簡易な可視光通信装置として使うことができる. 入力デバイスのフォトトランジスターでこの点滅を受信し, 色ごとにことなった力の方向を発生させれば, プロジェクターで照射する図形によって力場を定義することができる [12].

たとえば, 手術をする際のメスの動きを仮想力覚によってガイドするような応用が考えられる. 患者の体表に力場をプロジェクションし, メスに光センサーと仮想力覚アクチュエーターを装備すれば, 切開すべき場所を力覚として感じることができるようになる.

5 関連研究

他の仮想力覚でバイスとの比較

前述したように, 回転体やクランク機構によって仮想力覚を生み出すデバイスが提案されている [14, 2, 3] が, 本論文が提案するデバイスと比較すると重量やサイズが大きなものとなっている. また, 回転体は急激に回転数を変化させたり逆転させることができない. このことにより, これらの装置では単方向の一定量の仮想力覚を作り出すことができても, その方向を逆転させたり, 力覚量を制御することが難しい. これに対して, 本論文で提案しているデバイスはより単純な構造ながら仮想力覚を提示することが可能で, 既存事例から大幅に小型軽量化されている (重量比で約 40 分の 1) 点が利点と言える.

また, [2] では評価実験の結果, 周期運動が 40 Hz 以上になると仮想力覚が消失していくことを報告している. しかし, 本論文のデバイスでは振動数が 150 Hz 付近でも安定して仮想力覚が感じられた. したがって従来の回転体に基づく方式と本研究の構成では仮想力覚の発生原理が異なっている可能性もある. これについては今後さらに精密な検証をすすめていきたい.

力覚によるナビゲーション

ActiveBelt はベルト状の装着型デバイスで, 複数個のアクチュエーターの振動によって, 利用者を誘導することができる [15]. 上述のように, Traxion では方向を仮想力覚で提示することで, 利用者を特定の方向に誘導することを試みている.

力覚提示素子

触覚によるフィードバックは, タッチパネルに組み込まれたものが多く提案, 実用化されている [4, 11]. われわれが利用している素子は, ゲームコントローラー用に開発されたものである [1] が, 他の原理による力覚提示素子でも仮想力覚が発生するかどうかは興味深い課題である. ワイヤーは機構を利用せずに触感フィードバックを発生させるものとしてアレイ状の超音波トランスデューサーによるものがある [6] が, 本論文で示しているような力覚をつくりだすものではない.

6 結論・今後の課題

本論文では仮想力覚を生成する触覚アクチュエーター Traxion について報告した. 振動子の駆動パターンを工夫することにより, 仮想的な「力」を生成することが可能になった. 評価実験によりこの効果を複数の利用者によって確認した. また, 仮想力覚の大きさも評価し, 約 30 g 相当の力が生成されていると推定した. このデバイスは既存事例の力覚装置と比較して重量比で約 1/40 であり, インタラクションデバイスとしての利用可能性が高まった.

仮想力覚は人間の刺激と知覚の関係に基づいて力覚を「錯覚」として感じさせるものであり, 物理的な力は発生していない. また, 従来の機構部品による力覚フィードバックのように, 利用者の行動を束縛するほどの大きな力を作り出すことはできない. 一方, 振動フィードバックとは異なり, 明確に力の方向を感じることができる. したがって, 従来の力

覚フィードバックとも振動フィードバックとも異なるタイプの触覚フィードバックだということができる．例えば，歩行者ナビゲーションで移動方向に牽引するアプリケーションを振動フィードバックで実現しようとすると,「右」「左」などの方向を特定の振動パターンと対応づける必要があり，利用者がこの対応づけを記憶理解しなければならない．仮想力覚ではこのようなマッピングが不要で直接方向を感じさせることができる．

現状では既存の特定のタクタイルアクチュエーターの駆動パターンによって力覚が発生することを確認した段階だが，他の振動子でも同じ効果が発生するかどうかは興味ある課題である．実際，tactor [8] と呼ばれるコイル型の振動子でも力覚発生量は弱いながらも類似の仮想力覚が発生することを確認した．また，今回利用したアクチュエーターは，上下方向に非対称の構造を持っており，上方向の力覚と下方向の力覚の量が若干異なるように感じられる．今後は各種のアクチュエーターの駆動信号と加速度を計測し，仮想力覚の発生原理をより厳密に確認していきたい．

一方，評価実験の結果，一次元の仮想力覚デバイスを単純に組み合わせても，合成力覚を作り出すことができないことが分かった．これについても振動子の運動や加速度を解析し，可能性を探求していきたい．

アプリケーションとして，DLP プロジェクターによる可視光通信により，プロジェクションされた平面に容易に力場を形成できることを示した．これを仮想力場と呼ぶ．この発想を発展させると，各種のサーフェスコンピューティングにおいて，投影されている物体と現実のインタラクティブデバイスの間に引力や斥力を作り出すようなインタフェースが可能になる．また，現実の物体に投影した場合には，その物体の解析結果と仮想力場が対応するような応用が可能になる．たとえば，サーモカメラで撮影した温度映像から，温度勾配を力覚として感じるようなインタフェースや，手術の際のメスの力覚によるナビゲーションなどが考えられる．

参考文献

[1] ALPS. Alps Electric Press Release Force Reactor™. http://www.alps.com/e/news release/2005/0608 01.html, 2008.

[2] T. Amemiya, H. Ando, and T. Maeda. Phantom-DRAWN: direction guidance using rapid and asymmetric acceleration weighted by nonlinearity of perception. In *Proceedings of the 2005 international conference on Augmented tele-existence*, ICAT '05, pp. 201–208, 2005.

[3] A. Badshah, S. Gupta, D. Morris, S. Patel, and D. Tan. GyroTab: a handheld device that provides reactive torque feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pp. 3153–3156, 2012.

[4] M. Fukumoto and T. Sugimura. ActiveClick: Tactile Feedback for Touch Panels. In *CHI 2001 summary*, pp. 121–122, 2001.

[5] Y. Hirata and M. Sato. 3-dimensional Interface Device For Virtual Work Space. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 889–896, 1992.

[6] T. Iwamoto, M. Tatezono, and H. Shinoda. Non-Contact Method for Producing Tactile Sensation Using Airborne Ultrasound. In *Proc. EuroHaptics 2008*, Vol. 5024 of *LNCS*, pp. 504–513. Springer, 2008.

[7] H. Iwata. Artificial reality with force-feedback: development of desktop virtual space with compact master manipulator. *SIGGRAPH Comput. Graph.*, 24(4):165–170, 1990.

[8] T. Labs. Haptuator Mark II Product specification. http://www.tactilelabs.com.

[9] M. Minsky, O. young Ming, O. Steele, J. Frederick P. Brooks, and M. Behensky. Feeling and seeing: issues in force display. *SIGGRAPH Comput. Graph.*, 24(2):235–241, 1990.

[10] N. Nakamura and Y. Fukui. Development of Fingertip Type Non-grounding Force Feedback Display. In *EuroHaptics Conference, 2007 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2007. Second Joint*, pp. 582–583, 2007.

[11] I. Poupyrev, S. Maruyama, and J. Rekimoto. Ambient touch: designing tactile interfaces for handheld devices. In *ACM UIST 2002*, pp. 51–60, 2002.

[12] J. Rekimoto. SenseableRays: Opto-Haptic Substitution for Touch-Enhanced Interactive Space. In *CHI '09 extended abstracts*. ACM, 2009.

[13] S. Stevens. On the Psychophysical Law. *Psychological Review*, 64(3):153–181, 1957.

[14] Y. Tanaka, M. Sakai, Y.Khono, Y. Fukui, J. Yamashita, and N. Nakamura. Mobile Torque Display and Haptic Characteristics of Human Palm. In *ICAT 2001*, pp. 115–120, 2001.

[15] K. Tsukada and M. Yasumura. Activebelt: Belt-type wearable tactile display for directional navigation. In *in Lecture Notes in Computer Science. Springer-Verlag GmbH, Oct 2004*, pp. 384–399. Springer, 2004.

[16] 雨宮 智浩, 杉山 久. "牽引力錯覚を利用した牽引式羅針盤による視覚障がい者の歩行誘導の評価". 信学技報, 109(27, WIT2009-43):215–220, 2009.

HandyScope：引き出しジェスチャを用いたテーブルトップ用遠隔地操作手法

栗原 拓郎　三田 裕策　大西 主紗　志築 文太郎　田中 二郎*

概要. 大型のマルチタッチテーブルトップでは，ユーザはその大きさのため遠隔地に手が届かず，操作が困難である場合がある．そのような場合，ユーザはその領域に手が届く位置まで移動することを求められる．そこで我々は遠隔地操作手法である HandyScope を示す．HandyScope では，ユーザは手元から遠隔地を操作することができる．また，ユーザは手元と遠隔地の間にてオブジェクトを移動させることができる．加えて，HandyScope の起動，操作には引き出しジェスチャを用いることにより，ユーザは従来のマルチタッチ操作と競合することなく素速くポインティングを行うことができる．我々は HandyScope の性能を測るために従来のタッチ操作と比較実験を行い，遠隔地を選択する場合 HandyScope は有用であること及び，オブジェクトまでの距離が遠くなると HandyScope がより有用な手法であることを確認した．

1 はじめに

大型のマルチタッチテーブルトップを使用する際，ユーザはテーブルトップの周りに立ち，その位置からタッチ操作を行う．このタッチ操作を行う範囲について，ユーザは自身の位置から 34cm 以内の位置においてその 90%を行っていると Toney らは述べており [11]，それ以上遠くに位置する手の届かない領域（遠隔地）に対してタッチ操作を行うことは困難である．

そこで我々は遠隔地操作手法である HandyScope を示す．HandyScope では，ユーザは 2 個の円形のウィジェットを用いて遠隔地を操作することができる．2 個のウィジェットとは，遠隔地へ送り操作範囲を決定するサークル（scope）と手元において操作するためのサークル（handler）である．scope 内の領域は handler にも表示され，ユーザは handler に表示された領域を操作することにより遠隔地を操作することができる．また，ウィジェットを経由することにより，ユーザは手元と遠隔地の間にてオブジェクトを移動させることができる．加えて，本手法では Control-Display（CD）比を動的に変更できるため，ユーザは素速く目標領域まで scope を移動させること（ポインティング）ができる．

本手法では吉川らの提案する，両手を用いた引き出しジェスチャ[13] を本手法の起動及び scope の位置を決定するために使用する．引き出しジェスチャでは，ユーザは別のデバイスを使用せずに、タッチ操作のみにて使用することができる。また，引き出しジェスチャは従来のマルチタッチ操作と競合しないため，複数ユーザが同時に使用できる．テーブルトップでは複数ユーザが同時に作業することも多く，本手法はその際にも有用である．

2 関連研究

テーブルトップ環境における遠隔地ポインティング手法については多くの研究がなされている．Parker らは，スタイラスの先端の影を用いてポインティングを行う手法を示した [9]．また Banerjee らは片手の指を指す動作によってポインティングを行う手法を示した [2]．これらの研究は，ポインティング手法を実現するために追加のデバイスが必要となる．一方，我々の手法は追加のデバイスを用いる必要がない．また，マルチタッチテーブルトップ環境における仮想マウスを提案した研究もある [4, 8]．これらの研究では手の接触形状を認識するマルチタッチテーブルトップが必要である．一方，我々の手法では手の接触形状を認識する必要がないため，複数のタッチ点を検出可能なマルチタッチテーブルトップのみにより実現できる．

また，両手を用いたポインティング手法の研究も行われている．I-Grabber[1] はマルチタッチインタラクションにより操作するウィジェットである．両手を用いたポインティングを採用している点において我々の手法と同様であるが，我々の手法では CD 比を動的に変更でき，かつポインティング開始からポインティング先の決定までを一連のジェスチャにて行うことができる．そのため，ユーザは素速く正確にポインティングを行うことができる．所らは 2 つの加速度センサを用いた両手によるポインティング手法を示した [10]．また，Malik らは画像処理を用いた両手によるポインティング手法を示した [7]．こ

Copyright is held by the author(s).

* Takuro Kuribara and Yusaku Mita, 筑波大学大学院システム情報工学研究科 コンピュータサイエンス専攻, Kazusa Onishi, 筑波大学 情報学群 情報科学類, Buntarou Shizuki and Jiro Tanaka, 筑波大学 システム情報系.

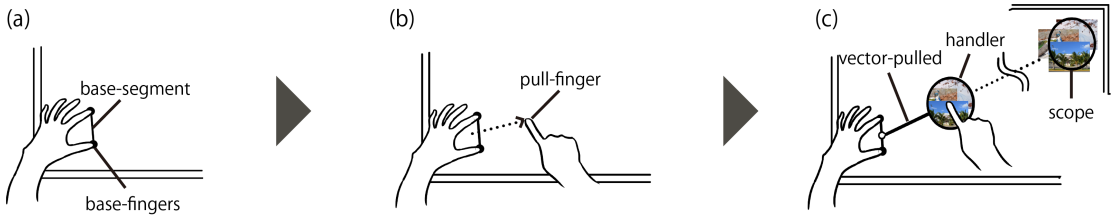


図 1. HandyScope の起動及び操作方法．a）2 本の指を置く，b）1 本の指にて 2 本の指の間を横切ることにより起動，c）HandyScope の操作

れらの研究と異なり，我々の手法ではテーブルトップを対象に，タッチを用いた両手によるポインティングを行う．そのため他のデバイスを必要としない．

3 HandyScope

本節では，HandyScope の操作手法及び，HandyScope を用いたオブジェクトの操作手法について述べる．

3.1 起動及び操作方法

HandyScope の起動及び操作方法を図 1 に示す．ユーザは最初に非利き手の 2 本の指（base-fingers）をテーブルトップ上に置く（図 1a）．次に，利き手の 1 本の指（pull-finger）により base-fingers の間（base-segment）を横切る操作を行う（図 1b）と，handler が pull-finger を中心とした位置に表示される．同時に，base-segment の中心点と pull-finger の間（vector-pulled）の半直線上に scope が表示される．ユーザが vector-pulled を変更すると，その変更に応じて scope の位置が変更される．図 1c に HandyScope を用いた操作の様子を示す．HandyScope は base-fingers をテーブルトップから離すことにより終了される．

3.2 scope の位置の決定

base-segment が生成されてから i フレーム後における scope の位置 P_i は以下の式により決定される．ここで S_0，S_i は図 2 に示すように base-segment の長さであり，S_0 は base-segment が生成された時の長さ，S_i は base-segment が生成されてから i フレーム後の長さである．また，G_i は base-segment の中心であり，V_i は G_i から pull-finger までの vector-pulled である．また，α は定数であり，k_i は CD 比を表す．base-segment を拡大すると CD 比を表す k_i が大きくなり，逆もまた同様である．すなわち，利き手もしくは非利き手を移動させることにより V を変化させると，その変位である ΔV_i に基づき，scope の位置が変化する．

$$
\begin{aligned}
P_i &= G_0 + \sum_{j}^{i} k_j \Delta V_j, \\
\Delta V_i &= V_i \quad V_{i-1}, \\
k_i &= \alpha \quad \frac{|S_i|}{|S_0|}.
\end{aligned} \tag{1}
$$

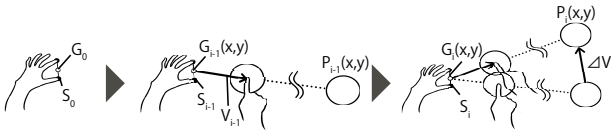


図 2. サークルの位置

3.3 CD 比の動的な変更

CD 比は base-segment の長さにより変更される．図 3 に CD 比と base-segment の長さの関係を示す．ユーザは CD 比の変更を行いながらポインティングを行うことができるため，大きな CD 比による大まかな操作と小さな CD 比による細かい操作を動的に切り替えて操作することができる．例えば，図 4 に示すように大きな CD 比により素早く scope を移動させ，小さな CD 比により微調整を行うことができる．

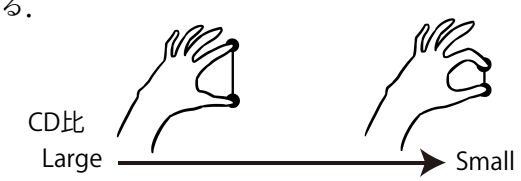


図 3. CD 比と base-segment の長さの関係

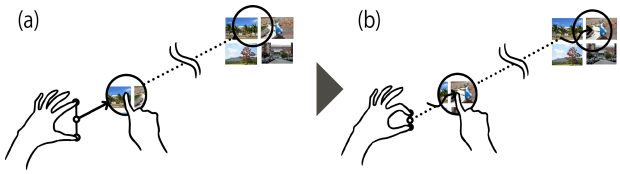


図 4. CD 比の動的な変更による scope の操作．a）大まかな移動によるポインティング先の決定．b）細かな移動による精密なポインティング先の決定

3.4 サークルを経由した遠隔地の操作

図 5 に示すように，handler 内には scope 内の領域が表示されている．また，handler 内での操作は全て scope 内に適用される．図 5a に遠隔地のオブジェクトを拡大縮小している様子を，図 5b にオブジェクトを回転している様子を示す．遠隔地に対する操作を全て handler 上にて行うことができるため，ユーザは自身が遠隔地に移動して操作することや，遠隔地のオブジェクトを手元に移動させることなく，遠隔地のオブジェクトを操作することができる．

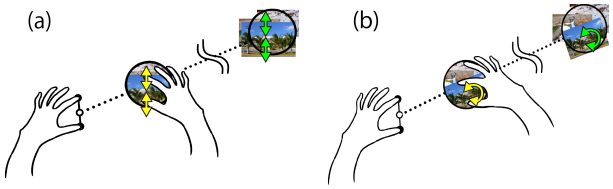


図 5. 遠隔地のオブジェクトの操作．a）オブジェクトの拡大縮小，b）オブジェクトの回転

3.5 サークル間のオブジェクトの移動

handler 内のオブジェクトをタッチしてサークル外までドラッグ操作を行うと，遠隔地のオブジェクトが手元に移動する．実際に移動している様子を図 6a に示す．同様に，図 6b に示すように手元にあるオブジェクトを handler 内にドラッグすることにより遠隔地にオブジェクトを移動させることができる．これにより，遠隔地と手元の間においてオブジェクトの移動を素早く行うことができる．

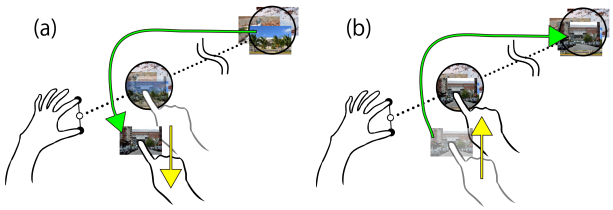


図 6. サークル間のオブジェクトの移動．a）遠隔地から手元にオブジェクトを移動，b）手元から遠隔地にオブジェクトを移動．

3.6 サークルの再移動

HandyScope の起動時は pull-finger を中心としてサークルが表示されるが，一度 pull-finger をテーブルトップから離すと，handler 内に対する操作は scope 内に対する操作として扱われる．そのため，図 7 に示すように再度サークルを移動させる場合には handler の中央ではなく，枠をドラッグする．これにより，scope の位置を移動させることができる．

3.7 サークルの大きさの変更

handler の枠をピンチ操作することにより．サークルの大きさを変更することができる．図 8 にサークルの大きさを変更している様子を示す．これにより，遠隔地の大きなオブジェクトを操作したい時や小さな範囲のみを操作したい時に対応することができる．

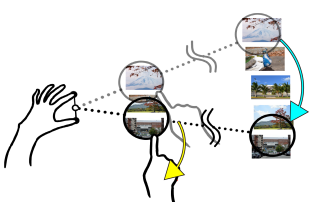


図 7. サークルの再移動

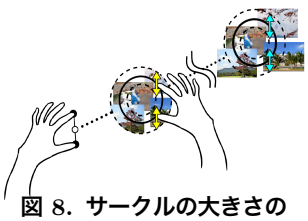


図 8. サークルの大きさの変更

3.8 HandyScope のメリット

HandyScope では，ユーザは非利き手の 2 本の指により CD 比を変更し，この 2 本の指の中心点と利き手の指の間の相対位置によりポインティング先を決定する．ポインティング先には scope を，手元には handler を表示し，ユーザはこのサークルを介して遠隔地に対するインタラクションを行う．これにより Frisbee[6] 及び Dynamic Portals[12] と同様に遠隔地を手元から操作することが可能になる．また，Frisbee では任意の位置から起動できず，Dynamic Portals では操作対象としたい遠隔地の指定に他者の協力を必要とするが，HandyScope では，ユーザは自身の操作のみにて任意の位置から起動し，動的な CD 比を用いて操作対象としたい遠隔地を素速く決定することができる．

4 実験

HandyScope の性能を調査するために，比較実験を行った．比較実験では HandyScope（HandyScope 条件）及び従来の直接タッチ（タッチ条件）の性能をテーブルトップ上の典型的な操作である選択，回転，拡大縮小の 3 種類のタスクにおいて比較した．

4.1 被験者及び実験環境

被験者は HandyScope を使用したことの無い 10 名（20 歳から 24 歳の大学生及び大学院生）であり，右利き 9 名，左利き 1 名であった．実験環境を図 9 に示す．実験に用いたテーブルトップは，画面サイズが 147cm　80cm の 60 インチディスプレイ（パイオニア社, PDP-607CMX[1]）に PQLab 社のマルチタッチフレーム（PQ Lab，Multi-Touch G^3 [2]）を装着することによりマルチタッチ機能を付したものである．なお，テーブルトップに関する幾つかの研究 [3, 5, 14] において，テーブルトップの高さを 91cm から 105cm としていたため，本実験ではテーブルトップの高さをその範囲内である 93cm とした．また，初期の CD 比にてテーブルトップの端までポインティング可能となるように式 1 における α を 12 とした．

4.2 実験手順

被験者には，選択タスク，回転タスク，拡大縮小タスクをこの順に課した．タスクは，テーブルトップ上でのポインティング手法を提案した研究 [2] に倣ったタスクである．各タスク前には，本番と同じタスクを本番の 1/4 の量だけ練習タスクとして課した．各タスクを行う際には，操作条件間の順序効果の打ち消しを狙い，被験者を HandyScope 条件を先

[1] http://pioneer.jp/biz/karte/PDP-607CMX.html
[2] http://multi-touch-screen.com/product_g3.html

に行う者とタッチ条件を先に行う者の 2 組に分けた．被験者には全てのタスクを終了した後にアンケートに回答して貰った．なお，被験者 1 人あたりの実験時間が約 1 時間半となったので，拘束時間に対する謝礼を渡した．

4.3 選択タスク

被験者には様々な位置に表示されるオブジェクトを選択して貰った．

まず，被験者は各試行の開始前にテーブルトップの短辺の中心（すなわち図 9 中の被験者の足元の床に，黒色のビニールテープにて示されている場所）に立つ．この状態から，テーブルトップ上のいずれかの提示位置に表示されるオブジェクトを選択する．なお，試行の開始前に開始点及びオブジェクトは表示されている．

テーブルトップ上に表示される開始点及びオブジェクトの提示位置を図 10 に示す．

HandyScope 条件では，被験者は開始点上において HandyScope を起動させることにより試行を開始する．次にオブジェクトに scope を合わせ，オブジェクトをタップする．オブジェクトがタップされると 1 回の試行が終了し，フィードバックとしてビープ音が発生する．タッチ条件では，被験者は開始点をタップすることにより試行を開始する．次に表示されたオブジェクトに手が届く位置まで移動して，オブジェクトをタップする．



図 9. 実験環境

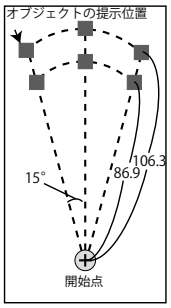


図 10. 開始点及び提示位置

本実験における独立変数は，開始点からオブジェクトまでの距離（86.9，106.3cm），開始点からオブジェクトまでの角度（ 15，0，15 度），オブジェクトの大きさ（3.9，5.8，7.7cm），操作条件（HandyScope，直接タッチ）である．被験者は各独立変数の組み合わせにおいて試行を 3 回ずつ，合計 108 試行（2 3 3 2 3）を行った．また，操作条件以外の試行の順序はランダムであった．

実験結果

両操作条件における 1 試行の所要時間を図 11 の左 2 つのグラフに示す．HandyScope 条件では所要時間は 1715ms，タッチ条件では 1943ms であった．被験者毎の所要時間において対応のある t 検定を行った結果，HandyScope 条件における操作が有意に速かった（$t(9) = 2.72, p = .011 < .05$）．

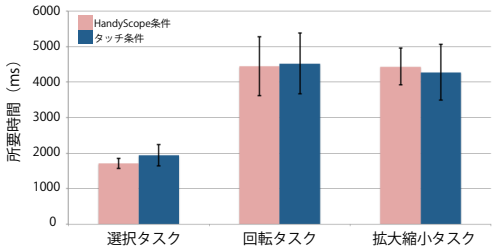


図 11. 各タスク毎の 1 試行の所要時間

4.4 回転タスク

被験者には様々な提示位置に様々な角度にて表示されるオブジェクトを回転させて，目標（ドック）に合わせて貰った．ドックの表示位置をオブジェクトと同位置，表示角度をオブジェクトと異なる角度とした．開始点及び提示位置，試行の開始方法は選択タスクと同じである．

HandyScope 条件では，被験者は HandyScope を使用してオブジェクトを回転させてドックに合わせる．オブジェクトとドックの角度が等しく（誤差 ±5 度以内）なるとオブジェクトの縁が赤色になる．この状態において操作を終了すると 1 回の試行が終了し，フィードバックとしてビープ音が鳴る．タッチ条件では，表示されたオブジェクトに手が届く位置まで移動して操作を行う．

本実験における独立変数は，開始点からドックまでの距離（86.9，106.3cm），開始点からドックまでの角度（ 15，0，15 度），ドックの大きさ（5.8，7.7cm），回転角度（ 45，45 度），操作条件（HandyScope，直接タッチ）である．被験者は各独立変数の組み合わせにおいて試行を 2 回ずつ，合計 92 試行（2 2 3 2 2 2）を行った．また，操作条件以外の試行の順序はランダムであった．

実験結果

両操作条件における 1 試行の所要時間を図 11 の中央 2 つのグラフに示す．HandyScope 条件では所要時間は 4444ms，タッチ操作では所要時間は 4520ms であった．被験者毎の所要時間において対応のある t 検定を行った結果，操作手法による所要時間に有意差はなかった（$t(9) = .267, p = .397 > .05$）．

4.5 拡大縮小タスク

被験者には提示位置に表示されるオブジェクトを拡大縮小させて，ドックに合わせて貰った．ドックはオブジェクトと同位置に異なる大きさで表示される．開始点及び提示位置，試行の開始方法は選択タスクと同じである．

HandyScope 条件では，被験者は HandyScope を使用してオブジェクトを拡大縮小してドックに合わ

せる．オブジェクトとドックの大きさが等しく（誤差 ± 4.8mm 未満）なるとオブジェクトの縁が赤色になる．この状態において操作を終了すると 1 回の試行が終了し，フィードバックとしてビープ音が鳴る．タッチ条件では，表示されたオブジェクトに手が届く位置まで移動して操作を行う．

本実験における独立変数は，開始点からドックまでの距離（86.9，106.3cm），開始点からドックまでの角度（ 15，0，15 度），ドックの大きさ（5.8，7.7cm），ドックに対するオブジェクトの拡大・縮小（1.5 倍，0.67 倍），操作条件（HandyScope，直接タッチ）である．被験者は各独立変数の組み合わせにおいて試行を 2 回ずつ，合計 92 試行（2　2　3　2　2　2）を行った．また，操作条件以外の試行の順序はランダムであった．

実験結果

両操作条件における 1 試行の所要時間を図 11 の右 2 つのグラフに示す．HandyScope 条件では所要時間は 4438ms，タッチ条件では所要時間は 4278ms であった．被験者毎の所要時間において対応のある t 検定を行った結果，操作条件間における所要時間に有意差はなかった（$t(9) = \quad .935, p = .187 > .05$）．

4.6　実験結果の考察

上述のように，選択タスクにおいては HandyScope を使用した操作が有意に速く，回転及び拡大縮小タスクにおいては有意差が現れなかった．この結果より遠隔地を選択する場合，HandyScope は有用であると言える．

一方，回転及び拡大縮小タスクにおいて操作条件間に有意差が見られなかった原因として，HandyScope の再起動に時間がかかることが挙げられる．実験において，被験者がオブジェクトをドックに合わせたつもりで HandyScope を終了したが，実際には合っていないため試行が終了せず，再度 HandyScope を起動してポインティングを行う様子が見られた．直接タッチ操作ではユーザがオブジェクト付近に移動して入力を行うため，再度入力を行う際に時間がかからないが，HandyScope は一度終了してしまうとポインティングを再び行う必要がある．これにより，HandyScope による回転及び拡大縮小タスクに時間がかかったと考える．

また，HandyScope は遠隔地を操作することを想定して設計されているため，より遠くの位置に対する操作では大きな差が出ることが予想される．HandyScope が遠距離の操作に対して有用な手法であることを確認するため，それぞれのタスクにおけるオブジェクトまでの距離毎の所要時間を用いて，操作条件における距離による特性を調べた．この結果を図 12 に示す．また，各タスクの各距離において，操作条件による有意差を確かめるため，対応のある t 検定を行った．選択タスクにおいて，オブジェクトまでの距離が 106.3cm の試行（$t(9) = 3.13, p = .006 < .01$）は，距離が 86.9cm の試行（$t(9) = 2.16, p = .029 < .05$）と比較して，操作条件間における所要時間の差がより大きかった．回転タスクにおいて，オブジェクトまでの距離が 106.3cm の試行（$t(9) = .772, p = .230 > .05$）と距離が 86.9cm の試行（$t(9) = \quad .357, p = .364 > .05$）は共に操作条件による有意性が現れなかったが，所要時間は距離が 106.3cm の試行では HandyScope 条件の方が速く，距離が 86.9cm の試行ではタッチ条件のほうが速かった．拡大縮小タスクにおいて，オブジェクトまでの距離が 3.9cm の試行（$t(9) = 1.48, p = .086 > .05$）と距離が 106.3cm の試行（$t(9) = \quad .321, p = .378 > .05$）は共に操作条件による有意性が現れなかった．これらの結果から，オブジェクトまでの距離が遠くなると HandyScope が有用な手法であることが確認できた．

本実験のタッチ操作において，遠隔地のオブジェクトを操作する際，ユーザはテーブルトップの外周に沿って 2，3 歩の移動を行っていた．オブジェクトまでの距離が遠くなると歩数が増えるため，操作条件間の有意差がより大きくなると考える．

また，本実験では HandyScope の初期の CD 比を，テーブルトップの端までポインティング可能となるように設定したため，CD 比の変更を行う被験者はほぼ見られなかった．このため，CD 比の動的な変更について評価を得るためには，改めて詳細な調査を行う必要がある．

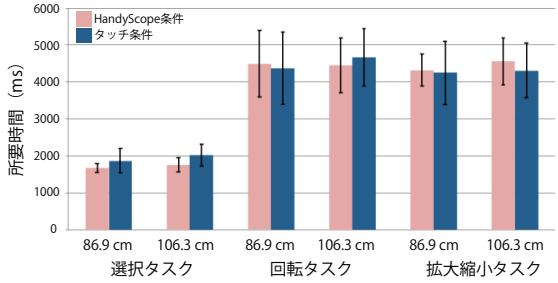


図 12. 各距離毎の 1 試行の所要時間

4.7　アンケート結果とその考察

それぞれのタスクにおける，HandyScope とタッチ操作のどちらの手法が好みであったかのアンケート結果を図 13 に示す．

選択タスクにおいては全ての被験者が HandyScope が好みであると答えた．また，拡大縮小タスクにおいては 8 人の被験者が HandyScope が好みであると答えた．この内全ての被験者がその理由として，タッチ操作ではオブジェクトに手が届く位置まで移動する必要があるが，HandyScope はその場にて操作を行うことができるためと答えた．

回転タスクにおいては好みが別れた．2 名の被験者は回転操作に両手を用いることができるため，タッ

チ操作のほうが好みであると答えた．また，別の 2 名の被験者は base-fingers を常にテーブルトップ上に接触させることが大変であったと答えた．別の被験者 1 名は試行の失敗時における HandyScope の再起動は手間であると答えた．

拡大縮小タスクにおいてタッチ操作が好みと答えた被験者 2 名は，その理由として常に base-fingers をテーブルトップ上に接触させることが大変であったと答えた．

そのため，今後は base-fingers をタッチパネル上から離しても遠隔地の操作を続けられるように，サークルを固定させる機能を実装する．

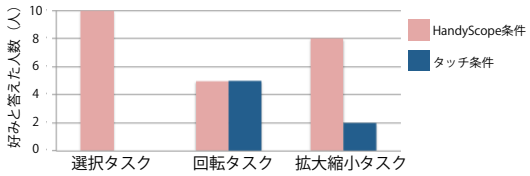


図 13. 手法の好みのアンケート結果

5 議論：引き出し方向

引き出しジェスチャは図 14 に示すように 2 種類の引き出し方向が考えられる．そこでどちらの引き出し方向が HandyScope に適しているかを調べるために，4.3 節に挙げた選択タスクを著者らが行った．その結果，図 14b の引き出し方向による操作が図 14a の引き出し方向による操作よりも所要時間が短くなった．所要時間に違いが生じた理由として，図 14a の引き出し方向では handler と scope 間の視線の移動距離が大きくなり，ポインティングが困難になることが考えられる．このため，HandyScope の引き出し方向を図 14b の方向に決定した．今後は被験者実験を行い，この妥当性を評価する予定である．

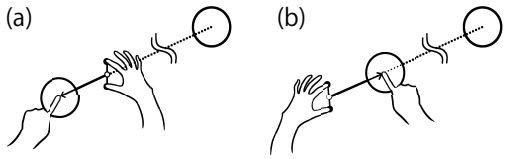


図 14. 引き出し方向

6 まとめ

我々は新たな遠隔地操作手法として HandyScope を示した．HandyScope を用いることにより，ユーザは 2 種類の円形のウィジェットを用いて遠隔地を素早くポインティングできる．ポインティング後は，遠隔地の直接操作や遠隔地と手元間におけるオブジェクトの移動が可能である．

評価実験の結果，遠隔地を選択する場合 HandyScope は有用であること及び，オブジェクトまでの距離が遠くなると HandyScope が有用な手法となることが確認できた．アンケートからは HandyScope がユーザに好まれることを確認した．今後は，base-fingers をテーブルトップから離しても遠隔地の操作を続けられるように，サークルを固定させる機能を実装する．また，CD 比を動的に変更できることの詳細な評価や，オブジェクトの移動について評価を行う．さらに，複数のディスプレイを連携させた際の本手法の有用性についても調査していきたい．

参考文献

[1] M. Abednego, J.-H. Lee, W. Moon, and J.-H. Park. I-Grabber: expanding physical reach in a large-display tabletop environment through the use of a virtual grabber. ITS ’09, pp. 61–64.

[2] A. Banerjee, J. Burstyn, A. Girouard, and R. Vertegaal. Pointable: an in-air pointing technique to manipulate out-of-reach targets on tabletops. ITS ’11, pp. 11–20.

[3] N. Banovic, F. C. Y. Li, D. Dearman, K. Yatani, and K. N. Truong. Design of unimanual multi-finger pie menu interaction. ITS ’11, pp. 120–129.

[4] T. Bartindale, C. Harrison, P. Olivier, and S. E. Hudson. SurfaceMouse: supplementing multi-touch interaction with a virtual mouse. TEI ’11, pp. 293–296.

[5] G. Furumi, D. Sakamoto, and T. Igarashi. SnapRail: a tabletop user interface widget for addressing occlusion by physical objects. ITS ’12, pp. 193–196.

[6] A. Khan, G. Fitzmaurice, D. Almeida, N. Burtnyk, and G. Kurtenbach. A remote control interface for large displays. UIST ’04, pp. 127–136.

[7] S. Malik, A. Ranjan, and R. Balakrishnan. Interacting with large displays from a distance with vision-tracked multi-finger gestural input. UIST ’05, pp. 43–52.

[8] J. Matejka, T. Grossman, J. Lo, and G. Fitzmaurice. The design and evaluation of multi-finger mouse emulation techniques. CHI ’09, pp. 1073–1082.

[9] J. K. Parker, R. L. Mandryk, and K. M. Inkpen. TractorBeam: seamless integration of local and remote pointing for tabletop displays. GI ’05, pp. 33–40.

[10] Y. Tokoro, T. Terada, and M. Tsukamoto. A pointing method using two accelerometers for wearable computing. SAC ’09, pp. 136–141.

[11] A. Toney and B. H. Thomas. Applying reach in direct manipulation user interfaces. OZCHI ’06, pp. 393–396.

[12] S. Voelker, M. Weiss, C. Wacharamanotham, and J. Borchers. Dynamic portals: a lightweight metaphor for fast object transfer on interactive surfaces. ITS ’11, pp. 158–161.

[13] T. Yoshikawa, B. Shizuki, and J. Tanaka. HandyWidgets: local widgets pulled-out from hands. ITS ’12, pp. 197–200.

[14] H. Zhang, X.-D. Yang, B. Ens, H.-N. Liang, P. Boulanger, and P. Irani. See me, see you: a lightweight method for discriminating user touches on tabletop displays. CHI ’12, pp. 2327–2336.

ライフログによる記憶拡張のための探索手法とその実践

中村 聡史*

概要． 技術の進歩により，短期作業記憶に格納したものの忘却してしまうような情報もデジタルカメラなどを利用することによって一部記録できるようになってきた．また，記録を何度も閲覧することにより長期記憶化することも可能となっている．近い将来，見たこと，聞いたこと，体験したことをすべて記録できるようになると期待される．しかし，記憶を記録できるようになったとしても，記録を検索して記憶として引き出すことができないようでは折角の記録も役に立たない．本稿では，デジタルカメラを利用してユーザ自身のタイミングで記録をする一人称視点での主観写真ライフログに注目し，写真ベースの膨大なライフログを効果的に活用する探索手法を提案および実装する．また，10 年近いライフログの利活用から見えてきた利点や問題点などを整理する．

1 はじめに

人間は記憶して，過去を振り返ることが可能な生き物である．

Atkinson ら [1] の二重貯蔵モデルによると，人間の記憶過程をまず入力刺激を瞬間的に記憶しすぐに忘れる感覚記憶，そして感覚記憶から選択的に入力され 20〜30 秒程度保持することが可能な短期作業記憶，短期作業記憶から処理を行った結果を格納する長期記憶とに分類される．

短期作業記憶は 20〜30 秒しか保持することができないため，多くの場合情報が忘却されていってしまい，思い出すことができなくなる．長期記憶化することによって思い出すことが可能となるが，長期記憶化にはリハーサルを繰返すことや，深い意味的処理を施すなどすることが必要であり，長期記憶を行うことは容易ではない．

ここで，短期作業記憶を長期的に記憶（記録）する方法として，デジタルカメラなどの記録装置を使う方法が考えられる．デジタルカメラにより，旅行で見かけた素晴らしい風景や現地で出会った印象的な人々，道端で出会った印象的な看板や参考になる講演スライド，配布資料に書き込んだメモや田舎のバス停の時刻表，日々食している料理からお店で食べた印象的な料理，友人と会うときに着ていた服や子どもの成長記録など，見て興味をもった様々なものを記録することが容易な時代になっている．つまり，すぐに揮発してしまうような短期作業記憶に記憶される情報をデバイス上に記録することによって，いつでも引き出すことが可能となる．

ここで，記憶を記録し，記録から記憶を呼び覚ますことを目的として個人の一人称視点および興味でデジタル写真を撮影することを，本稿では主観写真

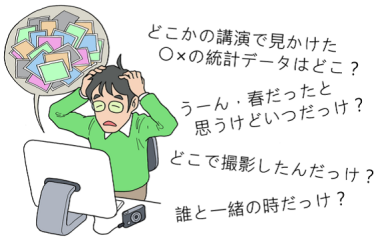


図 1. 探している写真はどこに？

ライフログ[2] と呼び，記憶されたデジタル写真をライフログ写真と呼ぶ．著者は 1998 年頃より 2005 年までは何らかのイベント時にデジタル写真を撮影し，2005 年 6 月 2 日から現在に至るまでは毎日ライフログ写真を撮影および記録しており，2013 年 7 月 24 日時点で合計 254,899 枚となっている（約 349GB）．緯度経度が付与されたものが 193,545 枚，Google Picasa によって顔認識されている人の数が 1,398 人，顔の数が 71,054 個となっている．

デジタル写真は手軽に撮影することができるため主観写真ライフログが容易になっているにも関わらず，一般的になっていない．その一番の理由はそもそも記録や管理するのが面倒というものであるが，それ以外にも検索が難しいという問題がある．ライフログ写真には一般的にテキスト情報が付与されていないため，キーワード検索は困難である．また，ユーザは過去の事になればなるほど記憶が定かではなくなっており，目的とするライフログ写真の撮影場所や撮影時間，一緒にいた人の名前などを忘れているだけでなく，間違って記憶していることも多い（図 1）．さらに，検索条件を明確にすることが難しいという問題もある．

そこで我々は，検索の問題を解決するため，時空間と人間関係を利用し，絞り込み探索を行うことで，目標とするライフログ写真への到達を支援する探索

Copyright is held by the author(s).

* 明治大学 総合数理学部

手法を実現する．また，ライフログ写真を長期記憶化するための受動的閲覧手法を実現する．さらに，長期的な利用からライフログ写真における人間関係がどの程度利用可能かといったことについて議論を行う．

2　写真ライフログ

ライフログにデジタル写真を使うメリットは，静止画であるため一覧化しやすく，現在発売されているデバイスを利用するだけで手軽にライフログを実践することが可能であることなどである．ここで，写真ベースのライフログは，主観的なライフログ（主観写真ライフログ）と客観的なライフログ（客観写真ライフログ）に大別される [2]．主観写真ライフログとは，撮影時においてユーザの明確な意図があり，ユーザのタイミングと視点で空間を切りとるものである．

客観写真ライフログとは，システムにより自動記録されるものであり，SenseCam などウェアラブルカメラを利用した自動記録がそれに相当する．また，ユビキタスホームなどでのユーザの記録も客観的ライフログといえる．SenseCam は首から下げて利用する客観写真ライフログ用のカメラであり，1 分間に 2 回または，センサ（照度，地磁気，加速度センサなど）で取得した外界の状況に応じて自動撮影されるようになっている．客観写真ライフログは自動撮影されるため，自然に撮影できることや，何らかの作業中に撮影できることなどメリットがある．一方，撮影するだけではユーザの視点が抜け落ちるため，後に活用しようと思った際に何をキーとして探したらよいかわからないなどの問題がある．

主観写真ライフログと客観写真ライフログの特性の違いについて詳細に分析した研究は存在しないが，Gurrin らは SenseCam を用いた客観写真ライフログによる 1 年間の記録（100 万枚）から無作為に 1000 枚を抽出および分析している [5]．その結果，客観写真ライフログには腕や手がよく写っており，仕事中（24.6 %）や運転中（16.2 %），会話中（13.3 %）などの写真が多く，人を自然に記録できるとしている．また，インドアの写真が 73.4 % と多く，風景写真は 1.1 % とかなり少ないことを報告している．さらに，29.9 % の写真に人が写っていることを報告している．一方，著者の主観写真ライフログの中から 1000 枚を無作為に抽出して調査したところ，作業中の写真はほとんどないが，作業の後などに撮影したものや，他人が作業している様子や，他の人が会話している様子を撮影していたものが多かった．また，65.6 % がインドア，34.4 % がアウトドアの写真であり 13.2 % が風景写真であった．さらに，44.0 % の写真に人が写っていた．単純な比較はできないが，主観，客観ともに人が写っている傾向が高いといえる．

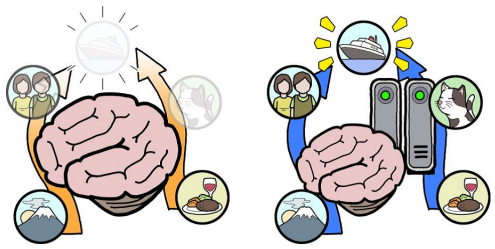


図 2. ライフログ（外部記憶）により記憶拡張

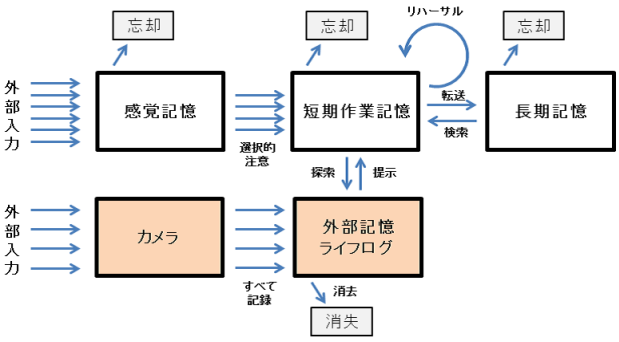


図 3. ライフログによる記憶のモデル

ここで，基本的に人間は画像を膨大かつ長期的に記憶することが可能である [7] ものの，Isola らの写真の覚えやすさに関する研究 [6] によると，写真のタイプごとに記憶の容易性が異なり，人を含んだ写真を最も覚えやすく，次に人の大きさのオブジェクトが被写体となっている写真を覚えやすいのに対し，風景のような写真は綺麗であっても覚えにくい事がわかっている．つまり，写真ライフログでは，時空間をベースとしつつも，人間関係を効果的に利用した探索手法および，人を考慮した長期記憶化手法が有効になるのではと考えられる．

そこで，本稿ではこのような人間関係を中心としたライフログ探索を可能とする仕組みを実現する．ここでは，ライフログ写真から人間関係を抽出および可視化し，対話的に探索可能とする．

3　ライフログの外部記憶化手法

先述の通り，短期作業記憶はすぐに失われてしまうものであり，長期記憶への定着も難しい．ここで，主観写真ライフログを記憶として使う（図 2）と，図 3 のように消去しない限り短期作業記憶へ何度も呼び出すことが可能となる．そこで，本稿ではライフログを外部記憶として扱うために，短期作業記憶へと呼び出すための探索手法と，何度もユーザに提示し，リハーサルを行えるようにすることで長期記憶に定着させるための受動的閲覧手法を提案する．

3.1 人間関係を利用した探索

デジタル写真を撮影すると，撮影日時やカメラの機種名，各種撮影情報などが EXIF 領域に自動記録される．また，最近では GPS 搭載のデジタルカメラも増えており．撮影場所情報も自動的に記録されるようになっている．さらに，Google Picasa や Apple iPhoto などのように，写真ライブラリ内の写真に対する顔画像認識を行うソフトウェアも近年広がりを見せている．我々は，こうした撮影日時，撮影場所，被写体情報を写真とともに取り込み，データベースに格納し，時間，空間，人間での検索を可能とする．

しかし，人の記憶は曖昧であり，数ヶ月，数年前のこととなると，目標とするライフログ写真をいつ撮影したのか，どこで撮影したのか，誰と撮影したのかといった事も正確に覚えていない．つまり，目標とするライフログ写真への到達も一筋縄ではいかない．ライフログ写真の検索および探索においては，それを思い出すきっかけとなる柔軟な絞り込みと周辺情報提示が必要となる．

ここで，ライフログ写真として蓄積されているものは，ユーザ（撮影者）が見て切り取ったものであり，それぞれの写真に写っている人物を認識すると，ユーザからみた人物関係（主観的視点におけるソーシャルネットワーク）が見えてくる．つまり，人を 1 つのノードとし，同一の写真に写っている人には関係があるとみなしてその人物間（ノード間）にエッジをはり，そのエッジがどのようにはられているかということを分析してクラスタリングを行うことで，コミュニティを検出することが可能となる．そこで，本稿ではある時空間での写真集合からすべての人間関係を抽出し，その人間関係を Newman 法 [3] によってハードクラスタリングした結果とともに提示することにより，人間関係を考慮した探索を可能とする．

なお，ハードクラスタリングした結果の人間関係をバネモデルを用いて提示し，提示されている人間関係をもとに検索する手法を実現する．ここでバネモデルにおけるノード間の引力はユーザ間の親密性の高さに基づき，強さを変更している．また，人と人との間の親密性については，「同じ写真に共起していること」「同じ時分に共起していること」「同じ時間に共起していること」「同じ日に共起していること」という 4 パターンを用意し，その共起頻度によって計算している．

3.2 受動的閲覧による長期記憶への格納

近年，横長の液晶ディスプレイが広く普及している．横長のディスプレイは，動画を視聴する場合や，ゲームをする場合などには効果的であるが，作業を行なっている場合には，2 つのウィンドウを表示するには少し狭すぎ，1 つのウィンドウを表示するには広すぎる．ここで，1 つのウィンドウのみを表示するのであれば，画面の両端に余裕がある．

受動的閲覧による，外部記憶からの長期記憶化では，こうした画面の両端のスペースを有効活用し，利用者とライフログ写真との遭遇を促進し，その事自体が短期記憶のリハーサルとして働くようにすることによって，長期記憶化を狙うものである．受動的閲覧において，システムは特に人間にフォーカスを当て，任意の人間を取り出し，その人間を中心とした写真（その人間と一緒に居た場所や，居た時間を利用）中心とした n 件のライフログ写真をデータベースから取り出し，1 枚ずつ写真を上から下にスクロールさせながらアニメーション提示するものである．

4 実装

提案システムを Microsoft Visual C++を利用して，データベースに SQLite を利用して実装した．また，サムネイル画像を必要に応じて自動生成することにより，20 万枚以上でも問題なく動作する仕組みを実現している．さらに，データベースやサムネイルはバックグラウンドで自動的に構築および生成されるため，初めて利用する際にも特に負荷なく利用可能である．システムは LifelogViewer という名前で Web 公開している[1]．

図 4 は，カレンダーインタフェース，地図インタフェース，人間関係ビューを利用して時間および空間を指定しながら目的とする食事のライフログ写真を探索している様子である．カレンダーインタフェースで表示月を変更している際に提示される人間関係ビューから，まずどの月かを限定し，次に地図インタフェースを用いてズームし表示領域を限定することで提示される人間関係のクラスタを限定している．さらに，そこから目的とする人が含まれるクラスタを指定して検索し，最終的に目的の写真にたどり着くことができている．図 5 は，ライフログの受動的閲覧手法を利用して，作業中にライフログを提示している様子である．

5 考察

時間，空間，人間関係を考慮して 20 万枚の写真ライフログを扱える環境は限られている．Apple の iPhoto は 1 操作辺りに数十秒かかるため利用できず，Google の Picasa も頻繁に固まってしまうという問題が生じた．速度面は無視して，著者がこれまでに必要とした様々なライフログ探索について提案システムと Picasa の比較を行った．大きな違いは，Picasa ではユーザは明確に場所や時間，被写体を選べる場合は効果的に目標とするライフログ写真を探しだすことができるものの，少しでも記憶が曖昧な

[1] http://calendar2.org/

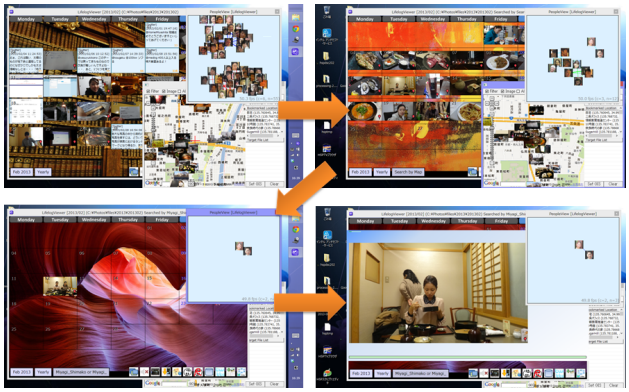


図 4. 時空間と人間関係を利用した写真探索

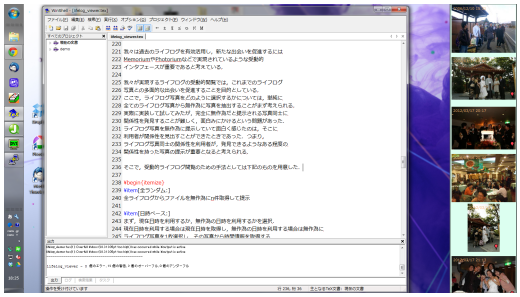


図 5. ライフログの受動的閲覧（画面右端）

場合には探すことが困難であるという点である．一方，提案システムでは，ユーザの記憶が曖昧な場合であっても対象となる写真の発見が可能であり，検索達成率は高かった．

本手法によってある特定の月の人間関係がどうなるかを示したものが図 6 である．上段は，私が京都大学から明治大学へと転職前（2013/02）と，転職後（2013/04）で大きな違いが現れており，2013/04 では転職先の大学での学生が登場し密なネットワークを作っている事がわかる．また，2013/03 では，転職元や転職先の方々，京都の友人の方々などとの交流があったため，様々な広がりのあるコミュニティができている．下段は，WISS が開催される 12 月について，2010，2011，2012 年を比較したものである．2010 は右上に，2011 は左下に，2012 は下にそれぞれ WISS のコミュニティが形作られている．また，2010 年は WISS のコミュニティは独立しているが，2011，2012 年については他のコミュニティとの関係が深まっていることが分かる．これは，2011 年，2012 年の 12 月にそれぞれ開かれたニコニコ学会に WISS 関係者が関与しているためである．

クラスタリングにおいて，知人と解釈する度合いを「同じ写真での共起」「同じ時分での共起」「同じ時間での共起」「同じ日での共起」と 4 つ用意し比較した．様々な時空間でのクラスタリング精度について実際に運用することで比較してみたところ，2005 年までは「同じ日での共起」，2006 年〜2008 年に

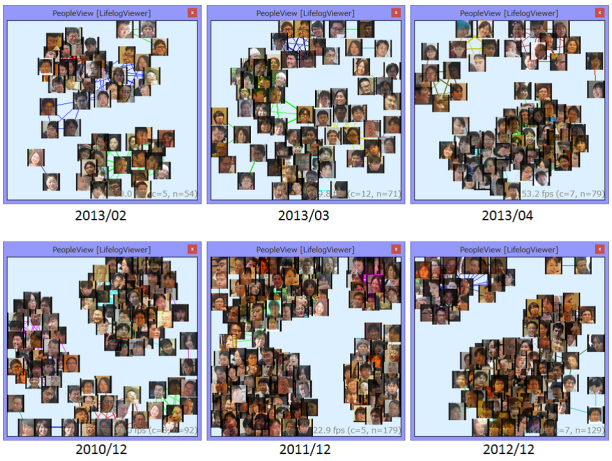


図 6. 人間関係の変化

ついては「同じ時間での共起」，2009 年以降については「同じ時分での共起」と「同じ写真での共起」が良い結果となっていた．これは，撮影頻度と，被写体として人間を選ぶ度合いとに深く依存しているのではないかと考えられる．実際，2005 年までは撮影頻度は高くなかったが，2006 年から 2008 年については撮影頻度があがり，2009 年以降も撮影頻度はこれまでと大きく変化していないものの，著者の撮影スタイルが変化していき人が被写体となっている量が増加したことも理由の 1 つとして考えられる．撮影頻度や撮影場所によってイベントを分離することが可能になれば，本問題は解決する可能性を秘めている．

受動的閲覧が作業の邪魔になるかどうかを調べるため，1 年以上の長期的な利用を実践した．その結果として，受動的閲覧の仕組みは，作業中はほとんど目に入らず，ふと気が緩んだ時に目に入る程度であり，著者自身は問題がなかった．むしろ，集中して作業をし，気を抜いた時に，知人の楽しげな写真や，娘や妻などの写真が目に入り，休憩時の癒やしとなっていた．また，受動的閲覧で提示される写真の中のグラフなどから着想を得て，講義資料に取り込んだこともあった．以上のことより，長期記憶に定着させるための方法として有効であったかは不明であるが，少なくとも受動的閲覧は有効であった．

この 1 年間でライフログが役に立った事例は多数あるが，そこから選別したものは下記の通りである．

- 出産育児一時金請求書を記入する際に医療施設の名称＆住所が必要となった．妻と沖縄で一緒だった時に出生届を書いた事を思い出し，地図 UI を開いて「沖縄」を指定．表示される人間関係 UI から妻の顔をダブルクリックし，カレンダー UI に提示された写真の中から妻がベッドに座っている写真を発見．そういえばこの時だったと，その日の写真を一覧化し出生届＆出生証明の写真とその情報を発見．

- 転居先の水道代自動引落しの手続き時に，銀行への届け出サインが必要に．銀行での契約時以降サインを使った覚えがないため日本語か英語か記憶が無い．写真の GPS 情報がずれているのか契約した場所で検索しても写真が見つからない．口座を使い始めたのは入籍時である事を思い出し，入籍日あたりを探索．契約後に，妻と飲みに行ったお店の店長が人間関係ビューに提示され，ふとその事を思い出して周辺を探索．書類を発見し，日本語だったことを確認して記入．

- 耐震工事のためのオフィスの一時的な引っ越しをしていて，1 年後に戻ることになった時，「前のオフィスに時計ってありましたっけ？」「あったとしたらどんなのでしたっけ？」という話になり，数名の同僚と一緒に時計が写っている写真を受動的閲覧で過去に見かけたことを思い出し，場所を人間関係 UI を提示しながら過去に遡ることによって，木の枠のこんなダサい感じの時計であったことを発見．

他にも，契約解除時にすでに忘れてしまった ID を調べたり，田舎で運行本数が極端に少ないバス停でカメラに運行表を記録させ後で活用したり，著者の妻が著者の友人に会った時にどのような服を着ていたかを探したりなど，その活用方法は様々である．

こうした探索は，時空間や人間関係があったからこそ可能となっているものが多い．また，受動的閲覧によって，何度も任意の人が被写体となっている写真と，その人の周辺の写真（一緒に作業した記録や，一緒に摂った食事，その前後のスケジュールに関するものなど）が連続的に提示されているため，自然と長期記憶に定着し，その定着した「人」と「周辺の写真」から目的とする写真を探索可能となっているケースも多い．

一方，長期的な利用から得られた問題としては，まず忘れることができないという事が大きな理由としてあげられる．例えば，ある時まで恋愛関係にあった人との別れは，その別れの前までは幸せな記録として残していたのに，別れの後では辛いものとしてただ存在するだけである．ライフログ写真の受動的閲覧では，こういった辛い過去もどんどん提示されてしまう．こういった，忘れることができない何かをどうするかというのは，今後の大きな課題の 1 つである．

また，美術館内や映画館内など，撮影できない場所は多数ある．こういった状況において，本人としては記録した覚えがあるのだけれど，実際には撮影および記録しておらず，探索しても見つからないということがある．そうした，存在しないライフログ（見つからない記録）をどう提示するかというのは課題の 1 つである．

ライフログを長く続けていくと，ライフログが無いと困るというライフログ依存症をといえるのではと思えることが多々ある．例えば，すぐにデジタルカメラにて記録することができるため，無理覚える必要がなく短期作業記憶が劣化しているように感じることや，受動的閲覧で何度でも見るため，ある出来事が遠い過去のはずなのについ最近であるかのように感じることなどである．こういった問題は，今後ライフログが世の中に広まると，社会問題となるのではと考えている．

先日，著者のパソコンの SSD が突如読み込めなくなってしまい，最後にバックアップを取った日から，20 日間分のライフログデータが失われてしまった．このようなデータの損失は，現代の記憶喪失といえるかもしれない．こういった戻せない記録をどう保管するか，また補完するかといったことも大きな課題となるのではないだろうか．

6 関連研究

SmartWrite/SmartCalendar[9] は，写真・メモベースのライフログであり，時空間をベースとした写真の探索を可能としているが人間関係は考慮されていない．我々の手法は時間のみならず，空間や人間関係を考慮した閲覧および検索を可能としており，対象の柔軟な探索が可能である．

五味らは，時間，空間，人間関係の 3 者を利用したライフログ写真の閲覧手法を提案している [10]．しかし，五味らの手法では，あらかじめ時間ベースで全ての写真のクラスタリングを行なってイベントを抽出し，イベントにおける共起度をもとに人物のクラスタリングを行なっているため，任意の時間や場所を指定した動的なクラスタリングは行えない．

捧ら [8] も，時間，空間，人間関係を利用してライフログ写真を閲覧する手法を提案している．ここでは，写真における共起性とその前後の写真から，人物同士の親密度を計算し，クラスタリングを行うというものである．ただし任意の時空間を対象とはしていないうえ，計算コストが高く，リアルタイムの分析には不適である．

Facebook，mixi などの SNS の友人関係からソーシャルグラフを作る試みは多数研究されている [4]．しかし，こうした研究では，グラフは明示的なリンク（フォロー，友達として登録など）を元にしている．一方，ライフログの中には明確な友人関係に関する情報は存在していない．また，こうした人間関係は日々変化するものであり，関係が強まったり（他機関の知人から同僚へ，赤の他人から人生の伴侶へなど），関係が弱まったり（交際相手からただの知人へ，クラスメートから他大学の友人へなど）する．我々は，ライフログ写真の中から 2 者間の関係性の強さを計算し，人間関係の変化を考慮しつつ，探索を可能とする手法を実現するものである．

7 まとめ

本稿では，主観写真ライフログの能動的探索および受動的閲覧手法を提案および実装し，その有効性について検討を行った．主観的な評価にすぎないが，長期にわたって継続利用しており，今まで到達できなかった探索を可能となり，主観写真ライフログの記憶拡張としての可能性を明らかにした．

システムは Web 公開および雑誌などで紹介されており，多数のダウンロードがある．今後は，主観写真ライフログを長期にわたって継続しているユーザに依頼し，システムの継続的な評価を行う予定である．

謝辞

本研究は文部科学省科学研究費補助金若手研究（A）（研究代表者：中村聡史，#23680006）の助成を受けたものです．

参考文献

[1] Atkinson, R.C., Shiffrin, R.M: Human memory: A proposed system and its control processes, The psychology of learning and motivation (Volume 2), pp.89-195 (1968).

[2] 中村聡史: 身近になったライフログ ～パソコンに眠る数万枚の写真をどう活用すればよいか～, 情報処理 54(2), pp.142-149 (2013).

[3] M.E.J. Newman：Fast algorithm for detecting community structure in networks. In Physical Review. E 69 (2004).

[4] Ken Wakita, Toshiyuki Tsurumi: Finding community structure in megascale social networks, Proceedings of the 16th international conference on World Wide Web, pp. 1275-1276 (2007).

[5] Cathal Gurrin, Alan F. Smeaton, Daragh Byrne, Neil O'Hare, Gareth J. F. Jones, Noel O'Connor: An examination of a large visual lifelog, Proc. of AIRS'08, pp. 537-542 (2008).

[6] Isola, P., Xiao, J., Torralba, A., and Oliva, A: What makes an image memorable? Proc. of the 24rd IEEE Conference on Computer Vision and Pattern Recognition, pp. 145-152 (2011).

[7] Lionel Standing: Learning 10,000 Pictures, Quarterly Journal of Experimental Psychology, vol.25, pp.207-222 (1973).

[8] 捧隆二, 佃洸摂, 中村聡史, 田中克己: 時間・空間・人物情報に基づくインタラクションによるライフログ画像の探索手法の提案, DEIM Forum 2012 D9-4.

[9] 美崎薫: SmartWrite/SmartCalendar 手軽に書けるメモとメモと写真を見続けるカレンダー環境の提案，情報処理学会ヒューマンインタフェース研究会報告 2005(71), 71-76 (2005).

[10] 五味愛, 伊藤貴之：「何時，何処で，誰と」3 つのメタ情報に基づく個人写真ブラウザ, 芸術科学会論文誌, Vol. 10, No. 1, pp. 36-47 (2011).

未来ビジョン

人の記憶は曖昧なものであり，過去に見聞きしたものを思い出せないことは多い．忘れることは必ずしも悪ではなく，痛ましい出来事や，友人知人との別れ，恐怖を覚えた出来事や，昔の彼女との辛い思い出など，思い出したくないことも多々あるであろう．しかし，思い出せないことによって失敗をしてしまうことも少なくない．

今後，技術の進歩により，ウェアラブルコンピューティングが今の携帯電話のように人々にとってありふれた普通のものとなり，また生活の中での視線検出技術の進歩や脳波測定技術の進歩などにより，主観的なライフログ写真の記録がより簡単になり，ライフログ写真のインデックスもより深いものになると期待される．将来，見聞きしたものを全て外部記憶に記録し，記録したものを検索などにより取り出すことも可能となるであろう．

つまり，忘れることができなくなる時代が到来することになる．もちろん，そうしたデバイスを使わないという選択しはありえるだろう．しかし，記憶拡張された人と記憶力という面で圧倒的に差が広がってしまい，嫌々ながらもそうしたデバイスを使うユーザが増えるのではと考えられる．

本研究はこうした忘れることができなくなることは，どういったメリットがあり，どういったデメリットがあるのかということを，実践的に取り組むものである．また，人の様々な記憶の仕組みとの関係を考えていくことにより，ライフログによる記憶拡張がどのように記憶を変えていくのか，社会を変革させていくのかについて考えるものである．

EpisoPass: エピソード記憶にもとづくパスワード管理

増井 俊之 *

概要． 忘れる可能性が低いエピソード記憶にもとづく秘密の質問を使って強力なパスワードを生成/管理するシステム「EpisoPass」を提案する．EpisoPass は，ユーザが作成した秘密の質問への回答にもとづいてシード文字列を換字することによってパスワードを生成する．シード文字列や回答のバリエーションにより異なるパスワードが生成されるので様々なサービスに対して異なるパスワードを生成できることに加え，シード文字列を逆計算することにより既存のパスワードの管理もできる．適切な運用により，パスワードに関連するあらゆる情報を秘密にすることなく強力なパスワードの生成/管理が可能である．

1 はじめに

個人認証のためにパスワードが現在広く利用されている．パスワード認証には多くの問題があることが知られているが [21]，今後も長期にわたって利用され続けると予想されるため [8]，問題点を認識しつつ適切に運用するための工夫が必要である．

パスワードの長期的記憶が難しいことはパスワード認証の大きな問題点のひとつである．安全に運用するためにはパスワードはランダムで長い文字列であることが望ましいが，そのようなものを頭の中に記憶しておくことは難しい．また複数のサービスを利用する場合，サービスごとに異なるパスワードを利用することが望ましいが，すべてのパスワードを記憶しておくことはほとんど不可能である．Florêncio の 2007 年の大規模な調査によれば，ユーザは平均 25 個のサイトで 6.5 個のパスワードを利用しており，3ヶ月間にユーザの 4.28%がパスワードを忘れていた [5]．また 2011 年の野村総研の調査によれば，一般的なユーザがパスワード認証を行なうサイトは平均 19.4 個で，利用しているパスワードは平均 3.1 個であった [23]．多数のパスワードを記憶することが困難であるため，多くのユーザが同じパスワードを複数サイトで使い回しているのだと思われる．

異なるパスワードをすべて記憶することは不可能なのでどこかに記録しておく必要があるが，パスワード文字列をそのまま記録するのは危険なので，複数のパスワードを秘密情報として扱うためのパスワード管理システムが利用されている．パスワード管理システムはひとつの「マスターパスワード」を利用して他のすべてのパスワードを管理するもので，暗号化されたデータベースにパスワードを格納するもの [1][3][12] が多いが，サービス名をもとにマスターパスワードを変換することによって複数のパスワードを生成するシステム [18] もある．両者ともにマスターパスワードの記憶が必須であり，マスターパスワードを盗まれたり忘れたりする危険が常に存在する．

一般にユーザはパスワードを忘れがちであるため，多くのサービスにおいてパスワードを復元したり初期化したりする手段が用意されている．ユーザが秘密の質問に対する答を登録し，質問に正しく回答することによってパスワードを復元したりリセットできるサービスは多いし，秘密の質問に答えることによってパスワード管理システムのマスターパスワードを復元するシステム [25] も提案されている．

新しく覚えた情報や新しく考えた情報はどうしても忘れてしまう可能性があるので，新しく作成したパスワード文字列を記憶して認証に利用することは本質的に無理がある．一方，既知で忘れることがないエピソード記憶を秘密の質問として認証のために直接利用することができれば，認証に必要な情報を忘れてしまうことがないはずである．多くの画像認証システム [2][19] は秘密の質問に対して適切な操作を行なうことによって認証を行なっているためパスワードのような特種な情報を記憶する必要がない．画像認証システムはまだ普及しておらず，利用できる環境は限られているが，忘れないエピソード記憶を利用した秘密の質問への回答を強力なパスワードに変換するシステムがあれば，通常のパスワード認証を用いた現在の様々なサービス上でも，認証方法を忘れる心配なく安全に認証を行なうことができるようになる．本論文ではこのようなシステム「EpisoPass」について述べる．

2 EpisoPass

2.1 EpisoPass の原理

EpisoPass は，ユーザが忘れることがない個人的なエピソード記憶を文字列に変換することによって安全なパスワードを生成するシステムである．パスワード文字列は以下の手順で生成される．

Copyright is held by the author(s).

* Toshiyuki Masui, 慶應義塾大学 環境情報学部

1. パスワード生成の「種」となる文字列を用意する．以下ではこれを「シード文字列」と表現する．

2. 忘れることがない個人的なエピソード記憶にもとづく秘密の質問を複数作成し，それぞれについてひとつの正答と複数の偽答を用意する．

3. 質問と回答の組にもとづいてシード文字列に換字操作を行なう．すべてに正しく回答したとき生成される文字列をパスワードとして利用する．

2.2 EpisoPass 利用例

2.2.1 ブラウザでの利用

筆者が twitter のパスワードを生成するためにブラウザで EpisoPass を利用している例を図 1 に示す[1]．シード文字列として「Twitter123456」という文字列を指定しており，4 個の秘密の質問に対する回答選択に応じて「Mfveabn574923」のようなパスワード候補が生成される．異なる答を選択したり異なるシードを指定すると全く異なる文字列が生成される．シード文字列の 8 文字目が数字である場合はパスワードの 8 文字目も数字になるなど，シード文字列の文字種に対応したパスワード候補が生成される．パスワードとして大文字/小文字英数字と記号をすべて利用しなければならないサービスの場合はシード文字列に「PassWord123!@#」のような文字列を指定すればよい．

最初の秘密の質問は筆者の小学校の同級生に関するもので，最後の質問は数年前の体験に関するものである．これらの質問は古いエピソード記憶にもとづいており，筆者が将来答を忘れることはほとんど考えられないが，本人以外がこのような質問に答えることは難しいので正しいパスワードを得ることはできない．

秘密の質問と答はブラウザで編集でき，右上の「サーバにセーブ」ボタンを押すことによりシード文字列，秘密の問題，答のリストがサーバにセーブされる．「ファイルにセーブ」ボタンを押すと JSON データをパソコンにダウンロードでき，パソコン上の JSON データをブラウザにドラッグ&ドロップするとサーバにアップロードできる．ユーザはどれが正答かを指定するわけではないので問題データを見てもユーザのパスワードはわからない．

シード文字列を「Facebook123456」に変更すると，生成されるパスワードは図 2 のように変化する．このように，サービスごとに異なるシード文字列を利用することによって様々なパスワードを簡単に生成できる．

[1] http://EpisoPass.com/masui

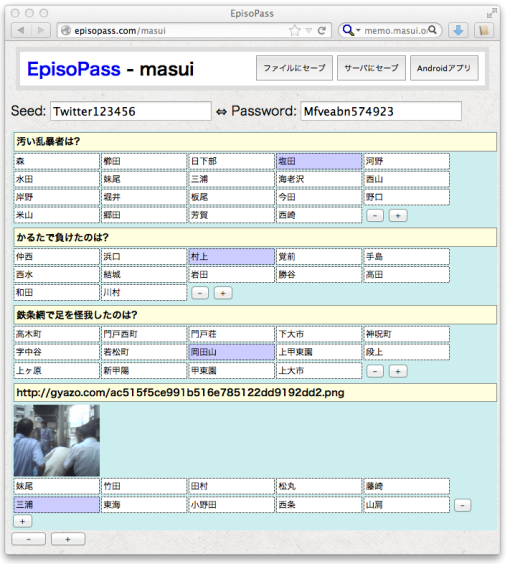


図 1. ブラウザ上で Twitter のパスワードを生成．



図 2. FaceBook のパスワードを計算．

2.2.2 既存パスワードの利用

現在 “Masui1234” のようなパスワードを利用している場合，図 3 のようにパスワード欄に現在のパスワードを入力すればそれを生成するシード文字列が自動生成されるので，生成されたシード文字列を記録しておけばよい．

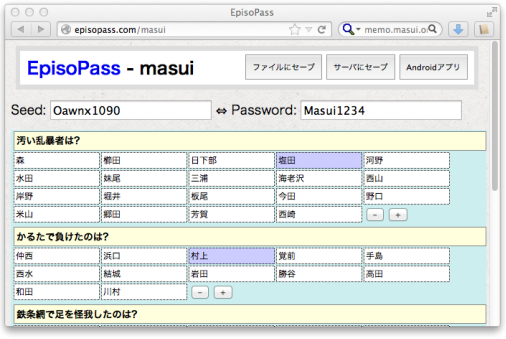


図 3. 既存のパスワードを利用．

既存のパスワード管理システムは，利用中のパスワードを記憶するもの [1][3][12] と新しいパスワードを生成するもの [18] に分類されるが，EpisoPass はこの両方をサポートしている．

2.2.3 Android アプリ

Web サービスを利用する場合，ブラウザとサーバとの間の通信を記録されたり盗み見されたりされる心配を完全に払拭することはできない．前述の例において，パスワードはブラウザ内部で JavaScript により生成されるので，一度ページを表示した後はネットワークを遮断してもパスワード計算を行なえるようになっているが，最初から全く通信を行なわずにパスワードを作成できる方がより安心であろう．このため，通信を全く行なわずにマシン単体でパスワード計算を行なうための Android アプリを用意した．ページの右上の「Android アプリ」ボタンを押すと，現在表示している秘密の問題と答を内蔵した Android アプリがサーバ上でビルドされてダウンロードされる．

シード文字列を設定して Android アプリを実行すると図 4 のように質問がひとつずつ表示され，ボタンを押してすべて回答するとパスワードが計算され表示される．

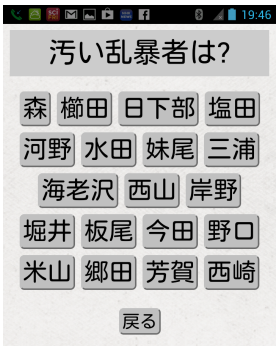


図 4. Android アプリ．

回答入力とパスワード計算は Android 端末で実行されるため，端末を「機内モード」に設定するなどの方法でネットワーク接続を遮断した状態でもパスワードを計算することができる．EpisoPass をインストールした Android 端末を持っていれば常に各種のパスワードを計算できるので，他人のマシンや公共の場所に設置されたパソコンなどでも容易にネットサービスを利用することができる．

前述の方法で EpisoPass アプリをサーバからダウンロードする場合は，ブラウザ上で秘密の問題をサーバに登録する必要があるが，秘密の問題を全くネット上に露出することなくアプリを利用することもできる．秘密の問題を含まない EpisoPass アプリを Google Play で公開 [2] しているので，これを端末にインストールした後，ローカルマシンで作成した秘密の質問を端末に転送すれば EpisoPass.com からダウンロードしたアプリと同様に利用できる．この手法を使うと秘密の質問が通信路を通ることがないので安全であるが，アプリのセットアップの手間は増える．

[2] https://play.google.com/store/apps/details?id=com.pitecan.episo_pass

2.3 パスワード文字列の計算方法

問題と回答から文字列を生成し，その MD5 値によってシード文字列を換字することによりパスワードを生成している．パスワード文字列の計算方法は附録に示す．

3 議論

3.1 運用形態と安全性

パスワード管理システムにおいて最も重要なのは安全性である．EpisoPass は様々な運用方法が可能であり，運用形態によって安全性の評価が異なるので，利用例にもとづいて EpisoPass の安全性を考える．

3.1.1 EpisoPass の利用を秘密にする場合

パスワード管理に EpisoPass を利用していることを公開せず，単なるパスワード生成機として利用する場合は，普通にパスワードを利用する場合に比べて安全度が低下することは無い．パスワードとして充分強力な文字列をシード文字列として利用すれば EpisoPass によって換字された文字列もパスワードとして強力だと考えられるし，推測しやすい文字列をシード文字列として設定した場合であってもランダムな換字操作によってより強力なパスワードが生成されるので，EpisoPass を利用するデメリットは無く，通常のパスワード管理システムと同様の方法で利用できる．

3.1.2 シード文字列を秘密にする場合

EpisoPass の秘密の質問を公開した場合でも，シード文字列を通常のパスワードと同じレベルで秘密にしておけば SuperGenPass[18] と同じレベルの利便性と安全性が確保できる．

3.1.3 すべて公開する場合

秘密の質問を解くことが不可能であれば，シード文字列と秘密の質問をすべて公開しても安全である．この場合，ユーザは秘密の質問とシード文字列を通常のテキストデータと同じように管理できるし，パスワード管理のために新しく記憶しなければならない情報が皆無なので，ユーザはパスワード管理について注意を払う必要が無くなり気が楽になる．

秘密の質問を解くことを困難にするためには質問の数と偽答の数が多くなければならないし，自分だけが知っているエピソード記憶をうまく秘密の質問にするにはコツが必要である．良い秘密の質問を作る方法に関しては 3.3 で議論するが，すべての質問を公開するのが心配な場合や，秘密の質問の数や質

がが充分でないと感じられる場合は 3.1.1 や 3.1.2 の手法で運用し，徐々に運用方法を変えていけば良いだろう．

3.2 秘密の質問の強度

パスワードは長年利用されているため強度や実際の運用に関して多くの研究が存在するが [7][11]，秘密の質問の強度に関しては充分な研究が行なわれていない．EpisoPass の運用実績は長くないが安全性などについて考察を行なう．

EpisoPass で選択枝が 10 個の秘密の質問を 8 個使用する場合，総当たりでパスワードを生成するには 1 億 (10^8) 通りの試行が必要であり，エントロピーは 26.6 ビットとなる．英字からランダムに 8 文字を並べて作成したパスワードのエントロピーは 37.6 ビットになるが，“pmvixuzq” のように全く意味のないパスワードを記憶して利用することは少ないため，実際に利用されるパスワードのエントロピーは 20 ビット程度と考えられているので [14]，秘密の質問と選択枝の数を 10 個程度用意すれば通常のパスワードと同程度の強度が期待できることになる．総当たり攻撃が可能なオフライン運用ではエントロピーの大きさは重要であるが，オンラインサービスではパスワード入力を何度か間違えるとサービスがブロックされるのが普通なので，それほど長いパスワードを用意する必要は無いと考えられている [6]．

一方，秘密の質問を利用する認証の脆弱性を利用した攻撃が近年問題になっている．パスワードを忘れたときのために，あらかじめ設定した秘密の質問に答えることによってパスワードをリセットできるサービスがあり，「母親の旧姓は?」や「最初に飼ったペットの名前は?」のような質問に対してユーザが答を登録するようになっている．このような問題は他人が調べたり推測したりすることが容易であるうえに秘密の質問の数は一般的に少なく，パスワードよりも脆弱だといえる [15]．ユーザが作成した秘密の質問を使えばこのような問題はなくなるはずであるが，他人に解かれにくい問題をユーザが作成することは難しく，またユーザ自身が答を忘れてしまうことも多いと考えられている [10][16]．

EpisoPass では，他人には解くことが難しく自分では忘れないような秘密の質問を自由にいくつでも利用できるようになっている．問題作成に慣れていないユーザには有効な秘密の質問を作成することは難しいかもしれないが，次節で述べるように，適切な質問を選ぶことによりこの問題を解決できるはずである．

3.3 秘密の質問の選択

EpisoPass 利用において秘密の質問の選択は非常に重要である．他人が推測することが難しく，自分が決して忘れないようなエピソード記憶を秘密の質問として利用すべきであり，以下のような性質をもつ記憶は秘密の質問として利用すべきではない．

- 自慢になるもの (何かの機会にうっかり他人に話してしまう可能性がある)
- ネット上に記録が残っているもの
- 他人と情報を共有しているもの
- 趣味や嗜好に関連するもの (他人に推測されやすいうえに嗜好が変化する可能性がある)

このようなものではなく，「わざわざ人に話すことはないが自分の記憶に強く残っているような無難なエピソード記憶」を秘密の質問として利用するのが良いであろう．具体例としては以下のようなものがある．

- 昔のちょっとした怪我の場所や種類
- 昔のちょっと悔しい思い出
- 昔何かを見つけた場所

たとえば図 1 の 3 問目のような経験は他人に話したことが無いが，痛い思いをしたことは忘れないし，偽答の地名を並べるのも簡単なので，認証のための秘密の質問として適切であると考えられる．

3.4 偽答の作成方法

秘密の質問の種類によっては偽答を用意するのが難しい場合があるが，正答として人名や地名を利用する場合，正答に似た人名や地名をリストすることは難しくない．「世田谷」が正答であるとき，「目黒」「杉並」のような偽答を用意するのは簡単である．正答と同じカテゴリに属する単語を自動的にリストすることができれば正答をもとにして簡単に偽答のリストを生成することができる．ひとつの単語もしくは単語の集合と同じカテゴリに属する単語を検索する手法は「同位語検索」と呼ばれ，Web のデータを利用した様々な同位語検索システムが提案されている [17][24]．

人名や地名の偽答を作成したい場合は人名や地名のデータベースを利用して偽答を生成することができる．市町村の人口ランキングや位置関係のデータなどを利用すれば似た地名を偽答としてリストすることが可能であるし，人名ランキングを利用すれば似た苗字を偽答とすることができる．たとえば日本の名字ランキングの 40 位近辺に「長谷川」「近藤」「石井」「斉藤」「坂本」「遠藤」「藤井」などの名字があるので，「石井」が正答のときこれらの名字を偽答にすればよい．しかし，「小学生のとき〜だった同級生は誰?」という秘密の質問の正答が「石井」であるときこの方法で偽答を生成すると，「長谷川」「近藤」などの同級生の存在を確認することにより正答が「石井」であることが判明してしまう可能性があるので注意が必要である．

3.5 パスワード漏洩時の問題

秘密の質問を公開している場合，シード文字列とパスワードの対応がひとつでも漏洩してしまうと，総当たり計算でチェックすることにより，すべて秘密の質問の正答が判明してしまう．秘密の質問の正答を知っていればシード文字列からパスワードを計算することができるので，漏洩した秘密の質問は利用不可能になってしまう．3.1.1 や 3.1.2 のような運用をしている場合はパスワードがひとつ漏洩しても他のパスワードは安全だが，3.1.3 のような運用をしている場合はひとつでもパスワードが漏洩するとあらゆるパスワードが漏洩してしまうことになる．

通常の Web サービスなどのパスワードが漏洩することは考えにくいが,「自転車の鍵番号」のように家族などで共有する可能性があるものに対して 3.1.3 のような運用を行なうと，鍵番号を知っている人物が総当たり攻撃を行なうことによって秘密の質問の答が判明してしまう．シード文字列や秘密の質問を秘密情報として扱わない場合はパスワードを他人と共有しないように注意する必要がある．

3.6 画像認証の利用

忘れにくいエピソード記憶を利用する認証手法として様々な画像認証システム [2][9][19] が提案されている．複数の画像の中から正答を選択するもの (Cognometric 方式)，ひとつの画像の中の特定の場所を指定するもの (Locimetric 方式)，画像の上で描画操作を行なうもの (Drawmetric 方式) が広く利用されているが [2][9]，EpisoPass は図 1 のように文字列のかわりに画像 URL を秘密の質問として利用する点が異なっている．Cognometric 方式はエピソード記憶を効果的に利用できるが，多数の偽答画像が必要だという問題がある．Locimetric 方式はエピソード記憶を効果的に利用できないことに加え，クリックしやすい「ホットスポット」は限られているため充分なエントロピーを確保できないことが問題になる [4]．また Drawmetric 方式もエピソード記憶を効果的に利用できないし，ユーザは似た傾向のストロークを選びがちであるため充分なエントロピーを確保しにくいことが知られている [13]．EpisoPass のように画像を秘密の質問として利用する場合，通常の秘密の質問の場合と同様に偽答を増やすことが容易であることに加え，画像に関連したエピソード記憶を有効に利用できるという利点がある [22]．

図 5 は本棚.org[20] のユーザのひとり [3] が利用している画像認証問題である．このユーザ以外には正答は見当もつかないが，本人にとっては忘れることがないエピソード記憶と結びついた画像だということであった．

[3] http://hondana.org/Leiko



図 5. エピソード記憶と深く結びついた画像.

4 結論

エピソード記憶に結びついた秘密の質問を利用してパスワードを生成/管理できるシステム EpisoPass を提案した．EpisoPass は単純な原理にもとづいており柔軟な利用形態が可能であり，強力な秘密の質問を用意することにより秘密情報を全く覚えることなく安全な認証を行なうことができる．強力な秘密の質問を作成して安全に運用が可能かどうかを長期的に評価したいと考えている．

参考文献

[1] AgileBits Inc. 1Password. https://agilebits.com/onepassword.

[2] R. Biddle, S. Chiasson, and P. Van Oorschot. Graphical passwords: Learning from the first twelve years. *ACM Comput. Surv.*, 44(4):19:1–19:41, Sept. 2012.

[3] Dashlane, Inc. Dashlane. https://www.dashlane.com/.

[4] A. E. Dirik, N. Memon, and J.-C. Birget. Modeling user choice in the PassPoints graphical password scheme. In *Proceedings of the 3rd symposium on Usable privacy and security*, SOUPS '07, pp. 20–28, 2007.

[5] D. Florêncio and C. Herley. A large-scale study of web password habits. In *Proceedings of the 16th international conference on World Wide Web*, WWW '07, pp. 657–666, 2007.

[6] D. Florêncio, C. Herley, and B. Coskun. Do strong web passwords accomplish anything? In *Proceedings of the 2nd USENIX workshop on Hot topics in security*, HOTSEC'07, pp. 10:1–10:6, Berkeley, CA, USA, 2007. USENIX Association.

[7] E. Hayashi and J. Hong. A diary study of password usage in daily life. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 2627–2630, 2011.

[8] C. Herley, P. C. Oorschot, and A. S. Patrick. Passwords: If We're So Smart, Why Are We Still Using Them? In R. Dingledine and P. Golle eds., *Financial Cryptography and Data Security*, pp. 230–237. Springer-Verlag, 2009.

[9] Internet Safety Project. Graphical Passwords. http://www.internetsafetyproject.org/wiki/graphical-passwords.

[10] M. Just and D. Aspinall. Personal choice and challenge questions: a security and usability assessment. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, SOUPS '09, pp. 8:1–8:11, 2009.

[11] S. Komanduri, R. Shay, P. G. Kelley, M. L. Mazurek, L. Bauer, N. Christin, L. F. Cranor, and S. Egelman. Of passwords and people: measuring the effect of password-composition policies. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 2595–2604, 2011.

[12] LastPass.com. LastPass. https://lastpass.com/.

[13] D. Nali and J. Thorp. Analyzing user choice in graphical passwords. Technical Report TR-04-01, School of Computer Science, Carleton University, Ottawa, 2004.

[14] National Institute of Standards and Technology. Electronic Authentication Guideline. NIST Special Publication 800-63-1. http://csrc.nist.gov/publications/nistpubs/800-63-1/SP-800-63-1.pdf, 2011.

[15] A. Rabkin. Personal knowledge questions for fallback authentication: security questions in the era of Facebook. In *Proceedings of the 4th symposium on Usable privacy and security*, SOUPS '08, pp. 13–23, 2008.

[16] S. Schechter, A. J. B. Brush, and S. Egelman. It's No Secret. Measuring the Security and Reliability of Authentication via 'Secret' Questions. In *Proceedings of the 2009 30th IEEE Symposium on Security and Privacy*, SP '09, pp. 375–390, 2009.

[17] R. C. Wang and W. W. Cohen. Language-Independent Set Expansion of Named Entities Using the Web. In *Proceedings of the 2007 Seventh IEEE International Conference on Data Mining*, ICDM '07, pp. 342–350, 2007.

[18] C. Zarate. SuperGenPass. http://supergenpass.com/.

[19] 小池 英樹, 増井 俊之, 高田 哲司. 画像を用いた個人認証手法. 情報処理, 47(5):479–484, May 2006.

[20] 増井 俊之. 本棚通信: 控え目なグループコミュニケーション. インタラクション 2005 論文集, pp. 135–142, February 2005.

[21] 増井 俊之. マイ認証. *Unix Magazine*, 21(3), March 2006. http://www.pitecan.com/UnixMagazine/PDF/if0603.pdf.

[22] 増井 俊之. パスワードとの闘い - パスワードなし認証システムの運用報告. コンピュータセキュリティシンポジウム 2009 論文集, 2009.

[23] 野村総合研究所. 利用者登録する商品・サービスを選別する傾向が強まった生活者と顧客情報の鮮度維持を望む事業者〜生活者と事業者を対象とした ID に関する実態調査〜. http://www.nri.co.jp/news/2012/120208.html, February 2012.

[24] 大島 裕明, 小山 聡, 田中 克己. Web 検索エンジンのインデックスを用いた同位語とそのコンテキストの発見. 情報処理学会論文誌. データベース, 47(19):98–112, December 2006.

[25] 平野 亮, 森井 昌克. パスワード運用管理に関する考察および提案とその開発. 電子情報通信学会技術研究報告. ISEC, 情報セキュリティ, 111(285):129–134, November 2011.

附録: パスワード計算アルゴリズム

パスワードは，秘密の質問と回答の組合せにもとづいてシード文字列を換字することによって計算される．換字は文字種ごとに行なわれる．たとえばシード文字列の 1 桁目が数字のときはパスワードの 1 桁目は数字に変換され，シード文字列の 1 桁目が記号のときはパスワードの 1 桁目は記号に変換される．

数字の換字の場合、シード文字列内の数字 A は，換字関数 $f_N()$ によってパスワード内の数字 B に変換される．$f_N()$ は以下のような関数である．

$$f_N(x) = (10 + N - x) \bmod 10$$

ここで N はシード文字列と回答の組み合わせから計算される自然数で，答の選択により変化する．たとえば $N = 5$ のとき，$f_5(x) = (10+5-x) \bmod 10$ となるので，x と $f_5(x)$ の対応は以下のようになる．

$$f_5(0) = 5$$
$$f_5(1) = 4$$
$$\cdots$$
$$f_5(8) = 7$$
$$f_5(9) = 6$$

$f_N()$ は N により変化するので，N がわからなければ $f_N()$ もわからない．N はシード文字列と回答の組み合わせから計算される自然数で，秘密の質問の答とシードを知らなければ N を計算することはできない．EpisoPass では以下のようにして $f_N()$ を計算している．

1. 問題文字列と選択した答の文字列の組を連結した文字列 S を生成
2. S の MD5 ハッシュ値 M(16 進 32 桁の文字列) を計算
3. シード文字列の k 桁目の文字に対し，M の $(k-1) \times 4$ 文字目から $(k-1) \times 4 + 3$ 文字目までの部分文字列 (16 進 4 桁) を取得し，それを N とする．たとえばハッシュ値 M が 12345678... だったとき，0x1234 = 4660 なので，1 桁目を計算する換字関数は $f_{4660}(x)$ となる．

研究分野と参照関係からたどる文献可視化インタフェース

中澤 里奈　　伊藤 貴之　　斎藤 隆文*

概要. 研究を進めていく上で，論文サーベイは非常に重要である．しかし，初学者は検索する上で適切なキーワードを全て知っているわけではないため，検索エンジンを用いた方法では，検索ワードによっては見落としがおこる可能性がある．一方で，より視覚的に参照関係をたどりながら文献を調査する方法の一つとして，文献の参照関係の可視化の研究が数多く行われてきた．これらの研究では，文献をノードとして時系列に並べ，参照関係をエッジとしたグラフによって，参照関係の構造を表現している．このような手法は，研究分野をまたいだ複雑な接続関係が視認しづらく，参考文献をたどっていくにはやや不向きである．そこで本報告では，研究分野と参照関係からたどる論文探し支援のための参照関係可視化手法を提案する．提案手法では，文献の内容をもとに文献を分類し，これらの分類と参照関係にもとに，可視化手法 FRUITS Net を用いて文献の配置を決定，エッジバンドリングを用いて参照関係情報の要約可視化を行う．これにより，分野間の参照関係をより簡単に把握することで文献調査を支援できるのではないかと考える．

1　研究背景

研究を進めていく際の文献調査は非常に重要な要素である．調査方法としては，検索エンジンを用いたり，既に読んでいる論文の参考文献をたどっていったりするといった方法がある．しかし，研究を始めたばかりの学生は検索のための適切なキーワードを全て知っているわけではないため，検索エンジンによる方法では，検索ワードによって見落としがおこってしまう上に，論文同士の関係を見いだすことができない．一方で，参考文献をたどることで視覚的に文献を調査する方法の一つとして，文献の参照関係の可視化の研究が数多く行われてきた [1] [2]．これらの研究では，各文献をノードとして時系列に並べ，参照関係をエッジとしたグラフによって，参照関係の構造を表現している．このような手法では，研究分野をまたいだ複雑な参照関係をグラフとして描画すると，非常に多くのエッジ間の交差が生じて視認性に悪影響をおよぼす．そこで本報告では，研究分野と参照関係からの文献探しに適した参照関係可視化手法を提案する．本手法では，文献の内容でクラスタリングされた参照関係グラフを描画し，クラスタ間の参照が多い場合にはエッジバンドリングを施す．この時，敢えて時系列順に並べないことでクラスタ間の参照関係の太さを視認しやすくする．これにより，一つのキーワードに関係ある論文が複数の分野にわたっている時に，それらを見逃さずに網羅的に発見でき，その全体的な構図も理解しやすくなると考えられる．

Copyright is held by the author(s).

* Rina Nakazawa and Takayuki Itoh, お茶の水女子大学大学院 人間文化創成科学研究科 理学専攻, Takafumi Saito, 東京農工大学大学院 工学研究院 先端情報科学部門

2　提案手法

2.1　文献のカテゴリ分類

本手法では，文献の内容から文献を研究の分野別にカテゴリ分類するために，今回は一例として潜在的ディリクレ配分法（LDA: Latent Dirichlet Allocation）[4] を用いた．LDA は，一つの文書に対して複数のトピックの存在を想定した，文書生成過程の確率的なトピックモデルである．LDA で推定されるトピックを一つの研究分野と見なし，各文献がどのような研究分野に属するかを分類する．LDA において，トピック数は予め決めておく値であるため，今回はトピック数を 10 とした．

2.2　参照関係の要約可視化

続いて，LDA を用いて推定したカテゴリ分類をもとに，FRUITS Net[3] の配置アルゴリズムを用いることでノードの配置を決定する．これにより，参照関係がある文献同士を近く，同じ研究分野に属する文献同士を近く配置し，ノード同士の重なりを回避しつつ，画面配置占有面積を低減することができる．ノードの大きさはその文献の被引用数に比例する．その上で，エッジの視認性を向上するためにエッジバンドリングを施すことで，大量の線情報を束ね，要約する．エッジを束ねる本数はユーザが閾値をスライダによって設定することができる．カテゴリ分類によってできた文献クラスタ間のエッジ本数を計算し，それが閾値以上であるならば，本数に比例させた太線を束として描く．閾値に満たない場合はエッジをそのまま描画する．

2.3　ユーザインタフェース上での描画と操作

図 1 は，本ユーザインタフェースにおける表示画面例である．ユーザは操作パネル上の (1)(3) で

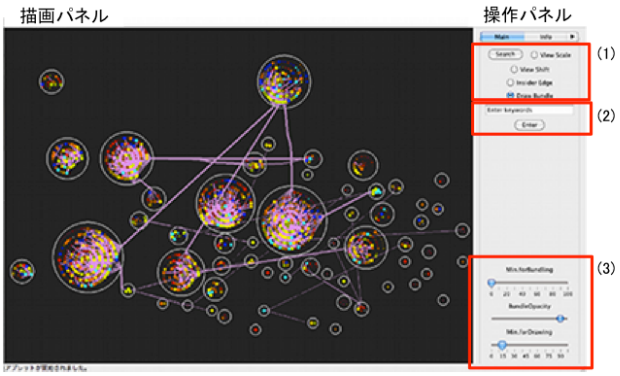


図 1. インタフェース表示画面例

可視化結果の拡大縮小や平行移動操作，エッジバンドリング適用の有無，その閾値の設定を行うことができる．また，各文献であるノードをクリックすると，その文献のタイトルや発行年，アブストラクトなどの詳細情報を表示する．同時に，クリックされたノードが持つエッジと，それに接続するノードが持つエッジの二世代分の参照関係をハイライト表示する．このエッジのハイライトは 2 つのノードについて同時に適用可能であり，これによって両ノードの接続関係を比較することができる．一方，最初に注目すべき文献を探すことは初学者にとっては容易ではない．そこでユーザは大まかな研究分野かキーワードを利用することで，画面上に表示させる文献をフィルタリングすることができる．自分の興味のある研究分野カテゴリを操作パネル上のボタンから選択した場合，選択した研究分野のみと分類された文献群をウィンドウ中央部に拡大表示される．また，(2) にキーワードを入力すると，入力された語がタイトル中に含まれている文献ノードのみが表示される．

3 実行結果

使用した文献データは 1990 から 94 年，2000 から 2010 年までに SIGGRAPH で発表された論文 1072 本である．例として，ユーザが写真や動画の処理関連の研究の文献を調査したい場合を考える．写真，動画関連の研究分野カテゴリを選択すると，表示結果は図 2 のようになる．図 2 の (A) が写真，動画の処理関連の分野であり，(A) に接続する束を追っていくと，(A) には (B) と (C) のどちらかと参照関係を持つ文献が多く存在するということがわかる．(B) は画像処理関連のクラスタ，(C) は大域照明，光反射や光拡散モデルなどと写真や動画の処理が組合わさったクラスタである．この結果から，写真や動画の処理に関する文献を探したい場合，優先して探したい内容がどちらかというと (B) の画像処理技術に関連寄りなのか (C) のライティング関連寄りなのかによって，その分野と接続している文献から先に読み進めていけばいいということがわかる．従って，本インタフェースを利用することで関係ありそうな論文を研究分野と参照関係から整理して，どの論文群から先に読むかを選択できるのではないかと考えられる．

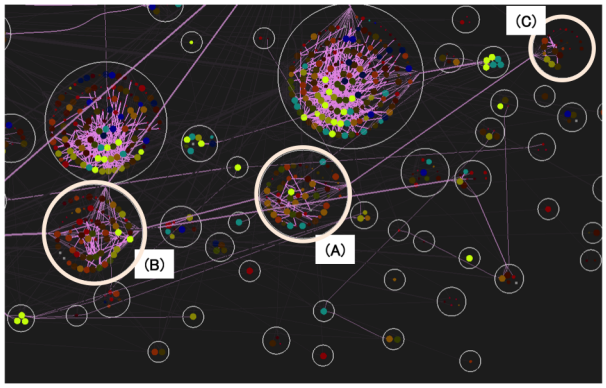


図 2. カテゴリ選択時の表示結果

4 まとめと今後の課題

本報告では分野間の繋がりを考慮した文献参照関係の可視化インタフェースを提案した．本手法によって，研究分野と参照関係から関連のありそうな文献を整理し，優先的に読むべき文献を選ぶことができるようになるのではないかと考えられる．現在はエッジの方向性を考慮した描画を行っていないため，今後はエッジの方向性を表現したエッジバンドリングと可視化を行っていく．さらに，今回手動で決定していたトピック数の推定や GUI 機能の拡充，評価実験なども進めていきたい．

参考文献

[1] J. Matejka, T. Grossman, and G. Fitzmaurice. Citeology: visualizing paper genealogy. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pp. 181-190, 2012.

[2] 岡田, 渡部, 斎藤. 時系列に基づいた文献参照関係の可視化-文献のクラスタリング-. 情報処理学会第 74 回全国大会, 3ZB-3, 2012.

[3] T. Itoh, C. Muelder, K. -L. Ma, and J. Sese. A Hybrid Space-Filling and Force-Directed Layout Method for Visualizing Multiple-Category Graphs. In *Proceedings of IEEE Pacific Visualization Symposium*, pp. 121-128, 2009.

[4] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent Dirichlet Allocation, Journal of Machine Learning Research, Vol. 3, pp. 993-1022, 2003.

偶然の再会を演出する消極的な人向け人間関係継続支援システム

芥川 洋平　　山本 景子　　倉本 到　　辻野 嘉宏 *

概要. 周囲の人と積極的に交流することが苦手な人付き合いに消極的な人は，日常生活において席を共にした初対面の人に連絡先を聞くことが難しく，そのときに新たに構築した人間関係を維持できない．この場合，その人との偶然の再会が人間関係を再開・継続するきっかけになることもありうるが，そのようなことはめったにないという問題がある．また，消極的な人はたとえ偶然の再会をすることがあっても，相手についての記憶の曖昧さ・人違いの危惧などから声をかけられず，結局その機会を有効に使えないという問題もある．そこで本研究では，以前に席を共にしたものの，そのときに連絡先を聞くことができず，それ以降にも連絡をとることができない人との関係を再開・継続させることを目的とする．これを実現するために，偶然の再会を消極的な人のために演出する手法として，以前に席を共にした消極的な人同士が同じ場所に向かうよう促す情報を提供し，近くに居合わせたときに相手の存在に気づける通知を行うシステムを提案する．

1　はじめに

私たちは日常生活において，同じ目的を持った人同士が同じ場所に居合わせたり，友人の友人などの繋がりから交友関係が広がったりすることによって，初対面の人と席を共にすることがある．例として懇親会や合コンなどが挙げられる．その際に，初対面の人と一緒に会話したりお酒を飲み交わしたりすることで，新たな人間関係を構築することができる．

この新たに構築した人間関係を維持する手段として，その人の連絡先を聞き連絡をとりあうことが挙げられる．しかし，周囲の人と積極的に交流することが苦手な人付き合いに消極的な人（以降,「消極的な人」）は，初対面の相手に対して連絡先を聞くことができない．また，消極的な人は，連絡先を知っている共通の知人や幹事を経由することで，後日に連絡先を聞くこともできない．とりわけこのような状況は，連絡先を聞きたい相手が異性の場合に多く起こりうる．このように，消極的な人が初対面の人の連絡先を聞くことができないという問題（以降,「問題点 1」）が起こる理由として，

- 関係の浅い相手に対して自分の連絡先を教えることに抵抗を感じ，同様に相手の連絡先を聞くことに抵抗を感じる
- 場の盛り上がりなどを考慮して，初対面の人の連絡先を聞くタイミングを逃してしまう
- 後日に連絡先を聞こうとしても，共通の知人や幹事に誰の連絡先を知りたいかという自身の興味を教えることに恥ずかしさを感じる

ということが考えられる．

Copyright is held by the author(s).

* Youhei Akutagawa, Keiko Yamamoto, Itaru Kuramoto and Yoshihiro Tsujino，京都工芸繊維大学

そこで本研究では，消極的な人が，以前に席を共にしたものの，そのときに連絡先を聞くことができず，それ以降にも連絡をとることができない人との関係を再開・継続することを目的とする．

一方，街を歩いているときに，以前に席を共にした初対面の人を偶然見かけることがある．このとき，その人に声をかけて再会することで，刹那的になっていた人間関係を再開・継続するきっかけを作ることができる．しかし，街を歩いているときに以前に席を共にした初対面の人を偶然見かけることは頻繁に起こる事象ではない．

また，以前に席を共にしたものの，連絡を取らなかった人を偶然見かけるまでの間にその人の顔や容姿を忘れてしまっていたり，曖昧になっていたりすることがある．その結果，消極的な人は両者がすぐ近くに居るにも関わらず，その人に声をかけずにそのまますれ違ってしまい，人間関係を再開･継続する機会を逃してしまうことになる．このように，消極的な人が偶然の再会を逃してしまうという問題（以降,「問題点 2」）が起こる理由として，

- 相手についての記憶の曖昧さから確信が持てず，別人だった場合を危惧する
- 自分が相手を覚えていても相手が自分を覚えていないかもしれないと懸念する

ということが考えられる．

以上より本研究では，消極的な人が相手の連絡先を聞けなくても，人間関係を再開・継続するきっかけになる偶然の再会を，消極的な人のために演出する．その手法として，以前に席を共にした消極的な人同士が同じ場所に向かうよう促す情報を提供し，近くに居合わせたときに相手の存在に気づける通知を行うシステムを提案する．

西田ら [1] や閑野ら [2] も初対面の人あるいは関係の浅い人との交流促進支援システムを提案してい



a) 店舗情報通知時　　b) 再会通知時

図 1. 提案システムの通知画面表示例

るが，西田らは，交流の場から離れた後も人間関係を継続させることについて，閑野らは，周囲の人と積極的に交流することが苦手な消極的な人がシステムを使うことについては考慮していない．

2　提案システム

以前に席を共にした消極的な人同士に，人間関係を継続するきっかけになる偶然の再会を提供するために，図 1.a) のように彼らに同じ店舗や施設の情報を表示し，同じ場所に行きやすくする．これにより，ユーザは人間関係を継続するきっかけになる偶然の再会をより多くすることができると考えられる．また，このように同じ場所に行くかもしれないという可能性の情報を通知することで，もしかしたら相手に会えるかもしれないという緊張感を持たせることができると考えられる．これにより，ユーザを吊り橋理論 [3] のような状態にすることができ，特に異性の場合において，より深い交友関係に繋げることができると考えられる．

これを問題点 1 を解消しつつ実現するために，ユーザが直接相手の連絡先を聞かなくても，システムが以前に席を共にしていた人を識別し，ユーザに偶然の再会の機会を与えるようにする．まず，以前に席を共にしていた人を識別する手法として，席を共にしていたユーザの携帯端末に共通の ID（以降,「イベント ID」）を持たせ，この ID を用いて席を共にしていたユーザかどうかを判別する．そして，イベントの開始時に，いつ・どこのイベントに参加したかという情報とイベント ID を各ユーザの携帯端末で共有させる．

また，問題点 2 を解消するために，以前に席を共にしていた人同士がその後再び近くに居合わせたときに，システムが図 1.b) のような画面表示と音や振動による通知を行うようにする．このように，いつ・どこのイベントで出会った人がいるかの通知を行うことによって，ユーザが周囲を見回して同じイベントに参加していた人を探すきっかけを与えたり，ユーザが容姿を忘れてしまったり曖昧になったりしていた相手について確信を持ったりすることができると考えられる．さらに，システムの通知に気づいて同じように携帯端末を持って周囲を見回している人がいることに気づき，その人に注目することで，相手を特定しやすくなると考えられる．各ユーザの距離を測定する手法として，各ユーザが所有している携帯端末の位置情報を取得し，取得した位置情報をサーバに送信し携帯端末間の距離を算出する．

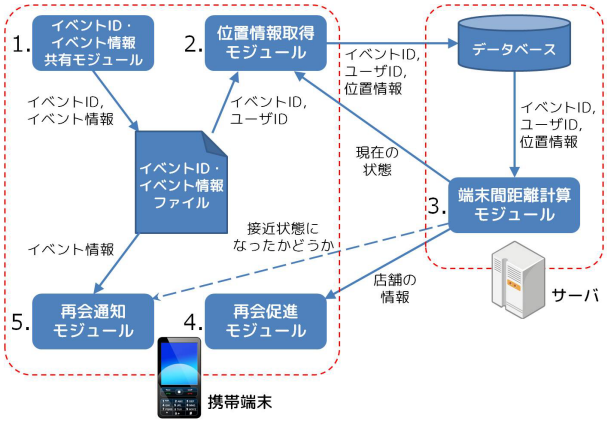


図 2. モジュール間の関係

提案システムは携帯端末上およびサーバ上に実装された以下の 5 つのモジュールで構成される．

1. イベント ID・イベント情報共有モジュール
2. 位置情報取得モジュール
3. 端末間距離計算モジュール
4. 再会促進モジュール
5. 再会通知モジュール

また，モジュール間の関係は図 2 のようになる．

3　今後の課題

再会したときの交流のさらなる支援のために，提案システムが近くに以前席を共にした人が居ることを通知するだけでなく，再会したときに話す内容に対する支援をすることが挙げられる．

参考文献

[1] 西田 健志, 濱崎 雅弘, 栗原 一貴. 超消極的な人でも安心して使える学会での交流促進システム. WISS2012, pp.103-108, 2012.

[2] 閑野 伊織, 田中 二郎. イベント開催前から開催後まで一連の流れに沿ってコミュニケーションを支援するシステム. マルチメディア, 分散, 協調とモバイル (DICOMO2013) シンポジウム, pp.56-63, 2013.

[3] D.Dutton and A.Aron. Some Evidence for Heightened Sexual Attraction under Conditions of High Anxiety. Journal of Personality and Social Psychology, Vol.30, No.4, pp.510-517, 1974.

ThermoGuide：表面温度分布を利用した加熱管理支援手法

斉藤 清隆　　高橋 伸　　田中 二郎 *

概要. 本稿では，サーモグラフィを利用し調理中のフライパンや食材の表面温度を分析することで，加熱調理を支援する手法を提案する．非接触で温度情報を取得可能なサーモグラフィを利用することで，調理者の作業を妨害することなく加熱状況の判別や材料投入の検知が実現可能となる．また，提案手法の適応例として卵焼き作成ナビゲーションシステムの設計及び実装を行った．これによって従来の曖昧な表現が用いられていたレシピとは異なる，定量的な判断基準に基づくナビゲーションの可能性を示した．

1　はじめに

加熱調理において温度管理は，料理の出来具合を左右する重要なファクタである．しかしながら，一般的な料理で多く参考にされているレシピ文では，曖昧な表現がなされていることが多く，調理初心者が適切な温度管理を行う上での障害となっている．一方で，正確な温度管理が必要な料理に関しては温度計を用いて調理を行うが，温度計が刺せる料理に限定される上，素早い作業が必要な調理では測定している暇がないという欠点もある．食材の温度分布の変化は一様ではなく，熱源からの距離や食材の種類によってもまちまちとなるため，1 点の温度では知り得る情報も限られてくる．

そこで我々は，非接触で温度分布を取得可能なサーモグラフィを利用し，調理中の加熱管理を支援する手法を提案する．既存の研究としてフライパンの中心温度に基づいたナビゲーションシステム [1] があるが，本研究では調理面の温度を監視することで，全体の加熱具合の判別から局所的な材料の火の通り具合まで幅広い加熱のサポートを目指す．

また，提案手法の適用例として卵焼き作成ナビゲーションシステムの設計および実装を行い，システムに関する実験を行った．

2　表面温度分布を利用した加熱管理支援

加熱調理をどれだけの時間行うべきかは，火加減，食材の量，調理器具の種類などの条件によって変化する [2]．そこで，加熱状況と密接に関係のある表面温度分布を取得し，調理状況を認識する．

具体的には，表面温度分布から計算できる適当な特徴量を，加熱を十分に行ったかどうかの指標として用いる．例えば，フライパンの余熱が完了したかどうかは，適切な温度になった領域が十分に広がっているかどうかで判断できる．また，材料の投入時には表面温度は一気に減少するため，高温領域の減少によって判断可能となる．このように，調理の指標としてフライパン面積に対する，ある温度以上/以下の領域の割合を利用する．また，単純な指標として調理領域の最低温度や最高温度も併せて利用する．

これらの手法を用いることで，フライパン自体の加熱や材料に火が通ったかどうかの判別あるいは加熱をし過ぎといった温度上昇に関する条件設定ができると同時に，食材が入ったタイミングの認識や加熱した食材を冷ます場合など温度低下に関する条件としても利用することが可能となる．これらを用いることで，従来曖昧な表現が用いられていたレシピ文において絶対的な指標を設けることが可能となる．

3　試作 : 卵焼き作成ナビゲーションシステム

本研究で提案する加熱管理手法の適用例の一つとして卵焼き作成ナビゲーションシステムを作成した．卵は固まり方が温度に大きく依存し，温度の調整ミスによる失敗が特に生じやすい．また，卵全体が固まるまでしっかり加熱できたか，部分的に加熱がされすぎていないか，といった状況は 1 点の温度だけでは判別が難しい．そのため，本研究で提案する表面温度分布を用いた状況推定が活かされると考えた．

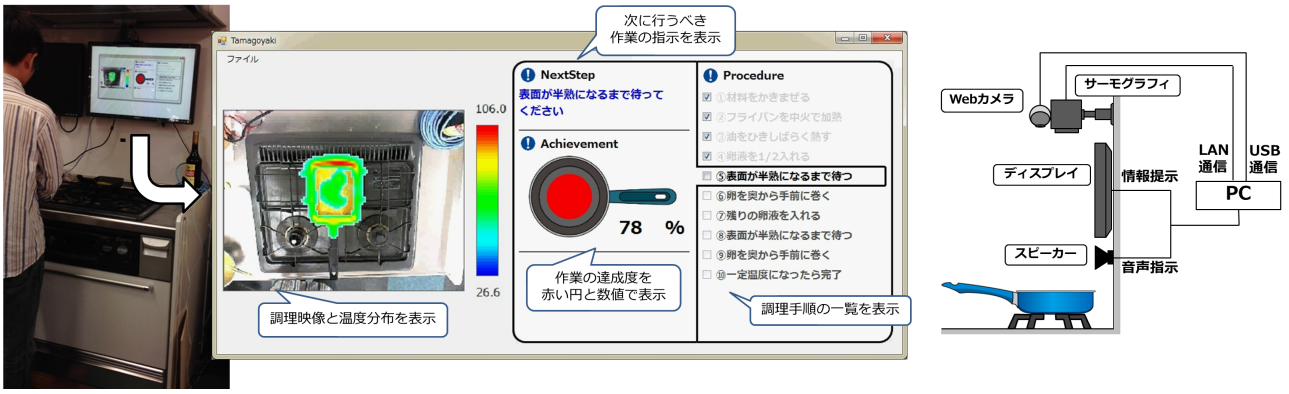
3.1　システム概要

本システムでは，壁面ディスプレイに表示される GUI およびスピーカーから流れる音声によってユーザをナビゲートする．システム利用時のキッチン内の様子およびナビゲーション GUI を図 1(a) に示す．

GUI 左部には調理時の映像とそれに対応する温度分布が重畳表示されている．ここでは目標温度に達している部分のみ熱画像表示を行っており，加熱が足りない部分はそのまま調理映像を表示している．そのため，どの箇所の加熱が足りないかをひと目で確認することが可能となる．また，システム側では逐次調理状況を監視しており，加熱状況に応じて達成度がイラストと数値で示される．そして，達成度

Copyright is held by the author(s).

* Kiyotaka Saito, 筑波大学大学院 システム情報工学研究科 コンピュータサイエンス専攻, Shin Takahashi and Jiro Tanaka, 筑波大学 システム情報系



(a) ナビゲーション GUI　　(b) システム構成図

図 1. ナビゲーションシステム

が 100 %になると自動でレシピがチェックされ，次の指示に切り替わると同時に音声によって指示文が読み上げられる．そのため，ユーザはシステムに対して入力等を行う必要はなく，映像や達成度といった情報を参考にしながら，自動で提示されていくナビゲーションに従って調理を行えばよい．

3.2 システム構成

本研究では温度取得のため，チノー社製熱画像センサ TP-H0225AN を使用した．このセンサは物体の赤外線放射エネルギーを検出することで非接触の温度検知を実現する．

システム構成図を図 1(b) に示す．キッチンの壁面にディスプレイとスピーカーが設置されており，その上部にサーモグラフィと web カメラが下向きで固定されている．web カメラから取得された映像は USB 経由で，サーモグラフィから取得された情報は LAN 経由でそれぞれコンピュータに送信される．

4 評価実験

4.1 実験内容

ナビゲーションシステムを用いて調理初心者の被験者に卵焼きの作成を行ってもらった．被験者は調理初心者の大学生・大学院生 4 名（男性 3 名，女性 1 名）であり，いずれの被験者も卵焼きの調理は今回が初めてであった．

4.2 結果と考察

実験で被験者が作成した卵焼きの一覧を図 2 に示す．いずれの卵焼きも，表面が焦げていたり卵が固まりきっていないといったことはなく，おおむね調理を成功することができていた．このことより，今回初めて卵焼きに挑戦した被験者においても本研究で提案する表面温度分布に基づくナビゲーションにより適切な調理手順を導くことができたといえる．



図 2. 実験 調理結果

システムの使用感についてだが，システムが調理状況を認識し進み具合を自動でチェックする機能は非常に好評であった．また，達成度表示により現在の作業がいつ終了するかをあらかじめ予測できたため，素早く次の作業に移ることができていたことも利点として考えられる．しかし，調理中に卵の巻き方など動作の面において被験者が苦戦する様子が見られた．今後は文字による指示だけではなく，動作に関するナビゲーションが必要と考えられる．

5 まとめと今後の発展

本研究では，非接触で温度分布を取得可能なサーモグラフィを利用し調理中の加熱管理を支援する手法を提案した．また，提案手法の適用例として卵焼き作成ナビゲーションシステムの設計及び実装を行い，実験を通じて従来のレシピ文に代わるような定量的な基準に基づくナビゲーションの可能性を示した．

今後の発展として，卵焼き以外の適用例を実現し，様々な種類の料理においても提案手法が有効であるかについて検証を行っていく予定である．

参考文献

[1] D. Uriu, M. Namai, S. Tokuhisa, R. Kashiwagi, M. Inami, and N. Okude. panavi: Recipe Medium with a Sensors-Embedded Pan for Domestic Users to Master Professional Culinary Arts. In *Proc. of CHI'12*, pp. 129–138, 2012.

[2] 渋川 祥子. 調理における加熱の基本. 日本食生活学会誌, 17(2):89–93, 2006.

Harry: ペーパークラフトからの 3 次元モデル生成

三浦 元喜* 山本 将史†

概要. 一般に，3D のモデリングを行うためにはマウスやキーボードを駆使した操作が求められるため敷居が高い．そこで我々はペーパークラフト（紙工作）から 3D モデルを手軽に構築する手法 Harry を提案する．体験者がペーパークラフトを作成する際，紙から形状を切り取る前に，形状の輪郭線をディジタルペンで入力する．その後，紙を組み合わせたり接着した箇所をペンでなぞることによって，形状同士の接続点や折れ線を計算機に入力する．入力された輪郭線や接続点を考慮した物理演算を行うことにより，3 次元形状を構成する．システムを構築し，ラフな 3D モデルが手軽に作成できることを確認した．

1 はじめに

多くの人にとって，紙は最も身近で扱いやすい記録・表現媒体である．装飾や加工が容易であることから，幼児期から塗り絵や折り紙，切り紙や工作等の立体物を含む創作活動に用いられることが多い．また，紙での創作活動を拡張するための計算機支援研究やシステムも提案されている．例えば colAR App[2] は塗り絵をタブレットのカメラで取り込むと，3D のコンピュータグラフィックスに変換して，アニメーション表示するアプリケーションである．また三谷は 2 次元バーコードが両面に印刷された紙を用いた，紙の折り畳み構造を推定・認識する手法を構築している [3].

一般に，3DCG を製作するためのモデリングソフトや CAD システムは高機能である反面，操作が複雑であるため，マウスやキーボードを駆使した操作が必要である．このため，PC 操作に慣れていないユーザにとって，3D のモデルを自在に作成することは困難である．我々はより多くの人が 3DCG 製作に関わりをもてるようにするために，紙工作を作成する過程においてディジタルペンによる筆記情報を追加することで，3D モデルを生成する手法 Harry を提案する．

2 提案手法 Harry

我々が提案する 3D モデリング手法 Harry は，ペーパークラフトを実際に作成する過程において，紙のパーツの輪郭やパーツの接続点をディジタルペンによって指定することにより，パーツの空間的な関係を記述し，3D モデルを生成するものである．すなわち，紙に対して閉曲線を記述すれば，計算機上には同じような形をした閉曲面を出力し，紙の上の任意の二点を指定すれば，計算機上ではその 2 点をつないだ立体モデルを出力する．このように，ディジタルペンをペーパーカッタやステープラに見立てた操作を行うと，計算機上に紙を模した平面あるいは曲面のポリゴンモデルを出力する．平面または曲面の形状は，現実世界の紙の挙動を模した物理演算によって推定する．コンピュータ上に出力する面は，紙のような，曲げに対しての抵抗（ハリ）がある材質を想定している．そのため，平面上の 2 点をつなぐといった操作のみで，立体的なモデルを生成することができる．図 1 に，提案手法の概略図を示す．

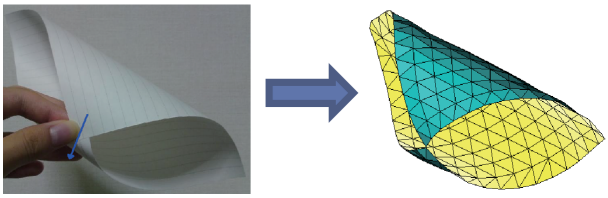


図 1. 提案手法の概略：ディジタルペンで紙の切取り形状と接続点を指定すると，対応する 3D モデルを生成

2.1 作成手順

モデル作成者は，最初にアノト方式のディジタルペンで紙にパーツの輪郭を自由な閉曲線で記入する．すると図 2 左に示すような平面モデルが画面に表示される．次にモデル作成者がペンで紙上の 2 点をタップ操作で指定するか，紙を重ねた状態で接続部分をなぞる操作 (スキャン) を行うと，図 2 右に示すように指定部分を接続した 3D モデルが画面に表示される．なお画面の平面モデルに対してマウスで接続点を追加したり，削除することも可能である．

2.2 提案手法の利点と欠点

紙に対してペンで記述するという作業は単純であり，かつ 3D モデリングに関する複雑な知識を必要としない．そのため，コンピュータに不慣れな人も

Copyright is held by the author(s).

* 九州工業大学 基礎科学研究系
† 九州工業大学 工学府 先端機能システム工学専攻

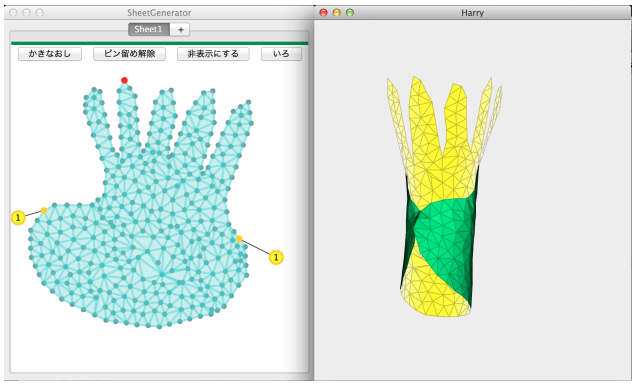


図 2. 平面モデル (左) と，立体モデル (右)

含めた多くの人が馴染みのある容易な作業であると思われる．このことから，この手法では幅広いユーザが，比較的容易にモデルを制作できるものと思われる．

また紙とペンを使用することによる利点もある．ペンデバイスは，マウスと比べると人間が正確に操作することができるものである．このことから，閉曲面の形状を指定する操作において，ユーザはより正確に，不快感なくイメージした形状を表現できると考えられる．特に，複雑な凹凸を持つような閉曲面の形状を入力する際に，ディジタルペンはマウスよりも容易に入力が行えるという利点がある．記述対象が紙であり，破棄・交換が容易に可能であるため，実際に曲げる・折るといった試行錯誤が容易に行えると言った長所もある．さらに，紙という実物体を操作する感覚を伴うため，タンジブルかつ直感的なメディアという優位性もある．

その反面，提案手法では，折り方の詳細（山折り，谷折り）や曲率の設定といった，複雑な立体形状を制作するには不向きである．おおまかな形状を作成する段階では紙とペンを用い，細かな調整についてはマウスを用い，画面の変化を観察しながら操作する方法が適切であると考えている．

3 関連研究

三谷 [3] は，折りたたまれた紙を撮影したディジタル画像から折りたたみ構造を認識し，計算機内にモデル化する手法を提案している．2 次元バーコードを表裏に格子状に複数印刷した紙を用い，折る操作を行うたびに撮影することによって，紙の構造を認識している．紙の状態を逐次記録し，3D モデル構成のための入力に用いるという点は類似性がある．ただし我々の手法は紙が重力で折れ曲がるといった，曲面を含むような形状についても構成できるという点が異なる．

五十嵐らの Teddy[1] は，初心者でも容易に 3D モデルを作成することができる手軽な 3D モデリングツールである．画面に対して輪郭形状を描くことで，計算機が適切な 3D モデルを推測し出力する．3D モデルの出力はリアルタイムに行われ，視点の変更や，追記を行うことができ，対話的にモデリングが行えるツールとなっている．提案手法は実在する「紙」に触れ，操作しながらモデリングを行えるという点が異なる．

入力デバイスを改良することによって，直感的なモデリングを可能にする研究として，脇田らが提案した pSurface[4] を挙げることができる．pSurface は，自由曲面のモデリングを行うことができる布製デバイスである．実際に手で変形させることで，その形状を入力することができる．また，マウス操作等によってモデルに行った修正を，実物体上にフィードバックすることもできる．pSurface も実体の操作により 3D モデルを操作することができる点は類似性があるが，センサやアクチュエータの制約により，作成可能な形状は本提案手法に比べると限定される．

4 まとめと今後の課題

紙の特性とディジタルペンを利用した 3D モデリング手法を提案した．紙とペンという慣れ親しんだ要素を用いることにより，これまでよりも幅広い方々に 3D モデリング環境を提供する基盤を提供できる．システムを構築し，ラフな 3D モデルが手軽に作成できることを確認した．本研究で述べた内容は，紙とペンの特性を利用した簡便な入力手法の 1 つの適用例であるが，汎用性が高いため他の創作支援や学習支援にも適用可能であると考えている．

謝辞

本研究の一部は JSPS 科研費（課題番号 23680078）の支援によるものです．

参考文献

[1] T. Igarashi, S. Matsuoka, and H. Tanaka. Teddy: a sketching interface for 3D freeform design. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques (SIGGRAPH '99)*, pp. 409–416, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.

[2] Puteko Limited. colAR App. http://colarapp.com/ (2013 年 10 月 18 日確認).

[3] 三谷 純. 2 次元バーコードを用いた紙の折りたたみ構造の認識とそのモデル化. 情報処理学会論文誌, 48(8):2859–2867, Aug. 2007.

[4] 脇田 玲, 上野 道彦, 中野 亜希人. pSurface: 自由曲面モデリングとアニメーションのための布製入出力デバイス. 情報処理学会 インタラクション 2011, Feb. 2011.

話者による正確なポインティングのできるプレゼンテーションシステム

若松 翔　　梅村 恭司　　三輪 多恵子　　岡部 正幸*

概要.　プレゼンテーションにおいてスライドのある部分に聴者の注目を集める方法として, ポインティングデバイスの利用やアニメーションの使用が挙げられる. 話している箇所の強調や情報の順序や流れを表現できるが, 使用するタイミングや指している箇所を誤ると聴者は話の流れが掴めなかったり, その後の説明についていけなくなる危険性がある. 本研究では, 話者自身がポインティングデバイスとなり発表を行うことができるプレゼンテーションツールにおいて, 話者が指定した位置に正確にマーキングする方法を示す. 話者の動作と連動しスライド上の表示ポイントにマーカが配置できるため, 明示的に注目箇所を示すことが可能である. さらに, 評価実験を行い一定の評価を得たので報告する.

1　はじめに

近年, 学会や勉強会のプレゼンテーションが録画され, ネットワークに配信されることが盛んに行われている. このような配信において, 会場の発表風景を直接撮影したものを配信する場合, 動画のビットレートによってはスライド及びポインティングの明瞭性が失われてしまう. また, 話者のコンピュータ出力のスライドと話者の映像を組み合わせたものを配信するという手法もあるが, ポインティングデバイスが配信上では反映されず, 話者が現在どの場所を指しているかが分からないという問題がある. この問題に対し, 我々は話者のシルエットをスライドの背景として表示し話者自身がポインティングデバイスとなり発表ができるシステムの開発を行った [3].

本稿では以前のシステムに欠けていた表示タイミングとインパクトに重点を置き, プレゼンテーションにおける分かりやすさを目的としたスライド位置マーキングの実装とその評価について述べる.

2　システムの概要

本システムは話者映像をシルエットに加工しスライドに重ねて表示することにより, スライドを主役としつつ話者の存在感と動作を補助として伝えるというものである. 話者がスライド上のどの部分を説明しているかを明確にするために, センサからの距離応じてシルエットの一部分の色付けも行っている. また, センサに向かって手を押し出すことによって, その部分に円形のスタンプイメージの設置が可能である (図 1). これは一定時間スライドの一部に注目を集める際に有効である.

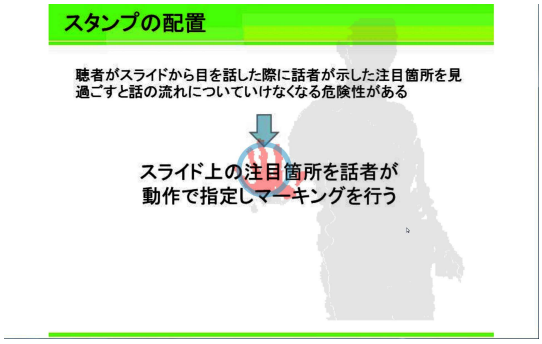


図 1. 話者シルエットを背景としたスライド表示と動作によるスタンプの配置

3　スライド位置マーキング

スライド作成時にあらかじめスタンプ表示ポイントを設定しておき, その部分に話者のシルエットが触れることによりスタンプイメージが表示される (図 2). アニメーション機能と同様にスライドに組み込むという前準備が必要であるが, 発表時に話者の指定箇所に正確にスタンプを配置することが出来る. センサに向かってプッシュ動作することによりスタンプを設置する動作マーキングに比べ, スタンプ表示ポイントを中心に正確にスタンプを配置でき, かつ動作マーキングにおけるスタンプ設置ミスがなくなるため, 任意のタイミングで確実に注目箇所を示すことができる. スタンプ表示ポイントは一度配置したスタンプがいつまでも残っているとスライド自体の妨げとなるため, スタンプは徐々に透明になっていき最終的には画面上から消えるといった挙動を行う. また, 通常のアニメーションによるコンテンツの動作は決まった順序でしか行われないが, 話者のインタラクションによりコンテンツが配置されるため, 状況によって変更が可能である.

Copyright is held by the author(s).

* Wakamatsu Sho and Umemura Kyoji, 豊橋技術科学大学 情報・知能工学系, Miwa Taeko, 豊橋創造大学 経済学部経営学科, Okabe Masayuki, 豊橋技術科学大学 情報メディア基板センター

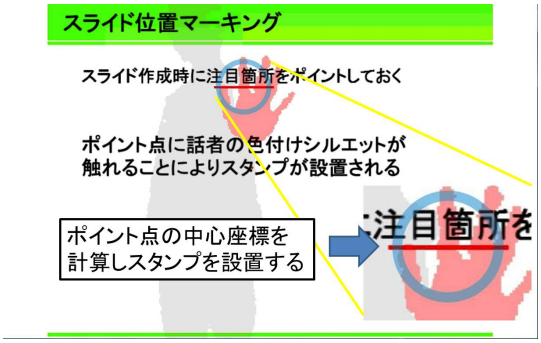


図 2. スライド位置によるスタンプの配置

4 評価

実装したスライド位置スタンプ機能が本システムに対して有効であるかの評価実験を行った.

実験手法は, 話者が提案システムを用いて聴衆にプレゼンテーションを行い, それぞれの機能を on/off して, 差があるかを評価用紙に記入してもらうという形式である. 実験は, 話者が前方に立ち聴衆に向けて説明するという一般的なプレゼンテーションの形式を取った. 話者はある程度システムに馴れているものであり, スライド内容は聴者にとって知識がある分野で, スライドは文字を中心としたシンプルな構成のもの, また下記の条件ごとに共通の話題ではあるが内容は異なっている, プレゼンテーション時間は条件毎に 10 分程度である, という条件のもと実験を行った.

プレゼン方法は以下の 3 パターンである.

- A:スライド + 話者のシルエット + 一部分の色付け
- B:スライド + 話者のシルエット + 一部分の色付け + 話者動作マーカ
- C:スライド + 話者のシルエット + 一部分の色付け + スライド位置マーカ

以上の条件を被験者に対して C, B, A, C(再度) という順番でプレゼンを行い, B の終わり, A の終わり, C の終わり (再度) に直ちに伝わりやすい条件はどちらかという設問に対し, どちらが伝わりやすいかを回答してもらった. 結果の分析には片側検定の符号検定を使用する.

4.1 結果と考察

評価結果を表 1 に示す. 表 1 よりマーカなしの場合に比べ, スライド位置マーカが加わったものの方が伝わりやすさにおいて統計的に有意であることが分かった. また, スライド作成時にマーカの出現ポイントを設置しておくスライド位置マーカと, 話者のプッシュ動作によりスタンプを配置することが出来る話者動作マーカの比較においても スライド位置マーカのほうが伝わりやすい結果となった. これは, 話者シルエットを用いる本システムでは, シルエットによるインパクトが充分あるため, 注目箇所に確実にマーキングを行うことができるスライド位置マーキングのほうが有効であると考えられる.

表 1. 各マーキング機能の評価結果 (被験者: 38 人)

	伝わりやすい	伝わりにくい	p 値
A と比べて C は	31	7	5.808 × 10e-5
B と比べて C は	25	13	0.03647

A:話者のシルエット + 一部分の色付け (マーカなし)
B:話者のシルエット + 一部分の色付け + 話者動作マーカ
C:話者のシルエット + 一部分の色付け + スライド位置マーカ

5 関連研究

本研究と同様に, 作業空間と人物画像の融合を行ったシステムとしては ClearBoard[1] がある. 本システムでは, 協同作業ではなくプレゼンテーションの問題設定なので, 話者画像の情報をシルエットに縮退する設計でシステムを作った. このため, 相手の視線, 表情のような人が見えていることによる伝達性を活かした ClearBoard とは問題設定が異なる. 本システムは話者の動作がスクリーンに投影されるため, 聴講者の視線の移動の量が少なくなる. これは, スクリーンの前でプレゼンテーションを行うことを支援するシステムである MAEDE[2] と同様な効果が得られることが期待できる. 本研究では, 話者が直接示すことができないような大型のスクリーンの利用を考慮しているため, その点で異なる.

6 まとめ

本研究では, ネットワーク配信での利用を想定した話者の映像をシルエットに変換してスライドの背景として表示するプレゼンテーションツールにおいて, シルエットが触れることによりスライドの注目箇所をマーキングする方法を示した. また, 評価実験によりマーカを用いた方が有効であり, 本システムに対しての実装が妥当であることが分かった.

参考文献

[1] 小林稔, 石井裕, ClearBoard-2 における協同作業空間と会話空間のシームレスな融合, 情報処理学会情報処理学会研究報告, 1993, pp.43-50.

[2] 前田晴己, 黒澤祐也, 栗原一貴, 宮下芳明, MAEDE: スクリーン前でのプレゼンテーションスタイル, 日本ソフトウェア科学会研究会 WISS2011, 2011, pp.164-166.

[3] 若松翔, 梅村恭司, 三輪多恵子, 岡部正幸, 話者動作による確実なポインティングのできるプレゼンテーションシステム, 情報処理学会 インタラクション 2013, 2013, pp.750-755.

「へやくる！」だらだらと賃貸物件情報をみることについて

大坪 五郎*

概要. 従来の情報検索インタフェースは,「ユーザは詳細な条件を入力しそれに対する正確な回答を期待する」という前提に基づくものが多かった．しかしながら，日常的に行う検索においてはユーザがそもそも何を入力すれば良いのかのイメージを持っていない場合も多い．不動産情報の検索はそのような性質を持っているという想定の元，まずユーザに情報を提示し，そこから気が向くままに条件を加え，変更，削除してもらうことを目的とした iPad アプリ「へやくる！」を開発した．既存の不動産情報検索アプリと比較し，一つの条件を設定するのに必要な手順が劇的に削減されており，ユーザは不動産情報について学びつつ様々な探索を試みることができる．

1 はじめに

従来の情報検索インタフェースは，ある程度自分が何を探そうとしているかユーザが把握していることを前提とし，設定された詳細な条件に対して正確な結果を返すことを目標としたものが多かった．しかしながら，問題領域によってはユーザが自分が何を探そうとしているか明確に自覚しているとは限らない．筆者が開発した Gards[3] では「毎日のお昼に何を食べるか選択する」という極めて曖昧な問題意識の元では，とにかくユーザに情報を提示し，その後のフィードバックを受け付け，それに対した情報を素早く提示することでユーザが当初想定もしていなかった「回答」に辿り着けることを示した．筆者が現在勤務している企業は HOME'S[1] という不動産情報検索サイトを運営している．自分が住む場所の選択というのは極めて重要な判断であるが，一生のうちに何度も行うものではなく普段は関心も薄い．その結果としてユーザは「自分が住みたいのはどのような物件か」「どのような物件に住めるのか」「そもそも条件として何が設定できるのか」といった問題に対して明確なイメージを持たずに検索を行うこともあると考えられる．そのため Gards と共通するコンセプトに基づく検索インタフェースを提供することも有効ではないか，という仮説の元 iPad アプリ「へやくる！」を開発した．

2 「へやくる！」について

2.1 コンセプト

前記の考察に基づくと，目指すべきユーザインタラクションの特性が明らかになる．すなわちユーザはどのような物件情報がそもそも存在しているのか，あるいはどのような検索条件が設定可能であり，それが何を意味するのかについて十分な知識を持っていないと想定する．このような場合には情報検索インタフェースはモードレス [4] なインタラクションで操作できることが必要であると考えた．すなわちユーザが自分の意思を反映させようとした際に，一連の手続きを強いる-すなわち特定の「モード」に入らせる-のではなく，興味，気分が赴くまま自由に最短手順での条件変更を可能とし，かつその結果をすぐに返す．これによりユーザは条件を設定することが，どのような意味を持つのか学びつつ物件情報を探索することができると考えられる．

このような特性を考慮し設計した「へやくる！」の画面イメージを図 1 に示す．このアプリでは，起動した瞬間から基本的に画面右側に物件情報が表示され続ける．ユーザはそれに対して，左側に並んでいるアイコンをドラッグ・アンド・ドロップすることで条件設定/解除を自由に行うことができる．多くの検索条件は最長 2 タップで設定，変更することが可能であり，変更後は即座に結果が変更される．ユーザが興味をもった物件がでてきた場合には，その上をタップすることで詳細情報を閲覧することができる．

2.2 実装

「へやくる！」は iPad 上のアプリケーションとして実装されている．検索条件を指定した際のソートアルゴリズムは，HOME'S サイトと異なる独自の物を用いているため，サーバー側に専用モジュールを追加している．

2.3 評価

「へやくる！」は商用サービスとしてリリースするため，ユーザビリティテストを行い要改善点を洗い出しその結果を仕様に反映させた．また目標とした「モードレスさ」について，他社製不動産情報検索用 iPad アプリとの比較を行った．その結果，他社製 iPad アプリでは一つの条件項目を設定，もしく

Copyright is held by the author(s).

* （株）ネクスト



図 1. へやくる！の画面

は変更するのに要する操作数は平均 4.3 であるのに対し,「へやくる!」では平均 1.7 であることが確認できた．このようにユーザが自分の意図を検索結果に反映させるために必要な操作数が少ないので，それだけユーザは思いつきを簡単に検索結果に反映させることができる．またインタフェースデザインには，開発とほぼ同時期に発表された iOS7 の新しいデザインコンセプトを反映している．アプリの主役はあくまでも物件情報（コンテンツ）であり，それ以外の不要な要素は極力排除し，画面デザインの面からも物件情報とユーザ間の距離を縮める事を狙っている．そうしたデザイン方針を徹底した結果，iOS7 全体のデザイン思想と合致した簡素で美しい画面デザインが実現できた．

3 関連研究

検索結果に対して，ユーザが能動的に自分の意図を反映させていく，という点では山本，中村らの Rerank-By-Example[2] がある．一度でユーザが意図する検索結果を得ることを期待せず，積極的にその意思を反映させることで望む結果にたどり着くことを目指している点が「へやくる！」と共通する．

4 まとめと今後の予定

「へやくる!」は現状細かい条件設定が必須となっている不動産情報検索に「モードレスさ」を持込み，ユーザがより気ままに閲覧できることを可能とした試みである．

今後実際に使用されたログデータを蓄積し，我々の狙いがどの程度成功を収めたか．また今後どのような方向に進むべきかについて検討する予定である．

参考文献

[1] HOME'S. http://www.homes.co.jp.

[2] 山本岳洋，中村聡史，田中克己: Rerank-By-Example：編集操作の意図伝播によるウェブ検索結果のリランキング情報処理学会論文誌（トランザクション）データベース, Vol.49, No.SIG7 (TOD37), pp.16-28, 2008 年 3 月.

[3] 大坪五郎: Gards - 変化し続ける興味に対応する情報推 薦, WISS2005 論文集, pp.31-36, 2005.

[4] ソシオメディア:「モードレス」https://www.sociomedia.co.jp/326.

ポータブルビジュアルプレイリストGRAPEの主観評価

魚田 知美　　伊藤 貴之*

概要. 個人が持ち歩く楽曲数は膨大になり，音楽コレクションの内容を理解することは困難になった．また，多くの人は音楽鑑賞の際，次に聞く曲を1曲ずつ逐一選択するよりも，プレイリストやアルバムといった音楽のまとまりを選択することが多いと考える．そこで著者らは，個人の音楽鑑賞の操作を支援するために，膨大になった音楽コレクションをプレイリスト単位で可視化するGRAPEを提案している．GRAPEではプレイリストを構成する各曲に対して，特徴量にもとづいて色を割り当て，特徴量の類似度に基づいて位置を決定することにより，プレイリストの内訳をグラデーション画像で表現する．この画像群を文字情報の代わりに表示するインタフェースを搭載した音楽プレイヤーを，PC上とAndroid音楽プレイヤー上で実装した．本報告では，いくつかのプレイリストを用いてユーザテストを実施した結果を示す.

1　はじめに

音楽の内容を把握するには，実際に聞く必要があるため，内容把握には時間がかかる．膨大かつ時間的メディアである音楽の内容を短時間で直感的に把握する一手段として，画像で全体像を表現する可視化が有用であると考える．また予備調査から，音楽鑑賞では楽曲を1曲ごとに逐一選曲する方法よりも，プレイリストやアルバム等の音楽のまとまりを選択する方法のほうが一般的だと考える．しかし既存の音楽可視化手法は歌手やジャンルに注目したものが多く，アルバムやプレイリストに着目した研究は少ない．特にプレイリストに関しては2009年時点で0件と指摘された[1].

そこで著者らは「1画像で1プレイリスト」を表すプレイリスト単位の音楽の可視化，GRAPE (GRadation Arranged Playlist Environment) を提案している．現在の音楽プレイヤーでは一般的に，プレイリストはその名前のみで表示されるものであり，その名前を知らない場合にはプレイリストの全体像はもとより，プレイリスト内の個々の楽曲の印象を推察することは難しい．それに対してGRAPEは，色のついた正方形のタイルを並べてプレイリストを表現することで，プレイリスト全体を通してどんな印象を有するかという全体像と，個々の曲の印象と，両方を同時に表現する．本報告では，GRAPEに関する主観評価結果を示す.

2　GRAPE

GRAPEでは，以下の処理によって生成されるグラデーション画像により，プレイリストを表現する．まず各楽曲の特徴量を算出する．続いて抽出した特徴量から，プレイリストを構成する各楽曲の位置を決定する．特徴量が類似する楽曲を近くに配置することで，プレイリスト全体の印象を視認しやすくする．さらに抽出した特徴量から，各楽曲の色を決定する．この各色で所定の位置を塗りつぶすことで，プレイリストの可視化結果となるグラデーション画像が生成される．この画像をユーザインタフェース上に搭載することで，ポータブルビジュアルプレイリストとして機能する．処理の詳細については[2]を参照されたい．可視化表示例を図1に示す.

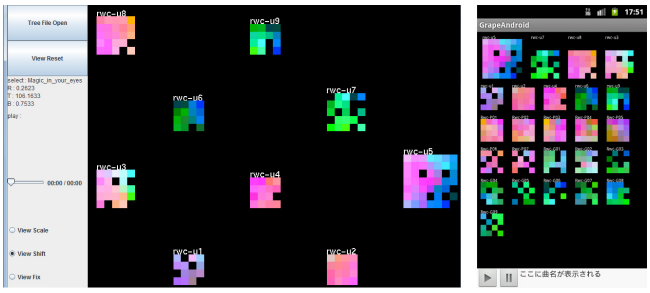


図 1. 可視化表示例 (左)PC (右)Android アプリケーション

3　ユーザテスト

被験者に可視化結果画像を観察してもらい，以下2つのユーザテストを実施した.

Q1：既存の楽曲群表現方法との比較

プレイリスト1, 2, 3, 4に対して，(A)GRAPEによる画像，(B) ジャンルアイコン，(C)CDジャケットを図2(左) のように用意した．それぞれのプレイリストに対して，プレイリストの中身やプレイリストの全体像が1番予想でき選曲の参考にできそうなものはどれか．(回答数138)

Copyright is held by the author(s).

* Tomomi Uota and Takayuki Itoh お茶の水女子大学大学院 人間文化創成科学研究科 理学専攻,

結果を図 2(右) に示す．プレイリスト 1，2，3，4 には市販されている楽曲を使用しているため，ユーザによってはCD ジャケットが既知であり，楽曲が予想できる可能性があった．プレイリスト 3，4 については (A) を選んだ人が 1 番多い結果となった．プレイリスト 1，2 については (A) を選んだ人は (B)(C) よりは少ないとはいえ約 4 分の 1 に達しており，よって (B)(C) よりも (A) を支持する人は一定数いると考えられる．併せて設けたコメント欄には下記が寄せられた．抜粋して下記に紹介する．

- 楽曲色をユーザが正確に読み取れるのか，デモ等で楽曲を聞いてみないと GRAPE の有用性がわからない (類似 6 件)
- アルバム単位でしか楽曲を聞かないため，CD ジャケットで十分 (5 件)
- 内容がわからない楽曲群はないため，有用性を感じない (4 件)
- 人と楽曲群を共有する際に有用と感じる (3 件)
- 聴きたい曲が具体的にきまっていない際，CD ジャケットのような具体的なものより，GRAPE の画像の方が有用だと感じる (3 件)

以上のコメントから，対象ユーザや使用場面を設問中に明確に設定し，かつ楽曲を試聴できる環境で再調査をしたいと考える．

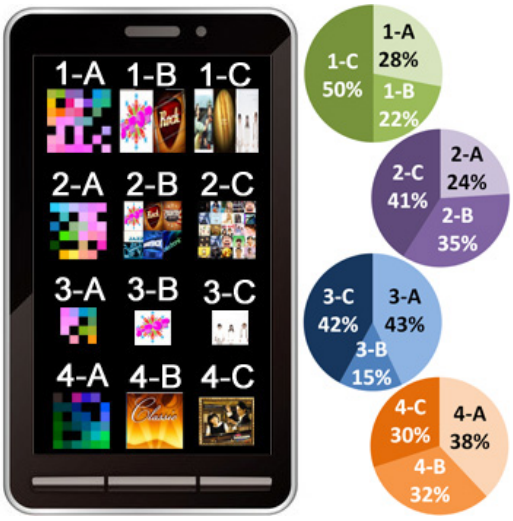


図 2. (左)Q1 に用いた画像 (右) 結果

Q2：プレイリスト選択の満足度

図 1 に示すプレイリスト 9 つの中から，今聞きたいプレイリストを GRAPE を用いて選択してもらい，自分の希望に沿ったプレイリストであったか実際に試聴して評価．また同時にランダムに選出したプレイリストについても試聴して評価．(4 段階評価，回答数 9)

使用した 9 つのプレイリストは RWC 研究用音楽データベース [3] の楽曲を利用して作成したものである．いずれの被験者もこの音楽データベースの収録曲を聴いたことがなく，GRAPE の可視化画像から得られる情報だけでプレイリストを選択したと言える．表 1 にまとめた回答結果から，全てのユーザにおいて，ランダム選出されたプレイリストより，GRAPE を利用して選択したプレイリストの評価の方が高かった．これにより，GRAPE が提示する画像はプレイリスト選択に有益な情報を提供していることが確認できた．

表 1. Q2 の結果
(希望に沿った選択ができた 4 ⇔できなかった 1)

被験者	プレイリストの選択方法			
	GRAPE		ランダム選出	
	PlaylistNo.	評価	PlaylistNo.	評価
1	7	4	8	2
2	2	3	5	2
3	5	3	8	2
4	7	2	1	1
5	6	4	7	3
6	5	3	1	2
7	2	4	9	3
8	7	4	3	2
9	2	3	6	2

4 まとめと今後の課題

本報告では，著者らが既に提案しているポータブルビジュアルプレイリスト GRAPE の主観評価結果を示した．ユーザテストから，既存の楽曲群表現方法に対するある程度の有用性と，未知のプレイリスト群選択時の有用性について議論した．

今後の課題として，各楽曲の色や位置が，人間が直感的に感じる色，それに伴った配置に近づくよう工夫することが挙げられる．また，GRAPE 上で簡単にプレイリストを作成・試聴・共有できる機能を設け，GRAPE の効果が最も期待できるプレイリストによる音楽鑑賞活動の活発化に繋げていきたい．

参考文献

[1] J. Donaldson and P. Lamere. Using visualizations for music discovery. In Tutorial at the 10th Int. Conf. on Music Information Retrieval, 2009.

[2] 魚田知美, 伊藤貴之, GRAPE: グラデーション画像によるポータブルビジュアルプレイリスト, 可視化情報シンポジウム, B215, 2013.

[3] 後藤真孝, 橋口博樹, 西村拓一, 岡隆一. RWC 研究用音楽データベース: 研究目的で利用可能な著作権処理済み楽曲・楽器音データベース. 情報処理学会論文誌, Vol. 45, No. 3, pp. 728-738, 2004.

肌マイクロスコープを用いた肌微細構造パラメータの診断

番場 文枝[1]　猪股 真美[1]　伊藤 貴之[1]　豊田 成人[2]　大高 瞳[2]　笹本 裕美[2]*

概要. 化粧品にはその様々な質感により肌を美しく演出する効果がある．肌の印象の違いは肌表面形状に起因しており，その効果を把握するために肌の微細構造と質感印象評価の関係を調べる必要がある．本報告では，肌マイクロスコープという専用カメラで撮影した肌の実写画像から肌微細構造パラメータを診断する一手法を示す．本手法では実写画像に対し二値化処理を適用し，続いてテンプレートマッチングに類似した手法で黒領域を探索することで，毛穴半径・皮溝幅・皮溝角度依存性を診断する．

1 はじめに

化粧品は女性にとって不可欠な商品であり，今日では多くの商品が販売されている．肌の質感印象と肌の表面形状の関係性を調べることはメーキャップ化粧品の効果を知る上でとても重要とされている．しかし，現在の印象評価では実際のモデルに化粧品を塗布して撮影した写真を使った伝達方法が主である．この方法では多くのバリエーションを再現するのはコストがかかるとともに，工学的な変化に大きな影響がある肌微細形状の要因は考慮されにくい．

そこで我々は，コンピュータグラフィクス（CG）を用いて微細形状を考慮した肌を生成し，肌の微細構造の違いによる肌質感変化のシミュレーションシステム [1] の開発に取り組んでいる．CG 画像の採用によって，従来よりも多彩な肌の印象評価が可能になると考えられる．本報告では上述の我々の研究のうち，肌撮影専用カメラの一種である肌マイクロスコープ（図 1 参照）で撮影した実写画像から，微細構造に関する数値を測定する手法を提案する．現時点での我々の実装では，まず肌の撮影画像を二値化して，続いてテンプレートマッチングに類似した手法で黒領域を検出し，これを毛穴や皮溝と判断する．以上の処理によって検出された毛穴や皮溝から，毛穴半径・皮溝幅・皮溝角度依存性に関する測定値のヒストグラムを生成する．

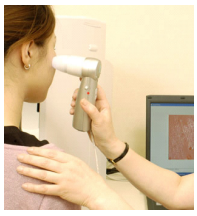


図 **1.** 肌マイクロスコープ.

Copyright is held by the author(s).

* 1) お茶の水女子大学　2) 資生堂リサーチセンター

2 肌微細構造パラメータの診断

本報告で提案する肌微細構造パラメータの診断手法は，実写画像を二値化した画像から円形領域を検出する，という単純な処理に基づいている．

2.1 二値化

まず入力画像に適応的二値化処理を適用する．各座標における R,G,B 値の総和が閾値以上の場合は黒画素に，さもなければ白画素に変換する．現時点では我々の経験に基づいて閾値を固定しているが，将来的には判別分析などの諸手法を適用して動的に閾値を設定できるようにしたい．

2.2 毛穴検出

続いて，二値化後の白黒画像に対して円形の黒領域を検出し，これを毛穴とみなす．ここでは図 2(左) に示すように，円形の黒領域をテンプレートとして，テンプレートマッチングに類似した以下の手法で円形の黒領域を探索する．

1. 円領域 (直径 $2r$) を画像内の全ての領域に走査させて，各位置において以下を実行する．
   (a) 円領域内部における黒画素数を集計する．
   (b) 円内部における黒画素数が閾値以上であれば，その円を毛穴と判別し，中心座標と円直径を記録する．
   (c) 毛穴と判別された円内部の黒画素を白画素に変換する．この変換は毛穴位置を除いた皮溝検出のために必要である．
2. 直径 $2r = 2r - 1$ として 1. に戻る．
3. 円の大きさがユーザ指定の最小値になるまで 1. と 2. を繰り返す．

以上により，各毛穴の中心座標，直径を記録する．

2.3 皮溝検出

続いて，毛穴以外の画素において細く伸びる黒領域を検出し，これを皮溝とみなす．ここでは図 2(右)

に示すように，16 方向の黒い線分のいずれかを引いた円をテンプレートとして，テンプレートマッチングに類似した以下の手法で細長い黒領域を探索する．

1. テンプレートを画像内の全ての領域に走査させて，各位置において以下を実行する．

   (a) テンプレートに引かれた黒い線分上における黒画素数を集計する．

   (b) 黒画素数が閾値以下であれば，その領域は皮溝ではないと判別し，1. に戻る．

   (c) さもなければ，その領域は皮溝であると判別する．テンプレートの平行移動によって皮溝の太さを推定し，テンプレートの中心座標，およびテンプレート上の線分の角度とともに記録する．

2. 16 種類のテンプレート全てに 1. を適用する．

以上の処理により，各皮溝の中心座標，角度，太さを記録する．

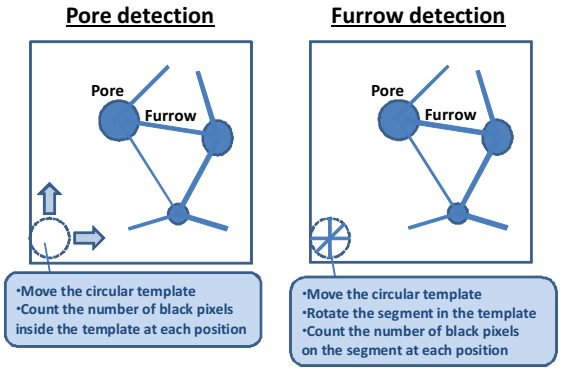


図 2. (左) 毛穴検出．(右) 皮溝検出．

3 実行例

本報告では，実写画像として代表的な特徴を持つ肌の実写画像サンプル 3 枚を入力として実験結果を生成した結果を示す．

健康な肌 特徴として，皮溝から生成される三角形が正三角形に近い形をしており均一であることが挙げられる．

乾燥肌 皮溝が一方向へ流れるため三角形を生成できていない場合が多い．

毛穴が目立つ肌 各々の毛穴の直径が広い．

上述のサンプル画像 3 枚に対して毛穴と皮溝を検出した結果を図 3 に示す．毛穴を水色で，皮溝をピンクで色分けしてある．乾燥した肌では毛穴が少なく皮溝も一方向に流れている様子が検出できている．また，3 枚とも毛穴の位置，毛穴の大きさ，皮溝の位置を概ね捉えられている．

続いて，検出された皮溝を角度別皮溝数で集計し，ヒストグラム化したものを図 4 に示す．横軸は 16 種類の角度，縦軸はその角度における皮溝数を表している．乾燥肌では単峰性の分布，つまり特定の一方向の皮溝が極端に多いことを表しており，乾燥肌の典型的な特徴が読み取れる．健康な肌，毛穴の目立つ肌では乾燥肌に比べ多峰性の分布を持ち，皮溝が流れていないことがわかる．

今後の課題として，実写画像から読み取れない立体的なパラメータの推定手法の確立や，リアルタイム処理を実現するための GPU 実装等があげられる．

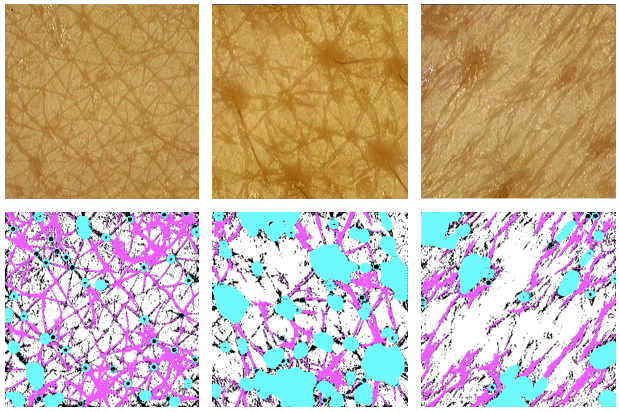


図 3. 実写画像と検出結果．(左) 健康な肌．(中) 毛穴の目立つ肌．(右) 乾燥肌．

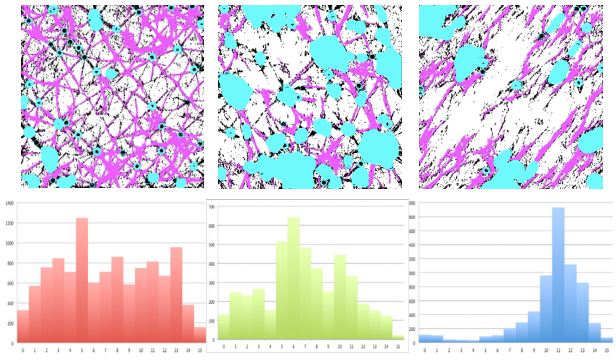


図 4. 角度別皮溝数のヒストグラム．(左) 健康な肌．(中) 毛穴の目立つ肌．(右) 乾燥肌．

参考文献

[1] T. Itoh, F. Banba, M. Inomata, M. Kurokawa, N. Toyoda, H. Otaka H. Sasamoto, Micro-Geometric Skin Simulation for Face Impression Analysis, NICOGRAPH International 2013.

多様性と類似性を重視したアパレル商品閲覧システム

小池 恵里子　伊藤 貴之*

概要. 現実の買い物では様々な商品を見て回る人は多く，特に女性にその傾向が強く見られる．本研究では，そのような買物に関する心理に着目し，効果的に商品を閲覧できるシステムを提案し，アパレル商品を例にした実装を示す．本手法では，キーワードの組み合わせに沿ってアイコンを生成し，対話型進化計算を用いた推薦アルゴリズムによって多様に選ばれたアイコンを表示することで様々な商品の閲覧を可能にする．一方でアイコンクリック時には画面を切り替え，アイコンに対応する類似商品を表示し比較しやすい商品提示を行う．このような，多様性と類似性に着目した商品提示を行うことで，現実のウィンドウショッピングのように演出を行い，効果的な買物行動を支援する．

1　はじめに

現実の店舗では，ウィンドウショッピングという名前にある通り，買い物時に様々な商品を見て回る人は多い．特に女性にその傾向が強く見られ，原因として「要求が曖昧.」「好きなもの以外にも注意が引かれやすい.」「細かいデザインにこだわりたい.」などがあげられる [1]．しかし，多くのサイトの検索システムはユーザが積極的にクエリを発する能動的検索になるため，要求が曖昧である場合には，クエリをつくる過程で精神的ストレスが生じることがある．例えば，ユーザは「かわいいセータが欲しい」という曖昧な要求をもっていたとする．この時，ユーザは”セータ”というクエリをシステムに発行するが，検索結果の商品件数が多いと，ユーザは商品の件数を減らすために自分の曖昧な要求を明確化して，クエリを書き換えなければいけない．また，発したクエリに何件の商品が検索されるか，ユーザは事前に知ることができないため，”該当する商品が無い”という最悪の状態も起こりうる．さらに，検索結果からその都度判断してクエリの修正を行うため，クエリ書き換え回数が過多になりやすい．その上，商品件数を絞って閲覧した後も”他の商品を見たい・気に入るデザインがなかった”というようなことから，検索は引き続き行われる．結果としてユーザ自身がかなり積極的になってクエリ修正を行わないと，様々な商品を閲覧することはできないということが起こる．これを解決するための一手段として，受動的に様々な商品を見ることができる商品閲覧システムを提案し，アパレル商品を例にした実装を示す．

2　要件定義

現実の買物では要求が漠然としていても目線を遠くにすれば多様な種類の商品が目に入り，そこから気になる商品に近づいて詳しく見ることができる．また，多くの店舗ではカテゴリ別（T シャツなど）・系統別（モード系など）というように類似した商品ごとに並べられ，近くで気になる商品を閲覧する際は，それらを細かく比較しながら自分の好みの商品を選ぶことができる．このような「遠くから見て多様な商品群を見る」と「近くから見て類似した商品を見る」という行為を繰り返すことによって「要求が曖昧でも様々なものを沢山見てみたい」という無意識な欲求を満たすことが可能である．このような現実の買物行動の観察から，以下の要件を満たすシステムを提案する．

要件 1: 曖昧な要求に対しても満足できる買物ができるように，明確なクエリを必要としない．

要件 2: 様々なものに興味を引けるように，ユーザの嗜好に沿いつつも多様な商品群の提示を反復的に行う．

要件 3: 比較しやすいように商品を類似品ごとに分類する．

3　研究概要

本研究では，開発者が用意したキーワードを各商品に付与し，そのキーワードを検索時に参照する．また，前処理としてキーワードそれぞれにデザインを用意し，キーワードの組み合わせを列挙して合成したアイコンを生成する．図 1 は商品の種別（T シャツ），色（赤），柄（ドット）のデザイン画像の合成によってアイコン画像を生成した例である．

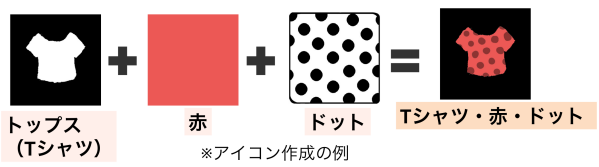


図 1. アイコン生成の例

Copyright is held by the author(s).

* お茶の水女子大学人文化創成科学研究科理学専攻

本研究は商品の多様性と類似性を考慮するためにアイコン画面と商品画面の２種類の画面にわける．図２にその表示例を示す． アイコン画面では Search



図 2. 左：アイコン画面　右：商品画面

ボタンをクリックする毎に推薦アルゴリズムによって多様に選ばれたアイコンが表示することで要件１と要件２を満たす．また，アイコンのクリック操作で商品画面に切り替わり，アイコンが示すキーワード群と同じキーワードを有する商品を表示することで要件３を満たす．

3.1　アイコン画面の可視化

本手法では可視化手法 FRUITSNet[2] のアルゴリズムをベースにして，類似したクエリをアイコン画面上で近くに配置することにより，検索結果の全体的分布を直感的に表現する．FRUITSNet では関連性のあるアイコン間をエッジで連結したネットワーク構造に対して，以下の条件を満たす一覧性の高い画面配置を実現する．

[**配置条件１**] 画像間の重なりの回避．
[**配置条件２**] 画像群の配置占領面積の低減．
[**配置条件３**] 関連性の高い画像群を近くに配置．

3.2　推薦アルゴリズム

本節では要件２を満たすために，以下の対話的進化計算アルゴリズムを導入している．

n: システムが用意するキーワード数
m: システムが一度に表示するアイコン数
$\vec{d} = \{d_1, ..., d_n\}$: 各商品について，各キーワードの付与の有無を表すベクトル．
$\vec{q} = \{q_1, ..., q_n\}$: ユーザの嗜好に関する各キーワードの重みを示すベクトル．Search ボタンを i 回押した時点での $\vec{q}$ の値を $\vec{q_i}$ とする．
A：現在表示されている x 個のアイコンの集合．
S：システムが優先的に表示するアイコンの集合．データ構造にキューを持つ．
C：種アイコンの集合．

この処理手順のうち Step1〜Step6 は，ユーザが Search ボタンを押すたびに実行されるものとする．

【Step0】ユーザの嗜好ベクトル $\vec{q_0}$ と優先アイコン集合 S を初期化する．

【Step1】種アイコン集合 C を空にする．表示アイコン群 A を空にし，A が m 個となるように優先アイコン群 S から順に追加する．

【Step2】表示アイコン群 A を表示する．

【Step3】Rocchio のアルゴリズム [3] に用いられるフィードバック算出式に類似した以下の式を利用し $\vec{q_i}$ を更新する．

D_i^+：i 回目にお気に入りに入れた商品の集合
D_i^-：i 回目に削除した商品の集合

$$\vec{q_i} = \vec{q}_{i-1} + \sum_{\vec{d^+} \in D_i^+} \vec{d^+} - \sum_{\vec{d^-} \in D_i^-} \vec{d^-} \qquad (A)$$

【Step4】ユーザが「お気に入り」に入れた商品の多いアイコンを，種アイコン集合 C に登録する．

【Step5】各々の種アイコン $c \in C$ に対して以下の処理を実行する．

1. 種アイコン集合に属するアイコン c について，ユーザの嗜好ベクトルからルーレットを用意する．
2. ルーレットから y 番目のキーワードを選ぶ．
3. 種アイコン c にて $d_x = 1$ であるキーワードのうち，y 番目のキーワードと排他的に出現すべき z 番目のキーワードについて $d_y = 0$, $d_z = 1$ としたベクトル d に対応するアイコンを，派生アイコン集合に登録する．このとき派生アイコンは i 回までに使用したアイコンと重複が無いように選択する．
4. この派生アイコンを，優先アイコン群 S に追加する．
5. 2.〜4. を，派生アイコンの個数だけ反復する．
6. 1.〜5. を，種アイコンの個数だけ反復する．

4　まとめと今後の課題

本報告ではウィンドウショッピングのような買物行動を考慮した検索支援システムを提案し，アパレル商品を例とした実装を示した．今後はユーザテストを行い，UI と推薦アルゴリズムの改善を行なっていきたい．

参考文献

[1] 木田理恵，彼女があのテレビを買ったワケ―男がわからなかった 女が商品を選ぶ本当の理由，エクスナレッジ，ISBN-978-4-7678-0697-6, 2008.

[2] T. Itoh, C. Muelder, K.-L. Ma, J. Sese: A Hybrid Space-Filling and Force-Directed Layout Method for Visualizing Multiple-Category Graphs, IEEE Pacific Visualization Symposium(2009),121-128.

[3] G. Saltom, The SMART retrieval system - experiments in automatic document processing, Prentice-Hall, 1971.

Touch-Shake:高齢者のコミュニケーションを促すインタラクティブシステムの開発と評価

山口 陽平　柳 英克　竹川 佳成 *

概要.　本研究は，高齢者のコミュニケーションを促すインタラクティブシステム“Touch-Shake”の開発と評価を行う．Touch-Shake は身体接触を検知する棒状のデバイスで，使用ユーザの静電容量に応じて Touch-Shake 本体から発する音と光が変化する．また，ユーザは接触相手や人数，個々の体格や体調に応じて様々なインタラクションを得られる．このため，ユーザは Touch-Shake を用いた身体接触を伴うコミュニケーションが楽しくなり，親密なコミュニケーションの機会が増える．本稿では介護老人保健施設の入居者が Touch-Shake を用いる試用実験を通じて，コミュニケーションの促進に関する有用性について検証した．

1　はじめに

我々のコミュニケーションには，言語と非言語の大きく 2 種類ある．非言語コミュニケーションには表情や視線，呼吸や身体接触などの非言語情報があり，コミュニケーションにおいて重要である．特に，身体接触を伴う対面コミュニケーションは親密なコミュニケーションであり，介護老人保健施設では，入居者にリラックス効果を与えるために用いられている [1]．また，Tatsumi ら [2] は，高齢者に意図的な身体接触を 3 分間行うと，高齢者は安心感を得ることを示した．しかし，日本人看護師と患者間で身体接触を行う時間は，コミュニケーションを行う総時間の 0.2%のみ [3] と大変少ない．

そこで本研究は，身体接触という親密なコミュニケーションを促すインタラクティブシステム，Touch-Shake の開発を行う．Touch-Shake には，身体接触を検知する静電容量センサが搭載されており，使用ユーザ個々の体格や体調，接触相手や接触の仕方に応じて光や音の反応を変化させる機能を持つ．この機能により，ユーザ同士の接触へのモチベーションが高まり，会話が生まれ，身体接触によるコミュニケーションの機会を増やせる．また，本稿は Touch-Shake[4] を高齢者向けに応用させたものである．

2　関連研究

Iida らが開発した Enhanced Touch[5] は，人体通信を用いた腕輪型のコミュニケーションツールで，人と接触すると LED が光る．Baba らが開発した Freqtric Drums[6] は，人に触れることで数種類のドラム音が鳴る電子楽器である．

これらは，人と接触すると音や光を発し，対面コミュニケーションを支援する点で本研究と類似している．しかし Touch-Shake は，使用ユーザ個々の体格や接触相手などに応じて楽曲と光が変化するため，接触行為に偶然性を取り入れる機能がある．また，接触中にデバイスを振って音や光を発するなど，接触中にもコミュニケーションを楽しめ，モチベーションを高められる点も他研究と異なる．

Copyright is held by the author(s).

* 山口 陽平　柳 英克　竹川 佳成

3　Touch-Shake の概要

3.1　システム構成

Touch-Shake の外観と使用中の様子を図 1 に示す．Touch-Shake は 2 本の棒状のデバイスから構成され，1 人が 1 本の棒を持って利用する．棒状を選定した理由は,「高齢者は見慣れない物を装着することに抵抗感を抱く」「マラカスやタンバリンなど，持って遊ぶ楽器を触ることに抵抗は無い」ことが，介護老人保健施設の事前調査で判明したからである．

3.2　Touch-Shake の機能

3.2.1　機能 (a)

図 2-(a) に示すように身体接触を検出すると，スピーカから楽曲が再生されて LED が光り，身体を離すと楽曲は止まり LED は消灯する．楽曲は，高齢者が聞き慣れている童謡かつ，Tatsumi ら [2] の研究に沿って身体接触時間が 3 分以上になるよう 1 曲 3 分以上の曲を選定した．

3.2.2　機能 (b)

図 2-(b-1) に示すように，静電容量センサの感度を高めることで閾値を設定し，ユーザが身体接触することで変化した静電容量に合わせて楽曲と LED が変化する．この特性を活用し，32 種類の楽曲と LED の光色パターンを用意することで，図 2-(b-2) に示すようにユーザの体格や接触する人数などによって楽曲や光色パターンが変わる．また，接触面積を大きくし，勢いよく相手と接触することでも楽曲と光色は変化する (図 2-(b-3))．

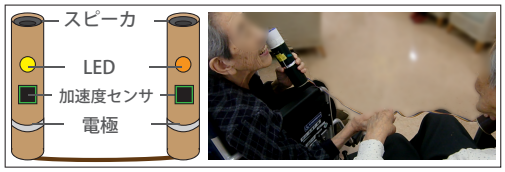


図 1. Touch-Shake の外観と基盤図

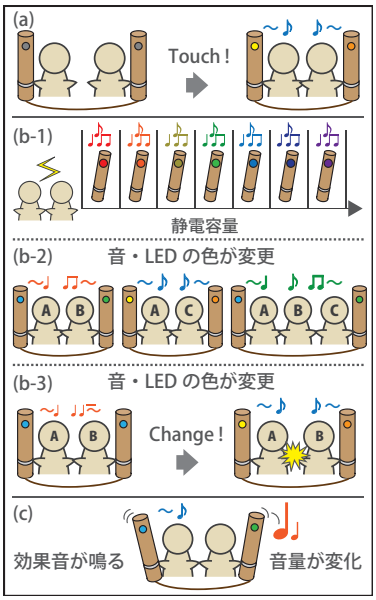


図 2. Touch-Shake の機能

3.2.3 機能 (c)

図 2-(c) のように，接触中に片方のデバイスを振ると効果音が鳴り，もう一方を振ると音量調節できる．

4 評価実験

4.1 実験方法

Touch-Shake が介護老人保健施設の入居者のコミュニケーションにどのような影響を与えるか実験を行なった．被験者は 80 代の方 1 人で，1ヶ月間介護者や入居者，家族と Touch-Shake を使用していただいた．実験するにあたり，介護者や看護士の方に被験者の様子を書いていただきながら調査を行った．

4.2 結果と考察

実験初日，被験者は初めて見る Touch-Shake に対して「ワクワクしてくるね」と期待感を抱いていた．初日は入居者と Touch-Shake を使用し，「よく思いつくね」「面白い」など会話をしながらお互い楽しんでいた．実験 3 日目には Touch-Shake の使用に慣れ，音楽に合わせて歌を歌い，喜ばれていた．また，相手によって楽曲や LED が変化することを不思議に思いながらも，笑顔で身体接触を行っていた．孫と使用した時は，お互い少し照れ笑いをしていたが，特に被験者は大変喜んだ顔で Touch-Shake を通じたコミュニケーションを交わしていた．しかし，難聴のため音が少し聞こえにくい様子であった．その後，被験者は計 9 名の方と共に Touch-Shake を 1 日 15-20 分使用しており，心身的に良効果を得られるとされている 3 分を大幅に超えていた．また，「色々な人と遊べて楽しい」と，Touch-Shake を使用するモチベーションも保たれていた．

以上の結果から，被験者は 1ヶ月の長期的期間でもモチベーションを保っていたことが判明した．また，相手と楽しく会話や歌を歌っていた上に，身体接触への抵抗感を抱いていなかった．これらより，親密なコミュニケーションを促進できたと考えられる．機能に関して，接触相手によって音や光の演出が変化することに楽しさを感じ，様々な人と接触するモチベーションに繋がっていた．また，Touch-Shake 自体に対しても抵抗感は無く，期待感を抱いていた．

5 まとめ

本研究では，高齢者のコミュニケーションを促すインタラクティブシステム’“Touch-Shake”を構築した．実験結果から，ユーザは Touch-Shake を用いた身体接触によるコミュニケーションが楽しくなり，笑顔で相手と会話や歌を歌っている様子が多く見られ，コミュニケーションを促進できた．

参考文献

[1] 浅井さおり, 田上明日香, 沼本教子, 西田真寿美, 高田早苗, 介護老人保健施設での看護場面におけるタッチの特徴, 老年看護学 : 日本老年看護学会誌 : journal of Japan Academy of Gerontological Nursing 7(1), 70-78, 2002-11.

[2] K. Tatsumi, Y. Adachi, Y. Yokota, M. Ashikaga, S. Tanaka, T. Sakai, Effects of Body Touching Therapy on the Elderly. In *Proceedings of the Journal of International Society of Life Information Science 18(1)*, pp. 246–253, 2000.

[3] 山口創, 看護師-患者間の非言語行動の実際と課題 -身体心理学の立場から- , 桜美林論考. 心理・教育学研究 2, 73-83, 2011-03.

[4] Yamaguchi, Y. Yanagi, H., Takegawa, Y. Touch-Shake: Design and Implementation of a Physical Contact Support Device for Face-to-Face Communication, *Proceeding of IEEE Global Conference on Consumer Electronics (GCCE2013)*, pp. 170–174 (Oct. 2013).

[5] K. Iida, K. Suzuki. Enhanced Touch a Wearable Device for Social Playware. In *Proceedings of the ACM 8th Advances in Computer Entertainment Technology Conference*, 2011.

[6] T. Baba, T. Ushiama, K.Tomimatsu, Freqtric Drums : a Musical Instrument that uses Skin Contact as an Interface. In *Proceedings of the 7th international conference on New interfaces for musical expression*, pp.386–387, 2007.

倍音分析によるいい声作りの支援アプリに向けて

菅原 衣織　伊藤 貴之*

概要. なぜこの人の演説は記憶に残るのか．なぜあの人の歌は心に響くのか．言葉やテクニックだけでは伝わらないこの熱い想い，それを伝えてくれるのが声色である．声色の表現次第でいい印象を与えることも悪い印象を与えることもできる．では，いい印象を与える声とはどんな声であるのか．我々は必要とされる声の印象は用途や状況によって変化するものであると考えた．そこで，本研究では倍音に着目し音声を分析し，倍音が与える印象に基づいて声の傾向を視覚化し提示することで，ユーザが自発的に各々の目指す声に近づけるような支援システムの開発を目標にしている．

1　はじめに

1.1　日常生活での声の役割

私たちが生活する上で切っても切り離せないコミュニケーション．そのコミュニケーションを円滑にし，私たちの人柄をことば以上に伝えてくれるのが，声である．明るく明瞭な声で話している人にはいい印象を受けることが多く，逆にぼそぼそとはっきりしない声で話す人には好印象を受け難いことが多い．また，メラビアンの法則 [1] として従来から知られているように，コミュニケーションにおいて人の行動が他人にどのように影響を及ぼすかという実験では，言語情報が 7 %，聴覚情報が 38 %，視覚情報が 55 %を占めるという考え方がある．このように，声の印象が自己の評価に与える影響は少なくない．

1.2　本研究におけるいい声

本研究では，声の印象，つまり声色の観点から「いい声」を追求し，これを支援するシステムの作成を目標にした．一般的に「いい声」と判断するためには，滑舌や話の速度，声の大きさなどいくつかの要素が複雑に絡んでくると考えられる．しかし，滑舌は早口言葉の練習を積むことで良くなり，声の大きさは拡声器を使うことで補正が可能であるように，いくつかの要素は既存の手法を用いて改善される．そこで本研究では声色の印象に絞って議論を進める．

声色の印象を左右する要素として，我々は倍音に着目した．音楽理論などにおいて倍音とは，楽音の音高とされる周波数に対して，2 以上の整数倍の周波数を持つ音の成分を指す．このとき楽音の音高を決定する成分を基音と呼ぶ．基音と倍音はフーリエ変換等の周波数変換手法を用いることで検出が可能である．以上のように定義される狭義の倍音を本報告では「整数次倍音」と呼び，基音に対して非整数倍の周波数を持つ音の成分を「非整数次倍音」と呼ぶ．図 1 は整数次倍音が大きく含まれる声，図 2 は非整数次倍音が大きく含まれる声に対して，横軸を周波数，縦軸を音量として周波数分布を示したものである．尺八奏者の中村は著書 [2] の中で，整数次倍音が強いとカリスマ性や明朗性，豊かさが印象付けられ，非整数次倍音が強いと情緒や親密感が印象付けられると述べている．本研究ではこれらの特徴を分析することで声色の印象を評価する．また本報告ではこれ以降，非整数次倍音を含む広義の倍音を「倍音」と称する．

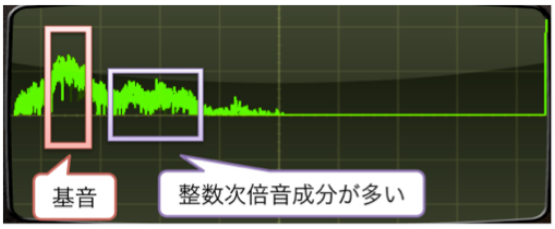


図 1. 整数次倍音が大きく含まれる声

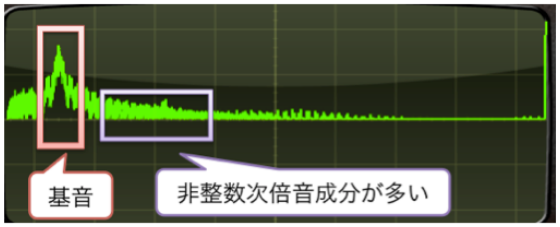


図 2. 非整数次倍音が大きく含まれる声

2　関連研究

人間の発声に着目し，ユーザがいい声を出せるように身体的スキルの向上を促すメディアの提案を目

Copyright is held by the author(s).

* お茶の水女子大学大学院人間文化創成科学研究科理学専攻情報科学コース

的とする発声のメタ認知促進システム”いい声マイク”の提案 [3] では，いい声を「音としてしっかりと発せられていて，響いている声」と定義している．

また，モテ声診断 VQ チェッカー [4] というウェブサイトおよびスマートフォン用の人気アプリケーションがある．このサービスでは「声の高さ」「声の大きさ」「一音の長さ」「耳への入りやすさ」「滑舌の良さ」の 5 つの項目から声の良さを定義している．

3 いい声作りの支援アプリに向けて

本研究では，ユーザが自身の声色を改善するための支援を目的としたアプリケーションを開発している．ユーザがモバイル端末上で本アプリケーションを用いて日常的に自身の声を省みることができるよう，我々は Android 上で本アプリケーションを開発している．本アプリケーションでは，まず Android の Visualizer クラスを用い音声に高速フーリエ変換を適用し，その周波数分布を取得する．そして取得した周波数分布から，倍音の傾向を導く指標として以下の二値を算出する．

- 倍音を構成する各周波数領域の音圧の総計 P_t
- 整数次倍音の音圧の倍音全体に対する比率 P_i

この算出結果により，自身の声にどの程度豊かな倍音が含まれているのか，また整数次倍音が強いのか非整数次倍音が強いのか，といった情報を得ることができる．

本アプリケーションの処理手順を図 3 に示す．声色を示す倍音成分を視覚的にユーザに提示する方法として，我々は現在以下の 2 種類を開発中である．

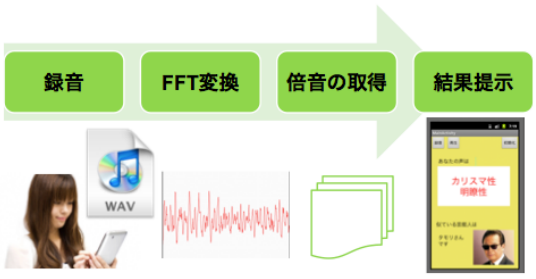


図 3. 本アプリケーションの処理手順

3.1 芸能人を例にした提示

本アプリケーションでは結果提示の際に，同じような倍音傾向にある芸能人を例に挙げて提示する．例えば，整数次倍音成分が大きければ（= P_i が大きければ）タモリ氏や黒柳徹子氏を，非整数次倍音成分が大きければ（= P_i が小さければ）明石屋さんま氏や桑田佳祐氏を，倍音成分が少なく基音に近い声の場合は（= P_t が小さければ）クリス・ペプラー氏を提示している．芸能人の倍音の特徴取得には，サンプルボイスやラジオやテレビ番組で取得した声，そして中村の著書 [2] の内容を参考にしている．

3.2 視覚変数による提示

我々は現在，本アプリケーションを拡張し，より感覚的に倍音成分を理解できるように，倍音の傾向や音圧などを視覚変数に割り当てて提示する機能を開発中である．この機能では，草間らの MusCat [5] のデザインを参考にした図 4 のようなデザインを採用している．

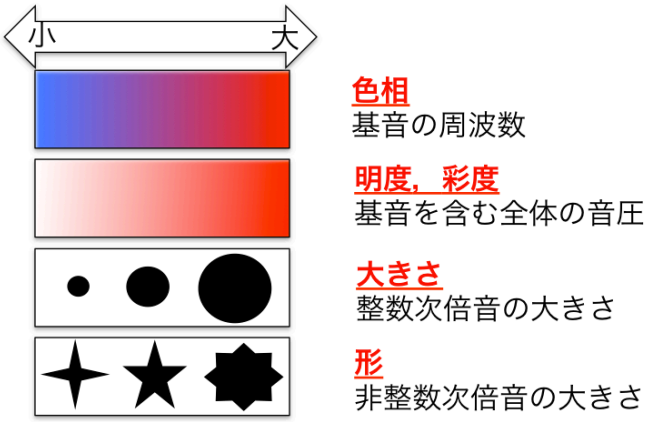


図 4. 倍音成分を視覚変数に割り当てたデザイン

4 まとめと今後の課題

本報告では，個人のニーズに応じた声色改善の支援を目的として，倍音に着目した声色提示のアプリケーションを提案した．本アプリケーションでは倍音成分の算出結果を，芸能人に喩える，視覚変数を用いる，の 2 種類を用いて提示する．今後の課題として，実際に被験者に本アプリケーションを一定期間使用してもらい，被験者の声がどのように変化するか観察したい．

参考文献

[1] A. Mehrabian: *Silent messages*, Wadsworth, Belmont, California (1971).

[2] 中村明一: 倍音 音・ことば・身体の文化誌, 春秋社 (2010).

[3] 矢島佳澄，筧康明，諏訪正樹: 発声のメタ認知促進システム”いい声マイク”の提案，情報処理学会インタラクション 2011 (2011).

[4] 声総研: モテ声診断 VQ チェッカー, http://www.koesouken.com/vqchecker/(2001).

[5] K. Kusama, T. Itoh, Abstract Picture Generation and Zooming User Interface for Intuitive Music Browsing, Springer Multimedia Tools and Applications, 10.1007/s11042-012-1108-y(2012).

Unisoner: 同一楽曲を歌った異なる歌声を重ね合わせる合唱制作支援インタフェース

都築 圭太　中野 倫靖　後藤 真孝　山田 武志　牧野 昭二*

概要.　本稿では，Web 上で公開されている「一つの曲を様々な歌手が歌った歌声」を複数重ね合わせて作られる合唱作品を対象とし，その制作と鑑賞を同時に実現できる合唱制作支援インタフェース Unisoner を提案する．従来，多様な楽曲を組み合わせて新たな作品を制作するインタフェースは提案されてきたが，同一楽曲の異なる歌声を重ね合わせる合唱作品のためのものはなかった．また，合唱作品を制作するために必要なソフトウェアやその操作は専門的で，伴奏音の抑制や波形に基づいた手作業での位置合わせが必要である等，制作の敷居が高かった．Unisoner は歌声の出現時刻と左右チャネルの音量比をマウス操作だけで指定でき，専門知識のないユーザでも手軽に合唱制作できる．ユーザはリアルタイムに生成される合唱を聴きながら，歌声を切り替えたり追加で重ね合わせたりするインタラクションによって合唱制作する．

1　はじめに

近年，様々なユーザが既存の楽曲を歌唱した音楽作品（以下，単に「歌声」と呼ぶ）が Web 上で多数公開されている．これによって，同一楽曲を多様な歌い回しや声質の歌声で音楽鑑賞できるようになり，これらを切り替えながら重ね合わせた合唱のような音楽作品（以下，単に「合唱」と呼ぶ）も制作されるようになった（図 1）．

合唱を制作するには，まず伴奏を抑制して個々の歌声を抽出する（カラオケ音源等を利用）．その後，それら複数の歌声波形の切り貼りや，左右チャネルの音量比調整など専門的で煩雑な手作業が必要となる．そのため，専門知識がない人の合唱制作は敷居が高い．もしこのような敷居を下げることができれば，誰でも合唱コンテンツの創作を楽しむことができる．さらに，それぞれの歌声が一つの作品であることから，手軽な合唱制作は音楽創作でありながら能動的な音楽鑑賞 [4] としての可能性も持つ．

そこで本稿では上記で述べたような手作業の一部を自動化するだけでなく，直感的な歌声の重ね合わせインタラクションが可能な合唱制作支援インタフェース Unisoner（ゆにぞな）を提案する．従来，複数の楽曲を対象としてそれらを連続的に再生するような音楽鑑賞インタフェース [5,7] や，楽曲をマッシュアップ（重ね合わせ）するインタフェース [1–3] は研究されてきたが，多様な歌声を重ねる合唱に着目したインタフェースはなかった．

Copyright is held by the author(s).

* Keita Tsuzuki, 筑波大学大学院 システム情報工学研究科, Tomoyasu Nakano and Masataka Goto, 産業技術総合研究所 (AIST), Takeshi Yamada and Shoji Makino, 筑波大学 システム情報系

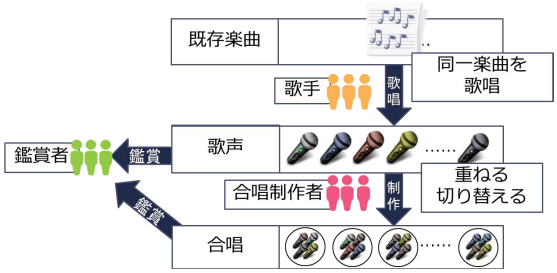


図 1. 歌声と合唱の派生関係.

2　手軽に合唱制作できる Unisoner

Unisoner の動作画面を図 2 に示す．Unisoner では個々の歌声をアイコンで表し (図 2 Ⓐ)，それらをステージ上 (図 2 Ⓑ) に配置するインタラクションによって合唱を制作する．このような歌声アイコンは自動推定された歌手の性別に応じて色が異なる (男声が青，女声が赤)．ここで，歌声を切り替えたり，新たに重ねる歌声を追加したりといった時刻の指定を歌詞に基づいて行い，そのような時刻間をセクションと呼ぶ．ユーザはリアルタイムに生成される合唱を聴きながら，セクションを区切ったり歌声を配置したりすることで，合唱の制作を行う．

以降，そのような切り替え時刻の指定と歌声の配置について説明する．

歌詞に基づくセクションの指定

歌詞は楽曲中の時刻を一覧性高く把握するために有用な情報であるため，これを活用する．従来の時間を用いた時刻指定に比べて特定の箇所の再生がしやすく，効率が良い．インタフェース左側に表示されている歌詞 (図 2 Ⓒ) は「セクション分割ボタン」(図 2 Ⓓ) によって分割できる．その区切り位置は歌

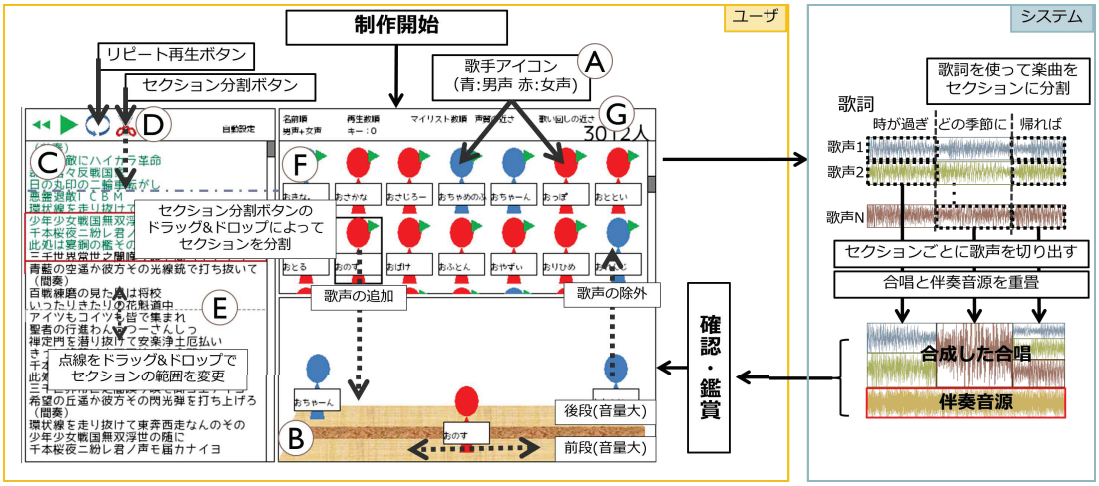


図 2. Unisoner の実行画面とインタラクションの流れ．ステージに女性の歌声と二人の男性の歌声を配置した例．

詞上に点線 (図 2 Ⓔ) として表され，マウスでドラッグ& ドロップすることでその位置を変更できる．

歌声アイコンのステージ上への配置

Unisoner ではセクション毎に配置する歌声や，それぞれの音量及び左右チャネルの音量比を，歌声アイコンのステージ上への配置によって決定する．

歌声アイコンは一覧 (図 2 Ⓕ) から選択し，マウス操作でステージ上へ配置するが，その際，選択を効率化するためのソート機能とフィルタリング機能を持つ (図 2 Ⓖ)．歌手名（動画投稿者名で代用），再生数，マイリスト数，歌い回しと声質の近さに基づくソートと，歌手の性別とキー（相対的な差）によるフィルタリングが可能である．

ステージにおける前段後段は音量の大きさを表現し，前段の方が大きい．また，ステージの左右は左右チャネルの音量比に対応しており，例えば左方向へ動かすと左チャネルの音量が大きくなる．

Unisoner の実現方法

Unisoner を実現するためには同一楽曲（原曲）を歌った複数の歌声と，それら歌声から伴奏を抑制するために，伴奏音源との時刻合わせとキーの違いを自動推定する信号処理技術が必要である．これら信号処理の詳細は文献 [6] で説明している．現在の実装では，歌声の性別はこのような信号処理結果によって，原曲の音高と比較してオクターブ異なっていた場合は，原曲と異なる性別を付与した．例えば，原曲が女性でそれよりも 1 オクターブ低い音高の歌声は男性とした．オクターブ以外の異なり方をしていた場合は，近い性別を選択して付与した．

Unisoner はニコニコ動画[1] において，最も多く歌われた一つの楽曲[2]を対象として動作を確認しており，それ以外の楽曲に対しても運用可能である．ここでは歌声に加えて，再生数，マイリスト数，動画投稿者のユーザ名も活用する．

[1] http://www.nicovideo.jp

3 おわりに

本稿では合唱制作支援インタフェースを提案した．今後は，それぞれの歌声の声質や歌い回しを可視化する等，創作に有益な情報を提示する機能の拡張を検討している．

謝辞

本研究の一部は，JST「OngaCREST プロジェクト」の支援を受けた．また，ニコニコ動画上の楽曲を扱う上で協力頂いた濱崎雅弘氏，石田啓介氏に感謝する．

参考文献

[1] M. Davies, *et al.* AutoMashUpper: An Automatic Multi-Song Mashup System. *Proc. ISMIR 2013*, 2013.
[2] N. Tokui. Massh!—A Web-based Collective Music Mashup System. *Proc. DIMEA 2008*, 2008.
[3] 宮島. Music Mosaic Generator: 高精度時系列メタデータを利用した音楽リミックスシステム. WISS 2007 論文集, 2007.
[4] 後藤. 音楽音響信号理解に基づく能動的音楽鑑賞インタフェース. 情処研報, 2007-MUS-96, 2007.
[5] 後藤 他. Musicream: 楽曲を流してくっつけて並べることのできる新たな音楽再生インタフェース. WISS 2004 論文集, 2004.
[6] 都築 他. 様々な歌手が同じ曲を歌った歌声の多様さを活用するシステム. 情処研報, 2013-MUS-100, 2013.
[7] 堀内 他. Song Surfing: 類似フレーズで音楽ライブラリを散策する音楽再生システム. *PIONEER R&D*, 17(2), 2007.

[2] ニコニコ動画における動画番号 sm15630734 を使用．2013 年 10 月現在，この楽曲に対して 5166 曲の歌声がある．

自由な筆記にロバストな手書き表認識システム

中洲 俊信　　柴田 智行　　井本 和範　　三原 功雄*

概要. 自由に手書きされた表から，罫線，文字列，行や列などの表構造を安定して認識できる手法を提案する．手書きの表には，線の歪みや線同士の接続ずれなどの変動が生じる．変動にロバストな認識を実現するため，ユーザが手書きした多くの表から主要な変動パターンを洗い出し，それらに基づいて，罫線の抽出，表構造の認識，セル内ストロークの抽出の 3 段階で表を認識する技術を開発した．手書きの表 880 サンプルを対象とした評価実験で認識率 98.6 %を達成し，提案手法の有効性を示した．また，罫線の直線化や罫線同士が正しく接続するような修正により，表を視覚的に整える表整形インタフェースを試作した．

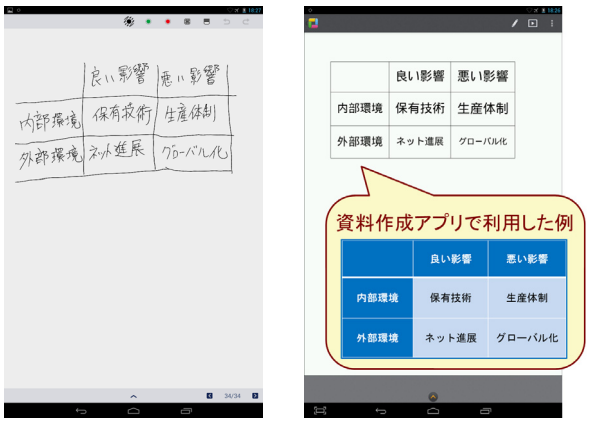
1　はじめに

近年，ペン入力に対応した携帯端末が注目され，タブレットやスマートフォンに書いた筆記をデジタルデータとして処理することが可能になった．頭の中のアイデアを思いつくままに自由に書きとめ，それをすぐに整理･活用できることが価値の高い発想につながると考えられる．本稿では，様々な情報の関係性をわかりやすく可視化し情報整理に役立つツールとして頻繁に使われる「表」に注目し，ユーザが自由に手書きした表から罫線，文字列，行や列などの表構造を認識し，即座に綺麗な表へと変換するシステムを提案する（図 1）．

手書きの表は，線に歪みがあったり線同士の接続がずれていたりする．また，ユーザによって書かれる表の形式も様々である．手書き表認識に関する従来研究では，このような手書きの表に含まれる変動を広く洗い出して整理したものは少ない．本稿では，事前検討として自由な手書きに伴う変動のパターンを洗い出し，主要な変動パターンに対してロバストな表認識手法を設計する．具体的には，筆跡の大まかな形状や関係性など特徴的な情報を利用し，単純な表のモデルに対応づけることでロバストに表罫線や表構造を認識する．Lin らによる従来技術 [2] も手書きに対するロバスト性を考慮しているものの，図 1(a) のように外周の罫線が書かれていない表を想定しておらず，認識が困難であった．本稿では，外周の罫線が書かれていない表も主要な変動パターンであると考え，認識できるように対応する．

2　手書きの表の特徴分析

ユーザが手書きした表などを参考に，手書きの表に含まれる変動のパターンを洗い出した．抽出した事例を図 2 に示す．これらに加え，(i) 複数のストローク（ペンが入力面に接してから離れるまでの 1 画の軌跡）を追記して 1 本の罫線が書かれている，(j) 1 本のストロークで複数の罫線が書かれているというパターンを抽出した．ここで，1 本の罫線とは，表の左端から右端までを通る各行の境界，および，表の上端から下端までを通る各列の境界である．



(a) 入力例　　(b) 認識・整形処理結果

図 1. 筆者らが開発した手書き表認識システム [1]

3　手書き表認識手法

本提案手法では，ペンデバイスから得られた筆跡を入力とし，罫線の抽出，表構造の認識，セル内ストロークの抽出の 3 ステップで表を認識する（図 3）．

3.1　STEP1：罫線の抽出

長さが閾値以上で傾きが水平／垂直に近い線分を含むストロークを，罫線を構成するストロークとして抽出する．その後，1 画で複数本の罫線が書かれている場合や複数画で 1 本の罫線が書かれている場合に対応できるよう，方向が変化する点でのストローク分割や，同方向で位置的に近いストローク同士の統合を行い，水平／垂直の各罫線を認識する．

Copyright is held by the author(s).

* 株式会社東芝 研究開発センター

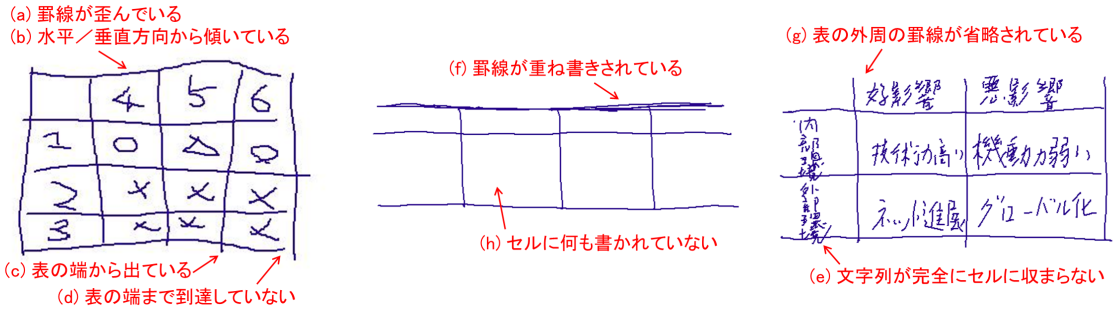


図 **2.** 手書きの表に含まれる変動パターン

3.2 STEP2：表構造の認識

抽出した罫線を直線近似し，歪みのある罫線で構成された手書きの表を単純化する．その後，罫線の接続関係や，外周の罫線が省略された意図的な罫線のはみ出しを認識する．意図的に罫線がはみ出して書かれている場合，はみ出し部分が相対的に長かったり，はみ出し部分で構築された領域に文字列が書かれていたりする．この知見から，罫線のはみ出し部分が他のセルのサイズに対して所定割合以上の長さである場合や，はみ出した罫線付近に閾値以上の長さのストローク群が存在する場合に，意図的な罫線のはみ出しと判定し外周罫線を補う．意図的なはみ出しと判定されなかった場合は，垂直（水平）罫線の端点を最も端の水平（垂直）罫線に接続する．

3.3 STEP3：セル内ストロークの抽出

セル内の文字列の一部がセルからはみ出す場合も想定し，各セルに半分以上含まれるストロークをそのセルに内包するストローク群として抽出する．抽出したストローク群は一般の手書き文字認識技術により文字として認識する．

4 表整形インタフェースの試作

ユーザが手書きした表（図 1(a)）に対し，罫線の直線化，外周罫線の補足，罫線の端点での接続処理を施し，即座に視覚的に綺麗な表へと変換する．また，セル内の文字列を所定のフォントに変換する．なお，その場での表編集や，資料作成アプリの表オブジェクトとしての活用も可能である（図 1(b)）．

5 認識性能評価実験

被験者 40 名が作成した 880 サンプルの表を対象に，本提案手法の認識性能を調べた．これらのサンプルは，外周の罫線が無いものやセル内が空欄になっているものなど 11 種類の多様な表を被験者が 2 回ずつ書き写して作成した．また，各表には，行数，列数，罫線を構成するストローク，各セルに含まれるストローク群が特定できるように正解を教示した．

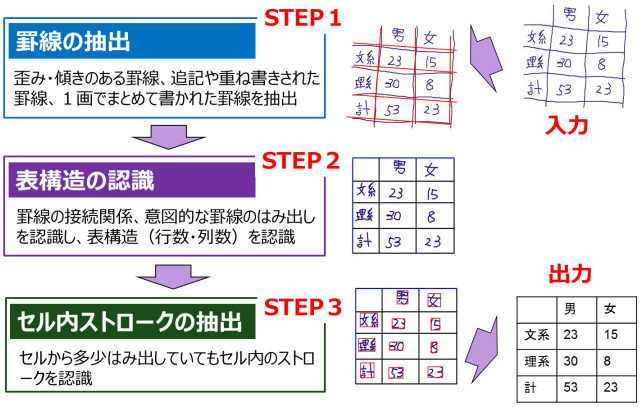


図 **3.** 手書き表認識システムの処理過程

本提案手法の認識結果の再現率と適合率の調和平均として認識率を算出した結果，98.6 %となり，様々な形式の表を高精度で認識できていることがわかった．また，様々な書き方（罫線の歪み，接続のずれ，追記された罫線など）にも対応できていることが確認でき，本提案手法の有効性が示された．

6 まとめ

手書きの表から罫線，文字列，表構造を認識可能な表認識技術を提案し，表整形インタフェースを試作した．なお，本技術はペン入力対応タブレットへの搭載という形で実用化された [1]．今後はシステムの操作感に関するユーザ評価を予定している．

※本稿に掲載の商品，機能等の名称は商標として使用している場合があります．

参考文献

[1] ”東芝製ペン入力タブレット　AT703”，東芝. http://www.toshiba.co.jp/regza/tablet/.

[2] Z. Lin, J. He, Z. Zhong, R. Wang, and H.-Y. Shum. Table Detection in Online Ink Notes. In *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.28, No.8*, pp. 1341–1346, 2006.

部品へのプロジェクションマッピングによる通電しない電子工作

秋山 耀　　宮下 芳明 *

概要.　提案システムは，タッチパネルディスプレイ上で電子部品を仮想的に結線し，プロジェクションマッピングによって各種 LED などを発光しているように見せる電子工作環境である．音やアニメーションによるブザー・バイブレータの疑似動作，画像認識などによる入力素子の動作にも対応している．指を近づけると端子が拡大する機能により，ドラッグ操作だけで結線・断線が行える．電子部品店に設置することで未開封のまま部品の動作や仕様を確認できるほか，複数人でわいわいと作業を行うこともできる．

1　はじめに

電子回路を作成する際に，作成者は回路図を参照して回路を作成するが，実際に配線をすると，難しい場合がある．そもそも回路図は，実体配線図とは異なり，回路図上での位置と実際に配置する場所が無関係であり，回路図そのままの配置や結線は現実的ではない．故に，初学者には回路図から回路を作成することが難しい場合がある．本稿では，回路図通りの配線を可能とするとともに，部品に通電しないどころか，部品を開封することさえせずに電子工作を行えるシステムを提案する．提案システムは，電子部品にプロジェクションマッピングを行うことによって，LED などの発光素子が実際に動作し，光っているように見せるものである．また，音やアニメーションによるブザーやバイブレータの疑似動作，画像認識などによる入力素子動作にも対応している．「回路図通りに配線すれば良いので」，初学者向けの電子工作ツールとすることができると考えている．ドラッグ操作だけで行えるデザインを目指し，そのための支援機構を用意した．

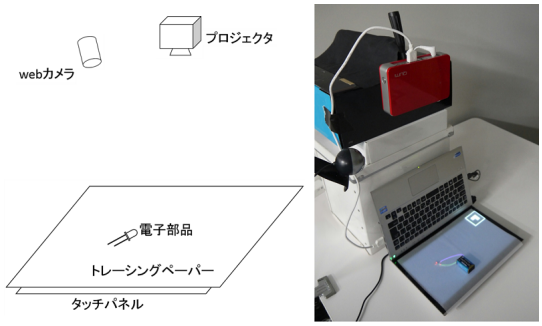


図 1: 提案システム全体図

Copyright is held by the author(s).

* Yoh Akiyama，明治大学 理工学部 情報科学科，Homei Miyashita, 明治大学 総合数理学部 先端メディアサイエンス学科, 独立行政法人科学技術振興機構, CREST

2　システム

提案システムはタッチパネルディスプレイとプロジェクタを用いている（図 1）．タッチパネル対応のディスプレイが水平に置かれており，上部にはプロジェクタと web カメラが設置されている．プロジェクタの表示範囲がディスプレイにちょうど重なるような高さにプロジェクタを固定し，表示位置とタッチ位置を合わせている．なお，ディスプレイ面における透過光（バックライト）と反射光（プロジェクタ）のバランスをとるためにディスプレイ上にトレーシングペーパーを敷いている．

提案システムは，ユーザの操作･作業の内容を検出する入力部と，素子にプロジェクションを行う出力部に分かれている．提案システムの実装は Processing で行った．AR マーカーを使用してディスプレイの表示領域とカメラの画像の位置合わせをするために NyARToolKit ライブラリを使用している．

2.1　入力

提案システムによる電子工作は，ユーザが素子の位置を設定するか，あるいはプログラムが指定する位置に素子を置いてから作業を開始する．

現在対応している素子は，7 セグ LED・マトリクス LED・フルカラー LED・豆電球の発光素子，発音素子などの出力素子がある．また，スイッチ，可変抵抗などの入力素子にも対応している．入力素子は，その位置のピクセルが肌色かどうかを識別して，指の有無を認識することで実装している．

電子部品の置かれているディスプレイをドラッグすることで線を引き，仮想的なジャンパ線で結線を行う．配線するには，素子の端子から端子へとドラッグし，断線するには，ジャンパ線を横切るようにドラッグする（図 2）．指を近づけると端子が拡大するなどの支援機能も搭載している．

2.2　出力

提案システムは，現時点では素子を設定通りに設置するためのガイドラインを表示している．これは

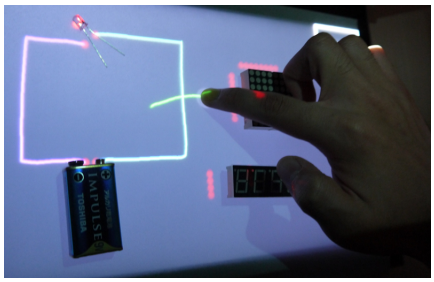


図 2: 仮想的な結線と断線操作

実行時には消すことができる．出力素子，特に発光素子にプロジェクションマッピングを行うことによって，図 2 のように，実際に通電せずに動作しているように見せることができる．仮想的な結線であるため，図 3 のような素子は袋から未開封のまま擬似動作させることも可能である．

LED 類は素子の設定位置に基づいて発光する箇所をプロジェクタで投影する．ブザーは，通電された時にパソコンからブザー音を出力することで実装した．バイブレータについてはブザーと同じくバイブレーションの動作音を鳴らし，かつ震えているようなアニメーションを周辺に表示することによって実現している．電池は現在 9V 角型電池と単 3 形電池の 2 種類に対応している．電池上にはプロジェクションマッピングを行わないが，プログラム内部で通電チェックをする起点となる．

また，配線されている箇所のペアを表示するジャンパ線を表示する．配線状況は回路行列で保存している．通電チェックアルゴリズムとしては，現在は電圧･電流･抵抗値などの計算はしておらず，閉回路の判定のみ行っている．電源素子のプラス極となっている頂点から深さ優先探索を行い，電源素子のマイナス極の頂点に辿り着いた時，それまで通った頂点に通電判定を出している．

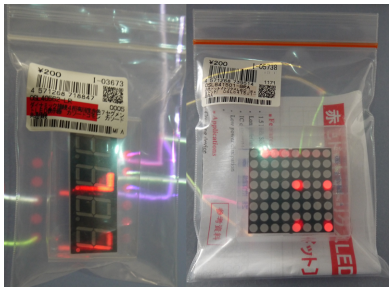


図 3: 未開封の 7 セグ LED とマトリクス LED へのプロジェクションマッピング

3 関連研究

フィジカルコンピューティングを容易にするアプローチとして Phidgets，Arduino，Gainer といったモジュールが広く使われている．Gainer はメディアアートにおける支援をひとつの目的としている．

他にも Visible Breadboard[1] や react3D Electric[2]，そして第二著者の HMMBB[3]，秋田の LED Tile 電子ブロックのように，AR 的に電流を可視化するシステムが挙げられる．

4 おわりに

このシステムは，教科書などの回路図と実際の回路制作とのギャップを埋めることができると考えている．また，このシステムを電子部品店に設置すれば，袋から未開封のまま部品の動作や仕様を確認できるため，部品選びをいままでより快適に行えるようになると考えている．また，複数人で回路を取り囲みながら，わいわいと作業を行うこともできる．

冒頭に述べた複数人での電子工作の可能性の他に，システムの改良予定を記す．まず，外部ソフトウェア，たとえば SPICE や Virtual BreadBoard などのシミュレータと連動してより正確な回路シミュレーションが行えるように改良を施したい．また，EAGLE や PasS などの CAD ソフトと連携し実体配線図などを出力することにより，ハンダ付けによる制作につなげることができると考えている．

現在，提案システムで構築できる回路は「配線に適切な抵抗値が設定されている」という前提で動作しているが，今後は抵抗器への対応，抵抗値の算出などが行えるようにしたいと考えている．これは SPICE や Virtual BreadBoard などの外部ソフトウェアとの連動により実装できると考える．

改良予定として，2 つの素子を選ぶと素子同士の理想的と思われる配線をする好意的解釈機能の実装がある．任意の 2 つの素子について，配線方法のテンプレートを事前に用意しておくことで，ユーザが大雑把な配線を行えるようにできると考えている．

現時点で提案システムが対応する素子や機能はまだ限られている．しかし，その応用可能性や，実在部品と仮想部品を混在させた電子工作の世界については，本ワークショップで大いに議論する価値あるものだと考えている．

参考文献

[1] 落合陽一.「電気がみえる」デバイス Visible Breadboard，日本バーチャルリアリティ学会論文誌，Vol.15，No.3，pp.463-466，2010.

[2] Frank Uhling. react3D Electric - Tangible User Interface. http://www.youtube.com/watch?v=6qTTQKWfV8Q

[3] 宮下芳明．フィジカルコンピューティングへのモチベーションを向上させるブレッドボード，夏のプログラミング・シンポジウム報告集，pp.1-4，2010.

VisualMimic: スクリーンショットを利用したGUI操作自動化のためのビジュアル開発環境

深堀 孔明　　坂本 大介　　五十嵐 健夫 *

概要. 本稿では，マウス操作やキー操作といったインタラクティブなアプリケーションに対する操作を自動化するためのスクリプト開発環境 Visual Mimic を提案する．本システムでは，ユーザは操作したい GUI 要素をスクリーンショット画像で指定することで，特定の API に依存しない GUI 操作を記述できる．具体的には，ユーザはまず自動化したい GUI 操作を例示により記述する．その後，生成されたスクリプトはグラフベースのビジュアルエディタ上に表示され，ユーザはドラッグ＆ドロップでスクリプトを編集する．結果として，ユーザは従来の開発環境に比べて少ない労力で GUI 操作を自動化することが可能になる．

1 はじめに

GUI アプリケーションや PC ゲームのようなインタラクティブなアプリケーションは，ユーザがマウス操作やキー操作といった GUI 操作を行うことで動作する．そのため，GUI アプリケーションの動作テストや，RPG ゲームにおけるキャラクタのレベル上げといった単調な作業を行う場合でも，常にユーザがパソコンを操作しつづける必要があり，ユーザに大きな負担がかかってしまう。Automator [1] のような GUI 操作を自動化するためのツールもあるが，専用の API を利用して各アプリケーションのインタフェースを操作することで実現しており，Flash ゲーム内のボタンのような，ユーザが独自に作成した GUI 要素への操作は記述できない．

一方，Sikuli [2] という GUI 操作の自動化ツールでは，操作したい GUI 要素をそのスクリーンショット画像で指定する．システムは指定された画像をデスクトップ画面から探し，マッチした領域に対して GUI 操作を実行する．このように GUI 要素の見た目を利用することで，任意の GUI 要素に対する操作を記述できる．しかし，Sikuli では各操作をテキストベースのスクリプトとして記述する必要があり，プログラミングが得意でないユーザにとって負担が大きいという問題がある．

そこで本研究では，Sikuli をエンドユーザ向けに改良した開発環境 Visual Mimic を提案する（図 1）．ユーザはまず、自動化したい GUI 操作を例示により記述する。生成されたスクリプトはグラフベースのビジュアルエディタ（以下キャンバスと呼ぶ）に表示され，ユーザはスクリプトをドラッグ＆ドロップで編集できる．本研究の最大の貢献は，例示プログラミングとビジュアルプログラミングを用いることで，ユーザが最小の労力で Sikuli と同様な GUI 操作の自動化スクリプトを作成できるようにしたことである．

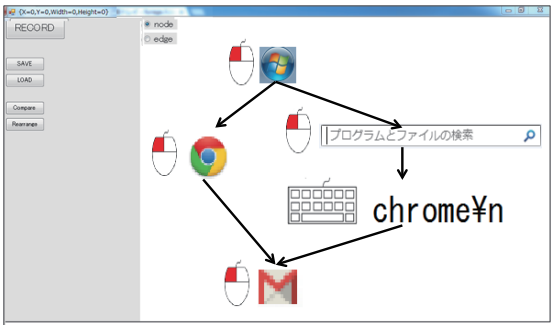


図 1. VisualMimic のインタフェース

2 Visual Mimic

ユーザが図 1 左上の録画ボタン（”RECORD”）を押すと，システムはバックグラウンドでユーザのマウス・キー操作の監視を始める．ユーザがマウスカーソルを動かすと，システムは現在マウスカーソルが接触している GUI 要素を自動で検出し，赤色の矩形でハイライトする（図 4-a 左）．ユーザがマウスクリックをすると，現在ハイライトされている GUI 要素をクリックする命令がキャンバス上に追加される（図 4-a 右）．同様に，ユーザが文字列をタイプすると，キー操作命令がキャンバスに追加される．

なお，マウスカーソルが接触している GUI 要素の検出には，デスクトップ画面の画素変化を利用している．マウスカーソルが動いた瞬間に，カーソル周辺で色が変化した領域を GUI 要素として識別する（図 2）．システムが望ましい領域をハイライトしない場合は，カーソルを大きく円状に動かすジェスチャによって，手動で領域を指定するモードへ移行することもできる（図 3）．

Copyright is held by the author(s).

* Koumei Fukahori, Daisuke Sakamoto, Takeo Igarashi, 東京大学



図 2. GUI 要素の自動検出アルゴリズム

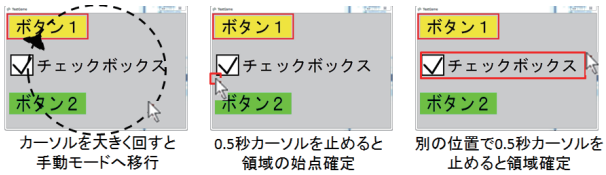


図 3. ハイライト領域の手動指定

例示により記述されたスクリプトはキャンバス上にグラフベースのビジュアル言語として表示される．各ノードはひとつの GUI 操作，エッジはスクリプトの処理の流れを表す．ユーザは任意のノードからスクリプトの実行を開始でき，システムはエッジにそってノードをたどり，対応する GUI 操作を順次実行する．例えば，図 4-b の一番上のノードが実行されると，システムはまずデスクトップ画面からボタン 1 の画像を探し，マッチした領域の中央にマウスカーソルを動かし，左クリックを行う．画像マッチングには OpenCV の *cvMatchTemplate()* 関数を利用しており，指定画像とのマッチ率が 80 % を超えた領域に対して各操作を行う．

ユーザはノード間にエッジを追加・削除することでスクリプトの処理の流れを編集できる（図 4-b）．枝分かれを作った場合，現在実行可能な GUI 操作の方へ処理が進む．例えば，ボタン 1 のクリック操作が終わったら，システムはデスクトップ画面からチェックボックスを探し，チェックがついている場合は左側のノードに，ついていない場合は右側のノードに処理が進む．また，後退辺を加えることで処理をループさせることができる．

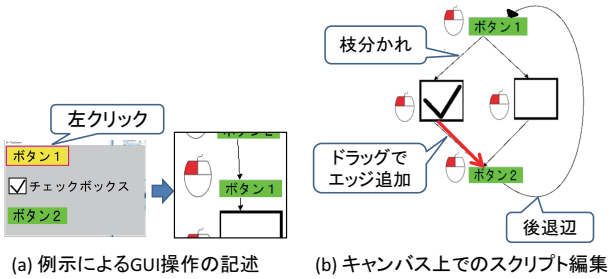


図 4. VisualMimic 上でのスクリプトの作成

3 利用例

図 5 は VisualMimic を用いて作成したスクリプトの例である．図 5-a のスクリプトでは，Google Chrome から GMail を開き，WISS の投稿システムから届いたメールをアーカイブする．また，図 5-b は YouTube 動画をリピート再生するスクリプトであり，動画の再生が終わるたびにリプレイボタンを自動でクリックする．これらのスクリプトでは処理の流れが矢印で表示されており，テキストベースのスクリプトと比べてスクリプトの構造が理解しやすくなっている．また，Sikuli のようにユーザがソースコードを書く必要がなく，比較的少ない手数でスクリプトを記述できる．例えば，図 5-b のスクリプトは 6 回のマウスクリックだけで作成できた．

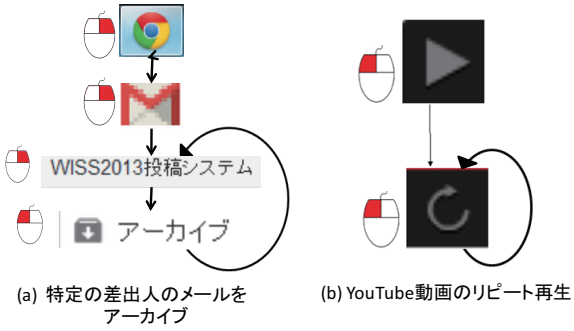


図 5. VisualMimic で作成したスクリプトの例

4 今後の課題

現状，本システムは変数を扱う仕組みを用意しておらずチューリング完全でない．したがって今後は，変数を扱うためのインタフェースをキャンバス上に実装する予定である．それに伴って for 文や if 文といった構文をキャンバス上で記述できるようにすることで，本システムの高機能化を進める．

また，本システムの GUI 要素の検出手法は完全ではない．現在の実装では，デスクトップ画面の色変化を利用して GUI 要素を識別するため，GUI 要素が色変化しない場合や，カーソルが動いた瞬間に GUI 要素でないオブジェクトがたまたま色変化した場合，本手法は正しく動作しない．今後は単純な色変化だけではなく，他のデスクトップ画面の画像特徴量も利用した GUI 要素の検出手法について検討していきたい．

参考文献

[1] Automator. http://support.apple.com/kb/HT2488.

[2] T. Yeh, T.-H. Chang, and R. C. Miller. Sikuli: using GUI screenshots for search and automation. In *Proc. UIST '09*, pp. 183–192, 2009.

災害対策本部における記録・文書管理システムの提案

仙道 航 * 瀬川 典久　澤本 潤　杉野 栄二 † 秋冨 慎司 ‡

概要.　2011 年 3 月に日本で発生した東日本大震災から約 2 年半が経過した。現在、日本では東日本大震災を受けて、将来に起こる災害に備えた検証を行うことが一つの課題となっている。東日本大震災のような大規模災害が発生した際に、様々な会議や組織が立ちあげられる。その中の一つに災害対応の要として運営される災害対策本部という組織がある。現在、災害対策本部では災害対応のための情報の集約は電話・無線通信・FAX、情報の共有はホワイトボードや展示用パネルへの紙文書の貼り付けによって行われている。災害対策本部では発災初動期から職員が息つく暇なく働いているため、紙ベースでの情報の記録をリアルタイムでデジタルに起こすことは難しい。我々は、災害対応中の記録が困難であるために検証のための記録が不足しているという問題を解決するために災害対策本部のための文書管理・追跡システムを開発した。

1　はじめに

現在、日本では東日本大震災での教訓を活かし将来に起こる災害に備えるための検証を行うことが一つの課題となっている。特に災害対応の要となる災害対策本部における意思決定や行動プロセス、運営に関する検証は将来に起こる災害のための検証として非常に重要なミッションであると考える。

現在、災害対策本部に関する検証を行っていくにあたって、議事録が未作成であったことが話題となっている [1]。このように、記録が不足している状況では、災害対策本部の指揮や下した命令が当時どのような情報を基に決定したのかが分からない。この問題に対し、当時の職員の振り返りによる議事録の作成を行っているが、人間の記憶に頼る物であり信頼性には欠ける。

災害対策本部における記録作業の障害の一つとして、発災初期（発災から 72 時間）の災害対応業務が非常に多いことが挙げられる。主な業務としては情報の集約と共有、関係機関との連携、集約した情報を基にした指揮と命令を下すことである。情報の集約は紙文書による FAX での送受信、共有は紙文書をホワイトボードやパネルに貼り付けることによって全体に共有したとすることが多い。災害対策本部で作成される紙文書や関係機関から集約される紙文書には災害対応のための判断に必要な被災状況に関する情報や、各機関の人員や災害対応に関する情報などが含まれる。

これらは、災害対策本部や関係機関の指揮命令に関わる判断材料としてだけでなく、当時得られた情報による状況判断が正しかったのかどうか、将来起こる災害に向けて検証を行っていくための重要な材料である。

しかし、災害対応に追われながら災害対策本部に集約された紙文書の変遷を記録することや作成した紙文書がどこまで伝わったかを全て記録することは困難である。また、紙文書は電子メールや電子的なファイルと異なり、コピー（ex. スキャナによるコピー）や送受信 (ex. FAX, 人手による配布) のログが文書を作成したコンピュータ上には残らず、後日、その軌跡を追うことは難しい。

こうした問題に対し、本研究では地震や津波といった大規模災害が発生した際に立ちあげられる災害対策本部のための文書管理・追跡システム（SID）を開発した。

2　システム概要

我々が開発した SID を災害対策本部に導入することで、紙文書が誰によってどのような情報を基に作成されたのかに関してコンピュータによって記録することが可能になる。また、紙文書に対しスキャナによるコピーや FAX による送信及び受信が行われても継続して追跡が可能になる。つまり、デジタルだけでなくアナログなインタラクションの記録が可能になる。

SID を災害対策本部に導入した際の紙文書に関する情報の流れを図 1 に簡潔に示す。

SID では、コンピュータ上で作成された (ex. Microsoft Office Word) 文書が印刷される際、プリンタに文書データが送信される前に文書に対し特殊なマーカ（QR コード）を付加する。この QR コードには作成者の情報や作成者の所属する組織の情報、文書が作成された時刻が埋め込まれている。印刷された紙文書に関する情報は印刷時点で外部サーバもしくはローカルサーバに送信されデータベースへと登録される。これにより、マーカを識別子として文書を一意に特定することができ、作成者や場所など

Copyright is held by the author(s).

* 岩手県立大学大学院 ソフトウェア情報学研究科
† 岩手県立大学 ソフトウェア情報学部
‡ 岩手医科大学

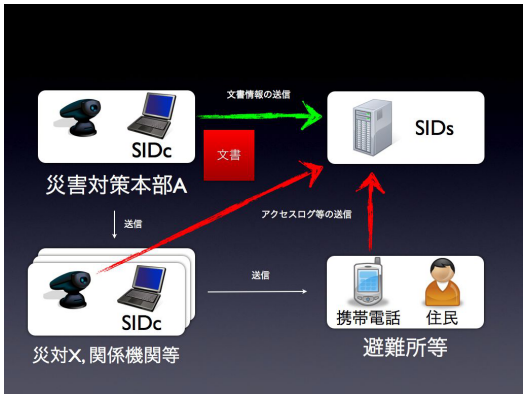


図 1. 紙文書に関する情報の流れ

の関連する情報の紐付けを行うことができる。

紙文書の作成者が所属する災害対策本部に貼り付けられた紙文書や他の災害対策本部や関連組織に送信され貼り付けられた紙文書を補足するために、災害対策本部内に全方位カメラを設置する。この全方位カメラによる映像から付与されたマーカの解析を行う。解析が終わるとサーバと通信を行い、紙文書がどの災害対策本部にいつ渡ったかといったようなログをサーバへと蓄積する。また、スマートフォンなどの QR コードリーダから QR コードを読み込んだ際には、該当する紙文書に関する情報を Web ブラウザ上で閲覧することができる。

このように、紙文書の作成時におけるそれに関わる情報と紙文書が共有されていくフローでの情報をそれぞれ蓄積することでこれまで困難であった記録と追跡を実現する。

3 システム実装

SID では、紙文書にマーカを付与する、サーバとクライアント間での各種情報のやりとり、Web ビューの開発などに主として PHP や PEAR などを用いている。

本稿では、特に SID による紙文書が作成され管理されるまでのフローについて図 2 を用いて詳しく述べる。SID を利用する災害対策本部職員（以下、ユーザ）は基本的に普段コンピュータ上で文書を作成する手順に加えて特別な操作を行うことはない。現在、文書を作成するためのアプリケーションとして MicroSoft Word for Mac 2011（以下、Word）をターゲットにしているため、Word を例として作成フローの一例について述べる。

ユーザは Word を使用し、文書を作成するとホワイトボード等へ貼り付けるための共有などを行うため印刷ボタンを押下する。SID ではこの際、文書データを直接プリンタへは送らずに一度 PDF として SID のワーキングディレクトリへと保存する。PDF には作成者やホスト、タイムスタンプを埋め込んだマーカを付与する処理やサーバへの文書情報の送信処理を行う。その後、通常どおりプリンタへと印刷指示をする。つまり、ユーザは普段通り印刷ボタンを押下しただけであるが、印刷された紙文書にはすでにマーカが付与され、紙文書に関する情報はすでに管理されているという状態になる。

次に、ユーザは文書をコンピュータ上に保存する。保存された際、SID では印刷された紙文書（コンピュータ上では PDF として扱っている）に関する情報を起点としてユーザが現在保存した文書（.Docx）を探索しサーバへと文書情報を送信する。探索の際には、文書内容の比較等を行っており、簡単な版管理のための処理を行っている。

サーバへと蓄積された情報は順次データベースへと登録される。文書の詳細や関連する文書等の関係を閲覧するための Web ページはデータベースに蓄積された情報を基に、新たな文書が管理（追加）される度に自動で生成される。

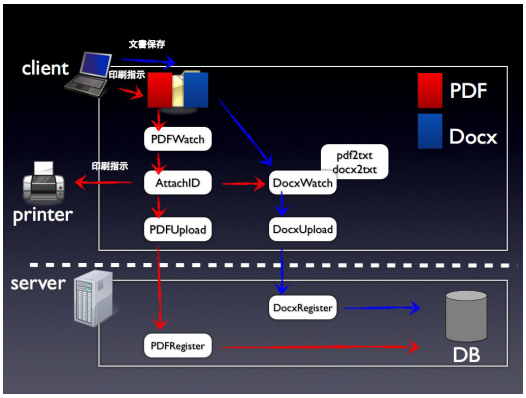


図 2. 紙文書の作成から管理までのフロー

4 おわりに

本稿では、災害対策本部における災害対応中の記録が難しいという問題に対し、紙文書にマーカを付与することによる記録・文書管理システムの提案を行った。

今後は、ユーザビリティを損なわない範囲での機能拡張とエラーハンドリングの強化を行う予定である。また、現在使用しているマーカは近距離用に使われるマーカであるため、今後長距離用のマーカの利用もしくは新規に開発のいずれかも行う予定である。

参考文献

[1] http://www.nikkei.com/article/DGXNASFS2700A_X20C12A1MM0000/?dg=1, 震災 10 会議で議事録なし ずさんな文書管理, 2012/1/27, 日本経済新聞

誰でも神プレイできるシューティングゲーム

簗瀬 洋平 *

概要. 本稿では，人間の感じる難易度と，実際のゲームの難易度の間には差がある事に着目し，実際の難易度に大きく影響しないやり方で見た目の難易度だけを変化させる2Dシューティングゲームのシステムを提案する．多くの人間は，ゲームを遊ぶ際に格好良いプレイをしたいと考えるものである．ゲームプレイの中でピンチを乗り切った際，あるいは自分が理想とする鮮やかな動きが出来た時，人はそこに喜びと達成感を感じ，ゲームを継続し，プレイに習熟していくための動機が得られる．しかし，この格好良い，うまくプレイ出来たという感覚は個人の能力や経験によって大きく異なる．そのため，上級者をターゲットにして難易度の調整を行えば初級者には難しすぎ，初級者をターゲットにすると上級者にはつまらないという状態になりがちである．このシステムを使用する事で，見た目の難易度を変えずに初級者と上級者，両方に対応する事が可能となる．

1 はじめに

過去，傑作と言われるゲームの中にはゼビウス[1]，スターソルジャー[2]など多くの２Dシューティングゲームが含まれている．2Dシューティングゲームは得点という一意の要素によって優劣が競われる点，画面上に存在する敵機の数，敵弾の数によってその難易度を客観的に見て取れる点などにより，プレイヤーの上手下手を見極めやすいゲームジャンルとなっており，シューティングゲームがうまいプレイヤーはゲームがうまいプレイヤーとして尊敬を勝ち得る事が出来た．

近年の2Dシューティングゲームにおいては，画面中に敵弾が放出され，それを紙一重で避けながらプレイを続ける様はWeb上の動画投稿や実況などで多くの視聴者を魅了するようになったものの，内容は高難易度化の一途を辿り，初心者のプレイヤーが気軽に挑戦してもすぐにゲームオーバーとなってしまうため，新規参入者が得られにくい．

そこで筆者は初心者でも上級者の様に多くの敵弾を避けながらプレイ出来るゲームシステムを考案し，デモ作成を行った．

2 ゲーム概要

2.1 ゲームデザイン

本稿のために作成したデモは，2Dシューティングと呼ばれるゲームジャンルの作品である．プレイヤーは立方体の自機を操作し，画面上に出現するカプセルを取る事で加点される．

カプセルを取るごとに画面上に出現する敵機の数が増え，出現した敵機は自機に向かって弾丸を発射する．この弾丸に自機が接触すると，自機は爆発し，ゲームオーバーとなる．ゲームオーバーとなるまでの間にどれだけの得点を取るかによって，成績の優劣を競う．

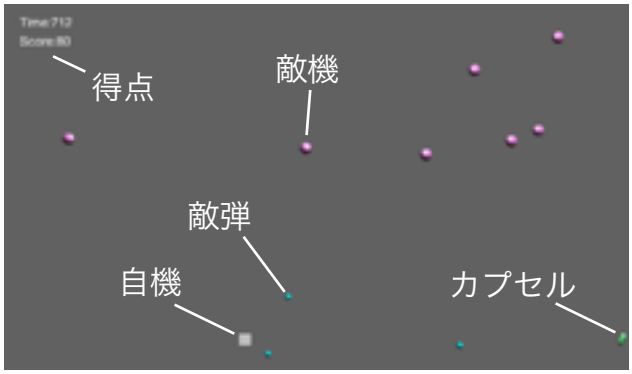


図 1. ゲーム画面

なお，一般的には上下左右に動け，自機から弾を発射して敵機を倒すデザインの2Dシューティングゲームの方が多い．本稿においては，PCでゲームをプレイする場合，上下の操作を入れるとキー配置の問題で初級者の操作難易度が上がってしまう点と，敵を倒すという要素を入れる事でプレイヤーの上手下手が判別しにくくなる点を考慮してこのような構成にしている．

2.2 ゲームシステム

敵機は自機が画面下方より出現し，わずかでもキーを入れた瞬間から出現するようになる．最初は1機のみ出現するが，カプセルを取得するごとに1機ずつ同時出現数が増えていく．

画面上方のランダムな位置に生成され，画面の左右，出現位置と反対の方向に向かって移動

Copyright is held by the author.

* Yohei Yanase, Unity Technologies Japan

を開始する．敵機は画面から消える前に一発だけ弾をプレイヤー方向に向かって発射する．
敵機が発射する弾は発射時点の自機の位置に向かって移動するタイプ1，自機の加速と移動方向を加味しそのままのキー操作を続けたら必ず当たるタイプ２，発射時点での加速速度の状態からは絶対に当たらない境界手前に撃つタイプ３の３種類である．

タイプ1の弾はプレイヤーを移動を移動させる消極的な動機として作用するが，このタイプの弾は一方向に移動し続ければ画面端での折り返し以外で接触する事がないため，避けるのが容易である．そこで，タイプ2の弾を使う意味が出て来る．タイプ2の弾は移動方向の先に向かって発射されるため，プレイヤーはその弾との接触位置を見極め，移動を中断する必要がある．タイプ1と2を混在させる事により，プレイヤーは移動と停止のどちらが適切かを判断し，ゲームプレイを行う必要が出て来る．

タイプ3の弾はどれだけ数が増えても難易度にはほとんど影響しない．よって，タイプ3の弾のみを増やしていく事により，実質の難易度を変えずに見た目の難易度を上げていく事が出来る．

3　デモの流れ

デモにおいては，まず体験者のゲームの習熟度について質問をする．この際，上級者と初級者が混じった集団であるとより効果的である．

まず，上級者に対してはタイプ1と2のみが発射され，得点の増加と共に敵機の数が増え，難易度が上昇していくゲームを遊んでもらう．多くのシューティングゲームでは，弾はプレイヤーの動きとは無関係に発射され，少数の弾だけがプレイヤーに向かって飛んでくるという構成であり，特にタイプ2のような偏差射撃が行われる事は非常に稀であるため，上級者でも比較的早くゲームオーバーになってしまう．

次に，初級者にほとんどがタイプ3で時々タイプ1や2が混じった弾を撃つ設定でゲームをプレイしてもらう．先の上級者のプレイを見ていると，ゲームが難しいという印象を受けるが，実際にはさほど難しいものではないため，順調に得点を重ね，上級者の得点を超える事も容易となる．しかし，プレイしている本人，それを見ている先のプレイヤー共にゲームそのものは同じに見えているため，何故そうなるのか疑問を抱く事になる．

4　問題点と今後の展望

本デモは現状実験を十分に行っていないため，タイプ1とタイプ2の弾の割合の変化による客観的，主観的な難易度の変化について定量的なデータを持っていない．また，一般的な2Dシューティングゲームのように上下左右に動け，敵を撃って倒せるようにした場合どのような変化があるのかという点についても予測がついていない．まずは本デモについての検証実験を行い，より実践的なゲームに活かせるよう知見を深めていくものとして位置づけている．

参考文献

[1] NAMCO. XEBIOUS, 1983

[2] Hudson Soft. Star Soldier, 1986

未来ビジョン

ゲーム制作には様々なHCI的知見が使われているが，その多くは膨大なシステムの中に埋もれており，ごく自然なゲーム体験の一部として受け入れられているためユーザーはもちろんゲームを研究している人間にすら知られていない．本研究はそれらを解き明かし，知見として共有するための一助とするべく行われている．

また，ゲームシステムや要素の提案を研究者が行う際，そのゲームが十分に良く出来ていない事により，本業の開発者に響かない，見向きもされないという事例も散見されるが，実際の制作過程で使われている様々な技術や必須の要素などを明らかにしていく事で，最低限の面白さを備えたデモの展示を実現し，デジタルゲームにおける産学の連携をより推進していきたいという思いもある．

本研究で作成したUnityプロジェクトは今後のゲームあるいはゲームの要素を研究していく方々が利用出来るよう，誰でもダウンロード出来るよう公開し，研究目的でも改変，再配布などを可能とする予定である．

叩打音を利用した操作インタフェースと浴槽への適用

伊藤 大毅* 平井 重行†

概要. 手指で「叩く」操作を元にしたインタラクションを浴槽に適用し，その叩打位置（叩く場所）や叩打音色（叩き方），叩打パターン（叩くリズム）の組み合わせに対応した操作を実現するシステムを提案する．それらの方法により，様々な機器やシステムの操作を行うほか，叩くこと自体をエンタテインメントとすることまでを考えている．本稿では主に叩打位置の検出処理について述べる．

1 はじめに

日常生活において快適さの追求だけでなく，エンタテインメント要素まで含めた QoL(Quality of Life) 向上を目指した住環境の研究が行われている．我々はこれまで，住環境の中でも特に水場である浴室での操作インタフェースと応用システムの研究を行ってきた．静電容量方式タッチセンサを浴槽裏側に設置して「触れる」インタフェースを構成する TubTouch[1] では，給湯器や照明などの機器，音楽プレーヤアプリの操作を浴槽縁で行うシステムを実現している．また，Bathcratch[2] では，浴槽裏側に設置したピエゾセンサで浴槽のこする音を検出することにより「こする」インタフェースを構成し，DJ スクラッチ演奏を行うアプリケーションを実現している．

本研究では，これらの操作に加え，浴槽を「叩く」ことによる操作インタフェースを提案する．ここでは，叩打位置（叩いた場所）と叩打音色（叩き方），さらに叩打パターン（叩くリズム）を利用し，浴室内の様々な機器やシステムの操作，さらに新たなアプリケーションの実現を含めて，日常生活環境に埋め込まれた入力インタフェースの活用を目指す．

2 システム概要

2.1 システム構成

日本国内で一般に普及しているシステムバスの浴槽において，その縁の裏側にピエゾセンサを取り付け，浴槽縁の叩打を元に様々な入力手法を実現する．浴槽叩打音を浴槽の個体振動としてピエゾセンサで直接計測し，その音響信号をリアルタイム処理することで操作イベントを出力するシステム構成とする（図 1）．ここでは，叩打音の音響信号に対し，「叩打位置」の検出処理（複数のセンサの信号を利用），「叩打音色」の認識処理を行い，それぞれの情報に基づいて叩打パターンの認識処理を行う．以下，本稿においては主に「叩打位置」検出について述べる．

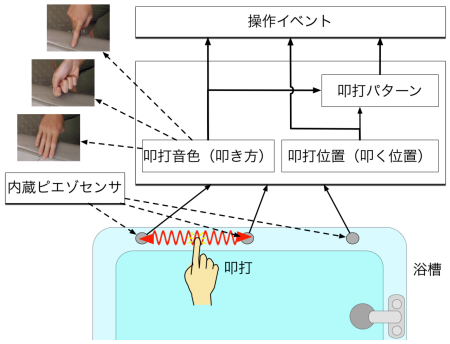


図 1. 浴槽インタラクションシステム構成

2.2 叩打位置検出手法

本システムでは浴槽裏側に設置したピエゾセンサによって叩打位置検出を行う．位置検出の仕組みとしては，センサ数の増加に伴う設置困難さと計算負荷の増加を考慮して，なるべく少ないセンサ数で実現できる方が良いと考える．このことから，ピエゾセンサを多数連続的に並べるのではなく，一定距離で離して数個設置し，叩いた振動がセンサに伝搬するまでの時間差を利用して叩打位置を算出する方法 [3] をとることにする．

2.3 初動検出処理

時間差による位置検出を実現するためには，叩打による振動がセンサへ到達する時間を正確に求める必要がある．そのため，適切な閾値を設定して叩打音の振幅がそれを超えた時刻を初動（振動の到達時刻）として検出する．ただ，閾値によってはわずかな叩き方の違いで振幅の変化の仕方が著しく変わり，検出時刻のばらつきが大きくなり十分な検出精度が得られない．そこでなるべく適切な処理を行うため，二乗平均平方根（RMS）により積分し，波形の時間変化をある程度平滑化して閾値処理する手法を用いる．これにより振動が増加し始める瞬間のみを的確に検出できる．

Copyright is held by the author(s).
* 京都産業大学大学院 先端情報学研究科
† 京都産業大学 コンピュータ理工学部

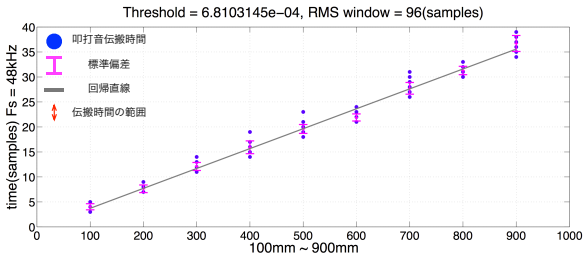


図 2. 各叩打位置の伝搬時間の標準偏差と回帰直線

3 叩打音位置検出のための振動解析

叩打音の初動を検出する適切な閾値は未知であり，ピエゾセンサの設置方法によってもかなり変化することが想定できる．また RMS 処理の際，波形を切りだす矩形窓の時間長も未知である．これら閾値と RMS 窓幅の長さの適切な値は，予め収録した叩打音データを解析することで求めた．具体的には，浴槽を複数回叩打した音データを用意し，閾値と RMS 窓幅を変化させながら，初動検出処理によって叩打位置からセンサに振動が伝搬するまでの時間を求める．この伝播時間の標準偏差が最小となる閾値と RMS 窓幅の値を適切な値とする．図 2 は，FRP の浴槽を指先で叩打した際の叩打音から求めた閾値と RMS 窓幅での各叩打位置の振動伝搬時間をプロットしたものである．この図は，叩打位置がセンサから 100〜300mm 離れた位置ならば伝搬時間が重複しておらず，その範囲までなら叩打位置が判別出来ることを示している．この結果から，適切な閾値，RMS 窓幅，位置検出可能な距離の範囲（必要なセンサ数）を求められる．この解析手法と結果の利用は，新たに適用する浴槽環境に対するキャリブレーション手法として利用できる．

4 叩打音を利用するアプリケーション

本システムの応用例として，叩打位置と叩打音色に対応したサウンドを鳴らすエンタテインメントシステム”BathDrum”を提案して実装している [4]．この BathDrum は現時点では叩打位置に合わせて単に音を鳴らすだけであるが，TubTouch や Bathcratch と同様に，浴室の天井裏設置のプロジェクタで浴槽縁へ画面投影することで，より様々な応用や楽しみ方も実現できると言える．例えば，もぐらたたきゲーム，音楽に合わせて浴槽を叩くリズムゲームの表示などが考えられ，これに叩打音色や叩打パターンの認識も含めた処理を考えると応用も拡がる．また，TubTouch システムではメニュー表示の On/Off を切り替える操作について，浴槽の特定位置を 2 度タップすることでユーザが明示的に操作するなど，他の操作手法と組み合わせる効果的な利用方法も考えられる．

5 おわりに

本論文では，浴槽を「叩く」インタフェースを提案し，そのシステム概要と叩打位置検出処理について述べた．現状では指先での叩打にのみ対応した処理であるため，手のひらや拳での叩打に対応させる予定である．また，検出精度向上のために周波数解析や機械学習などの処理も検討し，処理の反応時間も調査する必要がある．そして，それらは叩打音色に関する処理も合わせて行うことを検討していく．叩打音色の区別については，Lopes らのシステム [5] や Harrison らのシステム [6] で実装されているが，それらは多少のタイムラグが発生する手法のため瞬時の反応を必要とする処理にはあまり向いていない，一方で PosessingDrums[7] ではリアルタイムの叩打音色の変換を行っており，BathDrum のような瞬時の反応を必要とするシステムに応用できると考えている．それらの叩打音色の処理に加え，叩打パターン認識処理も合わせて応用できるようにしていく予定である．

謝辞

本研究の一部は，中山隼雄科学技術文化財団の研究助成によって行われました．

参考文献

[1] 榊原吉伸, 林宏憲, 平井重行,TubTouch: 湯水の影響や自由 形状への適用を考慮した浴槽タッチ UI 環境, 情報処理学会論文誌 Vol.53, No.4, 2013.

[2] Shigeyuki Hirai, Yoshinobu Sakakibara and Seiho Hayakawa. Bathcratch: Touch and Sound-Based DJ Controller Implemented on a Bathtub, Proc. of ACE2012, pp.44-56, 2012.

[3] Joseph A. Paradiso, Che King Leo, Nisha Checka, Kaijen Hsiao. Passive acoustic sensing for tracking knocks atop large interactive display, Proc. of the 2002 IEEE International Conference on Sensors, pp.512-527, 2002.

[4] 伊藤大毅, 平井重行: 浴槽叩打音を利用したお風呂ドラム BathDrum の叩打位置検出, 情処研究報告 2013-EC-29-1, 2013.

[5] Pedro Lopes, Ricardo Jota, Joaquim A. Jorge. Augmenting touch interaction through acoustic sensing, Proc. of ITS2011, pp.53-56, 2011.

[6] Chris Harrison, Julia Schwarz and Scott E. Hudson. TapSense: Enhancing Finger Interaction on Touch Surfaces. Proc. of UIST ’11, pp.627-636, 2011.

[7] Kazuhiko Yamamoto. Possessing Drums: An Interface of Musical Instruments that Assigns Arbitrary Timbres to Personal Belongings, Journal of Information Processing, Vol.21, No.2, pp.274-282, 2013.

住宅内の機器・サービスを統合するためのビジュアルプログラミング環境

橋岡 良* 平井 重行†

概要. 家電や様々なセンサ，インターネットサービス（天気予報やニュース，SNS など），コンテンツまでを様々に組み合わせて統合的に連携・動作させることを目的とした機器やサービスが次々に登場し，住宅のスマートハウス化を進めている．本研究は，その流れにおいてエンドユーザが様々な機器やサービスをより柔軟で自由に組み合わせられるようにすべく，直感的でわかりやすい操作を実現するビジュアルプログラミング環境を開発し，その内部処理を行うためのシステムも構築している．本稿では，そのシステム概要とビジュアルプログラミング環境について述べる．

1 はじめに

家電機器を中心に複数の機器を連携させるような統合制御を行うためのプログラミング手法およびそのユーザインタフェースの研究がある [1][2][3]．最近では，エンドユーザが住宅内で扱う機器はセンサ類を含んだ家電や住宅設備に留まらず，スマートフォンの普及に伴って様々なインターネットサービス (例えばニュースや天気予報，SNS など) までもが操作対象となりつつあり，幅広い統合や柔軟な連携を実現する必要が出てきている．一方で，現時点で住宅内の家電製品などはあまり対象でないが，ネットワーク上のサービス同士やガジェットなどとの連携をルール記述の形式でエンドユーザが作成し，ある種のプログラミングが行える IFTTT[1] や Zapier[2] といったサービスがインターネット上で提供されるに至っている．本研究は，このサービスや機器を連携させて住宅のスマート化を進めるサービスに対し，エンドユーザがより自由自在に動作を設定・連携させることができるビジュアルプログラミング環境とその動作を支えるシステムの開発を行っている．以下，本稿ではそのシステム概要と，ビジュアルプログラミング環境について述べ，どのようなプログラムが作成可能か例を挙げながら説明する．

2 システム概要

本研究システムはエンドユーザが利用するクライアント端末上で動作するビジュアルプログラミング環境と，アプリケーションサーバとで主に構成される．アプリケーションサーバは住宅内の様々な機器・センサ類，またインターネット上の様々なサービスとの間で情報のやりとりを行う (図 1 参照)．エンドユーザは PC やスマートフォン，タブレット機器などの端末を用いて，機器やサービスの連携をデータフロー型のビジュアルプログラミング環境でプログラムを作成し，そのプログラムは XML もしくは JSON 形式で WebSocket を介してアプリケーションサーバへ送信，データベースに保存される．アプリケーションサーバ側は，機器やサービスを統合制御するプログラムが動作しており，プログラムに従って状況に応じた機器制御やサービスを連係動作させる仕組みになっている．

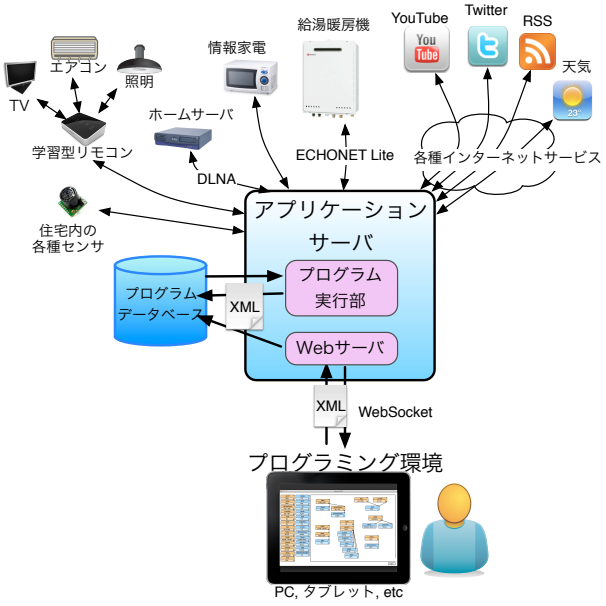


図 1. システム概要図

3 プログラミング環境

プログラミング環境は，Max/MSP や PureData, vvvv などのデータフロー型ビジュアルプログラミング環境として構成している．現時点で実装出来ているビジュアルプログラミング環境の外観は図 2 に

Copyright is held by the author(s).
* 京都産業大学大学院 先端情報学研究科
† 京都産業大学 コンピュータ理工学部
[1] IFTTT : https://ifttt.com/
[2] Zapier : https://zapier.com/

示す．その画面内部では，左側に機器や設備，サービスなどを表すオブジェクトが表示され，右側にプログラミングを行うプログラムフィールドがある．ここでは，オレンジ色のオブジェクトは，温度や扉の開閉など環境情報を知らせる機能，人の行動センシングに基づいて動作する機能，時間に伴って処理を行う機能などがあり，総じて入力オブジェクトと呼んでいる．水色のオブジェクトは，操作する機器や情報提示出力する機器・サービスを表す出力オブジェクトを表している．

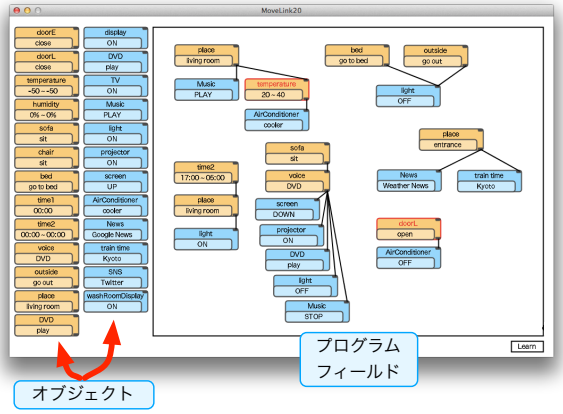


図 2. プログラミング環境概要図

ユーザは左側のオブジェクトリストからドラッグアンドドロップでプログラムフィールドへオブジェクトを配置し，オブジェクトのアウトレット (右下) と別オブジェクトのインレット (右上) との間をパッチコード (線) で接続することでプログラムを作成する．このパッチーコード間は，基本的にトリガー情報が流れることになり，そのトリガーを発する条件は，各オブジェクトの左下エリアに記載された内容に応じるものとなっている．各オブジェクトは図3のようにスライドバーで条件設定を行うことができる．なお，機器やサービスがオブジェクトの設定条件に合致した状態になっている場合には，オブジェクトの外枠と文字が赤くなる (図 4 参照)．

次に，プログラミング環境で作成したプログラム例を図 4 に示す．

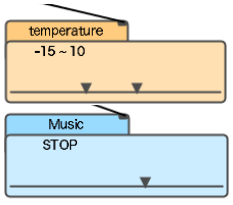


図 3. オブジェクトの条件設定例

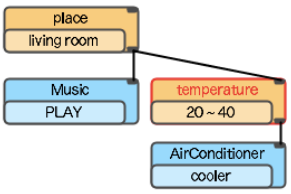


図 4. プログラム例

図 4 では，複数の入力オブジェクトと出力オブジェクトが接続され，複数の処理が一つにまとめられている．place オブジェクトと temperature オブジェクトがそれぞれ人の検出と設定温度範囲の条件判定を処理する入力オブジェクトである．ここでは place オブジェクトでリビングルームに人が居るか判定し，その結果に応じて接続された Music オブジェクトが音楽を再生する．一方で，place オブジェクトから temperature オブジェクトへ接続され，さらに AirConditioner オブジェクトへ接続された部分では，人がリビングルームにいると共に室温が 20〜40 度の範囲にある場合に冷房を ON にする，という処理を意味する．

これらエンドユーザが記述したプログラムは，プログラム作成者のユーザ ID と各プログラムに割り当てられた UUID と共に XML もしくは JSON のデータとしてサーバに送信し保存される．

4 まとめ

本研究は，スマートハウス研究の 1 つとして，家電や住宅設備などの既存機器や，様々なインターネットサービスを連携，統合的に制御するシステムと，そのためのエンドユーザが用いるプログラミング環境について扱っている．現時点は，ビジュアルプログラミング環境とバックエンドでプログラムを実行するサーバが一体化した形で，実際に幾つかのセンサや家電機器を統合制御でき，プログラミングも実際に行える状況にある．今後は，プログラムをサーバ側で行えるようにする．また，プログラミング環境については，操作方法や見た目のデザインについて改良を加えて，よりユーザが理解，操作しやすい形にし，PC だけでなくタブレット端末やスマートフォンでもプログラミングが行えるものにしていく予定である．

参考文献

[1] Masui Toshiyuki, Itiro Siio. Real-world graphical user interfaces. In Handheld and Ubiquitous Computing. Springer Berlin Heidelberg pp72-84, 2000.

[2] Kentaro Fukuchi, Takeo Igarashi, Maki Sugimoto, Charith Fernando, Masahiko Inami. Push-pins: Design-by-user Approach to Home Automation Programming. International Conference on Ubiquitous Computing, workshop 5, 2009.

[3] Shigeru Owada, Fumiaki Tokuhisa. Kadecot: HTML5-based visual novels development system for smart homes. Consumer Electronics, 2012 IEEE 1st Global Conference on, pp17-19, 2012.

1 つのセンサを用いたマルチタッチジェスチャセンサ

真鍋 宏幸*

概要. 1 つのセンサを用いた小型で安価なマルチタッチジェスチャセンサを提案する．検出対象ジェスチャはタッチとスワイプであり，さらにスワイプでは指の本数と方向を識別する．フォトリフレクタを用いた実装を行い，有効性を確認した．

1 はじめに

複数の指を用いることで，入力の自由度を増加させ，複雑な入力を素早く行えるマルチタッチジェスチャは，様々なデバイスへの入力手法として広く利用されている．タッチ検出面が広いほど，多くのジェスチャを組み込むことができるが，検出面が小さい場合であっても，マルチタッチジェスチャは有効である．例えば，マルチタッチジェスチャ入力を実装した 1.5 インチほどのディスプレイを備えた音楽プレーヤや，ツルの部分でジェスチャを行うメガネ型デバイスがすでに市販されている．通常，多点検出を行うために多数のセンサがアレイ状に並べられており，タッチ検出面上の任意の場所でジェスチャを行えるようになっている．一方，入力領域が小さいもしくは細長い場合など，ジェスチャが特定の場所で行われる時には，多数のセンサは必要ない．本稿ではこの点に着目し，1 つのセンサでマルチタッチジェスチャを検出する手法を提案し，フォトリフレクタを用いた実装を行った結果を報告する．

フォトリフレクタは小型かつ安価であり入手も容易なセンサであり，それを用いた入力手法の実装が数多く行われてきた．1 つのフォトリフレクタを用いた例として，近接センサ [2] やタッチセンサ [4] としての応用などがある．また，複数のフォトリフレクタを用いることで，より高度なインタラクションを目指した研究もある．例えば，複数の指を用いたインタラクション [1] や，指輪型デバイス [3]，柔軟物体への接触検出 [5] などがある．

2 提案手法

1 つのセンサを用いて，タッチとスワイプ，さらにスワイプを行った指の本数と方向を識別する手法を提案する．指と指を強く接触させた場合であっても，両指とタッチ検出面の間にはわずかな空間ができる．その間隙を検出することができれば，指の本数を数えることができる．また，通常の方向検出は複数のセンサを用いて実装されるのに対し，センサ感度をスワイプ方向に沿って変化させることで，1 つのセンサを用いた場合でも方向検出が可能となる．提案手法は様々なセンサに適用可能であると考えられるが，本稿では安価で実装が容易なフォトリフレクタへの適用を試みた．

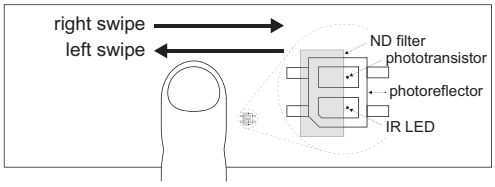


図 1. 製作したプロトタイプ．フォトリフレクタの上面半分には ND フィルタが設置されている.

2.1 実装

フォトリフレクタとして，検出距離が数 mm の GENIXTEK TPR-105（2.7 x 3.2mm）を用いてプロトタイプを製作した．検出距離が短いセンサを選定したのは，指と指とのわずかな間隙を検出するためである．また，スワイプ方向検出を目的としてセンサ感度に変化をもたらすために，フォトリフレクタの上面半分に ND フィルタを設置した．さらにその上面に，透明の樹脂板，指すべりを滑らかにするためのフィルムを貼付した．フォトリフレクタの出力はオペアンプで増幅した後に，5kHz でサンプリングし，PC で処理を行った．図 1 に製作したプロトタイプを示す．また，センサ出力を認識し，結果を表示する機能も加えた．

2.2 実験

プロトタイプを用いて，タッチ，1/2/3 本指スワイプを行った時の計測結果を図 2 にまとめて示す．タッチ時には比較的長時間の反応が観測されるのに対し，スワイプ時には短時間のピークが指の本数分だけ観測される．タップを行う場合，指先がセンサ表面に接触したことを知覚した後に，動作変更を行う必要があるため，指がセンサに触れている時間が長くなる．一方，スワイプを行う場合には，一連の動作は滑らかに行われるため，指がセンサに触れて

Copyright is held by the author(s).

* NTT ドコモ 先進技術研究所

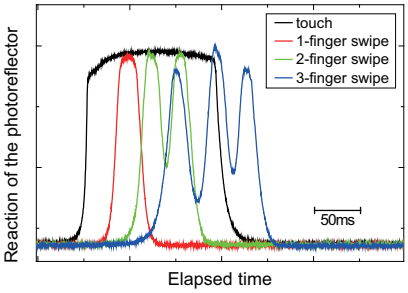


図 2. フォトリフレクタの出力例．ピークの数が指の本数に対応する．

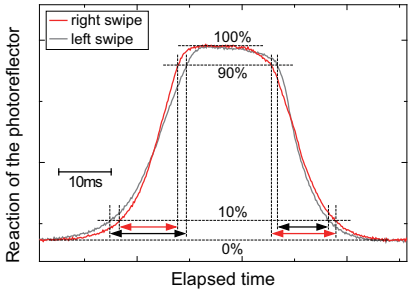


図 3. スワイプ方向の検出．スワイプの方向により傾きが変化する．

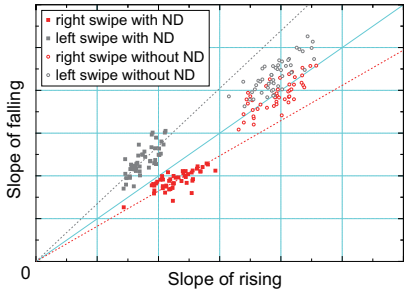


図 4. ND フィルタの効果．適用することで方向検出が可能となる．

いる時間は，スワイプの速さに依存することになる．そのため，極端に遅いスワイプを除き，センサに指が振れている時間はスワイプの方が短くなる．また，検出距離の短いフォトリフレクタを用いているため，指と指の間にできる間隙による反射光量の減少を明確に捉えることができている．

図 3 に，1 本指での右/左スワイプを行った時の計測結果を示す．右スワイプ時には左スワイプ時と比較して，急峻な立ち上がりと緩やかな立ち下がりが観測される．この違いは，次のように生じると考えられる．右スワイプを行う場合，指は ND フィルタが設置された方向からセンサに接近する．指が ND フィルタ上にある時点では，照射光と反射光の両方が減衰するため，センサの出力値は小さい．この状態はしばらく続くが，さらに指が進むと，ND フィルタで減衰されない強い照射光が指にあたるようになる．この段階で，反射光も減衰されなくなるため，出力値が大きくなる．指が完全にフォトリフレクタを覆っている間，センサの出力値はほぼ一定となる．さらに進むと，センサの出力は低下していくが，ND フィルタによる光の減衰がないため，指とセンサの距離がある程度離れてもセンサ出力は高くなる．結果として，出力の低下は緩やかなものとなる．

また，1 本指による左右のスワイプをそれぞれ 50 回行った時の，ピークの傾きの分布を図 4 に示す．立ち上がり/立ち下がりの傾きを横/縦軸にして，全てのスワイプをプロットした．なお，ピークの傾きは，図 3 に示すように，変化量が 10-90%となる区間に対して算出した．図 4 には，ND フィルタを取り外した時，つまりフォトリフレクタをそのまま用いた時，の結果も示した．ND フィルタがない場合でも，左右のスワイプ間にわずかな違いが見られるが，両者を正しく認識することは難しい．一方，ND フィルタを設置した場合には，その違いが拡大し，両者の分布は明確に分離されている．なお，ND フィルタにより，反射光量が減るため，ピーク値が小さくなり，傾きも小さくなる．また，図中の水色の実線は，2 つの傾きが同じになる点を結んでおり，右スワイプ時には立ち上がりの方が立ち下がりよりも傾きが大きいと単純化することもできる．

3 評価

4 人の被験者が，7 つのジェスチャ（タッチ，右/左方向の 1/2/3 本指スワイプ）をそれぞれ 10 回ずつ行った．次に行うジェスチャはランダムに指示される．全 280 試行で，7 回の false negative，2 回のジェスチャのカテゴリ誤り（3 本指スワイプを 2 本指，1 本指スワイプをタッチと誤認識），51 回の方向誤りがあり，認識率は 79%であった．スワイプ方向の検出を行わない場合の認識率は 97%，1 本指スワイプの方向正解率は 3 名の被験者で 100%，4 名全員でも 90%を超える結果が得られた．

複数指でのスワイプにおける方向検出の認識精度向上が必要であると考えられた．

4 まとめ

マルチタッチジェスチャを 1 つのセンサで検出する手法を提案し，フォトリフレクタをセンサに用いた実験により，有効性を確認した．今後，認識精度の向上，静電容量式のタッチセンサなど他のセンサへの適用に取り組んでいきたい．

参考文献

[1] A. Butler, S. Izadi, and S. Hodges. SideSight: multi-"touch" interaction around small devices. In *Proc. UIST '08*, pp. 201–204, 2008.

[2] K. Hinckley, J. Pierce, M. Sinclair, and E. Horvitz. Sensing techniques for mobile interaction. In *Proc. UIST '00*, pp. 91–100, 2000.

[3] M. Ogata, Y. Sugiura, H. Osawa, and M. Imai. iRing: intelligent ring using infrared reflection. In *Proc UIST '12*, pp. 131–136, 2012.

[4] V. Rana, N. Paliwal, and A. Chahar. Touch Sensor Assembly Using Infrared Radiations and its Applications. In *Proc. CICSyN '09*, pp. 96–98, 2009.

[5] Y. Sugiura, G. Kakehi, A. Withana, C. Lee, D. Sakamoto, M. Sugimoto, M. Inami, and T. Igarashi. Detecting shape deformation of soft objects using directional photoreflectivity measurement. In *Proc. UIST '11*, pp. 509–516, 2011.

ライフログに基づく音楽推薦システムMALLとその検証実験

宇野 愛　　伊藤 貴之*

概要. 聞きたい曲を選ぶ際に，メタデータから選ぶだけでなく，状況や環境（例えば，天気や季節，時間帯，場所など）に合った曲を選びたいときがある．これらの状況や環境の多くはライフログデータとして記録可能である．一方，我々は日常的に膨大な楽曲を所持していることが多く，状況にあった楽曲をその都度選ぶのは困難である．そこで我々は，ライフログデータから個人の選曲の傾向を把握し，それに従って楽曲を自動推薦する MALL (Music Adviser with Life Log) を提案している．MALL では楽曲再生時のライフログ情報と楽曲特徴量を照合し，その相関性ルールを導出する．そして様々な状況に応じて相関性ルールに該当する楽曲を推薦する．本報告では MALL に関する検証実験結果を報告する．

1 はじめに

音楽を選曲するときに多くの人は，メタデータを読んでいる．それに対して，その場の状況や環境に適合した楽曲を選びたい，という状況もよくあると考えられる．しかし一方で，ポータブル音楽プレイヤー等の普及，それらのメモリの大容量化に伴い，個人が持ち歩く楽曲の数は増大している．それにより，鑑賞時の状況に合わせた楽曲のみを手早く探し出すのは必ずしも簡単ではないと考えられる．

そこで我々は，状況や環境に適合した楽曲を自動推薦するシステムがあれば便利であると考える．状況や環境の変化に作用するものの多くは，ライフログデータとして記録することが可能である．そこで我々は，ライフログから選曲の傾向を把握し，楽曲を自動推薦する MALL を提案している．このシステムの実現により我々は，状況に合わせた楽曲の自動推薦はもちろん，今まで気づかなかった選曲傾向の把握も可能になると期待する．本報告では，MALL を使ったユーザテスト結果を示し，推薦結果の妥当性について検証する．

2 MALLによる楽曲推薦

2.1 楽曲特徴量抽出

我々の実装では楽曲特徴量抽出ソフトウェア MIRtoolbox[0] を用いて，各楽曲に対して RMS energy(音量の平均値)，Tempo，Rolloff(全体の 85%を占める低音域の割合)，Brightness(高音域の割合)，Roughness(不協和音の多さを示す値)，Spectral irregularity(曲の変化の大きさ)，Inharmonicity(根音に従っていない音の量)，Mode(major と minor の音量の差) の 8 つの特徴量を抽出している．

Copyright is held by the author(s).

* Ai Uno and Takayuki Itoh, お茶の水女子大学大学院 人間文化創成科学研究科 理学専攻

2.2 ライフログデータ収集

MALL では，ポータブル音楽プレイヤー等で音楽を鑑賞している人が，その曲が現在の状況に合っていると感じたら所定のボタンを押し，システムはその時点でのライフログ情報と一緒にその曲を記録する，というような処理手順を想定する．このアプリケーションの動作例を図に示す．図の右下にある鉛筆のボタンがライフログ情報を記録するためのボタンである．このボタンを押すことによって，日付，曜日，時間，その時聞いていた曲名，その曲のアーティスト名を記録することができる．このように MALL では，音楽を聞きながらボタンを押すという簡単な操作のみでライフログ情報を記録する．

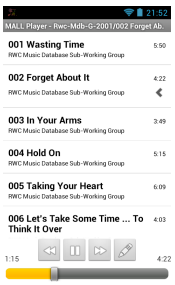


図 1. Android アプリケーションの動作例

2.3 相関性ルールの発見

続いて MALL では，2.1 節，2.2 節で収集した楽曲情報とライフログ情報との間の相関性ルール ($A \rightarrow B$) を求める．具体的には以下の式 (1) で算出する．ここで，A は時間帯や季節といったライフログに関する条件式を，B は楽曲特徴量の値の範囲を表す．相関性ルールを決めるにあたって二つの数値，支持度と確信度をもとに計算する．

支持度は, ライフログ条件と楽曲特徴量の全ての組み合わせの中から A，B を含むルールが登場する

頻度を $P(A,B)$ の確率で算出する．また支持度だけでは値の信頼性が低いため確信度を設ける．確信度は $P(B|A)$ で算出する．そして，支持度と確信度の双方が高い A,B の組み合わせを相関性ルールとして選出し，そのルールに該当する楽曲をライフログ条件ごとに推薦する．推薦された楽曲群はそれぞれライフログ条件ごとにプレイリストとしてユーザに返される．現時点での我々の実装では経験的に，支持度 0.05 以上，確信度 0.15 以上としている．

$$A \rightarrow B_1[a_{1min}, a_{1max}] \& B_2[a_{2min}, a_{2max}] \cdots B_n[a_{nmin}, a_{nmax}] \quad (1)$$

3 ユーザテスト

我々は JDK1.6.0，Android SDK 2.3.3，ADT プラグインを用いて音楽プレイヤーを実装し，SONY ポータブル音楽プレイヤー NW-Z1050 を用いて実行した．また JDK 1.6 を用いて相関性ルール導出部分を実装した．そして我々は，女子大学生 2 名を被験者としてユーザテストを実施した．被験者には一ヶ月間音楽プレイヤーを所持してもらい，プレイヤーに入っている音楽を聞き，その楽曲が状況・環境に合っている際に所定のボタンを押してもらった．楽曲データは，RWC 研究用音楽データベース [0] に収録されていた 313 曲を使用した．

以下，被験者 2 名のライフログデータをもとに MALL によって推薦された楽曲の主観評価結果について述べる．被験者にそれぞれのライフログデータから推薦されたプレイリストの楽曲を聞いてもらい，それらがライフログ条件に合っているかどうかを 5 段階で評価してもらった．評価結果とコメントは以下のとおりである．

ユーザ A に提示したプレイリスト

- 夜：5 夜に好んで聞く静かな曲が多かった．
- 水曜：4 好きな曲が多かったが，水曜特有の傾向は意識していなかった．
- 晴れ：4 晴れの日に好んで聞く明るい曲やさわやかな曲が多かった．

表 1. ユーザ B の“ 晴れ ”に合致した相関性ルール

ライフログ条件	楽曲特徴量
Sunny	RMSEnergy [0.03, 0.05]
	Roughness [5.0, 55.0]
Sunny	Rolloff [1500.0, 2000.0]
	Brightness [0.16, 0.21]

ユーザ B に提示したプレイリスト

- 朝：5 朝に好んで聞く明るい曲やさわやかな曲が多かった．
- 夜：4 夜に好んで聞くゆったりした曲や暗い曲が多かった．
- 晴れ：3 晴れに合う明るい曲もあったが、暗い曲も多かった．
- くもり：4 明るすぎず暗すぎない，くもりにあった曲が多かった．

これらの結果から，ユーザ B の“ 晴れ ”に対するプレイリストを除き，MALL がライフログ条件ごとに合う楽曲を適切に推薦できていることが示された．ユーザ B は，“ 晴れ ”のプレイリストには相反する 2 種類の印象を持つ楽曲が含まれていると述べた．ユーザ B の“ 晴れ ”に合致した相関性ルールを表に示す．一つ目の楽曲特徴量の組み合わせは，RMSenergy，Roughness 共に非常に小さい値となっている．この条件には，比較的純粋なハーモニーが使用されるエレクトリックでない楽曲が適合する．一方で，二つ目の組み合わせは，Rolloff は中間付近の値，Brightness は小さい値となっている．こちらの条件には，倍音をあまり含まないシンプルなアレンジの楽曲が適合する．このように，独立する 2 組の音楽特徴量ペアがユーザ B の“ 晴れ ”プレイリストの相関性ルールとして選出されたため，様々な印象を持つ楽曲を多く含んでしまう結果となったと考えられる．今後，この結果を改善させる方法を検討したい．

4 まとめ

本報告では，ライフログに基づいて状況・環境に合う楽曲を自動推薦する MALL を紹介し，ユーザテスト結果からその妥当性を検証した．ユーザテストの結果から，MALL による被験者 2 名への楽曲推薦結果が概ね妥当であることがわかった．また，あまりよくない結果が出た例について議論した．

今後の課題として，推薦結果をさらに向上させるため，相関性ルール選出のための支持度・確信度の閾値の再調整を行いたい．また収集するライフログデータの種類についても再検討したい．

参考文献

[1] O. Lartillot, “ MIRtoolbox ”, available from http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox

[2] RWC Music Database, available from http://staff.aist.go.jp/m.goto/RWC-MDB/

Autonomous Interactive Desk Environment の構築

藤田 真康　諏訪 博彦　栗原 聡 *

概要. オフィス等における PC を用いたデスクワークにおいては，各ウィンドウの配置や，机上のノート，ペン，マグカップといったオブジェクトの配置動作に煩わしさを感じることがある．この問題を解消するため，本研究ではユーザの代わりにウィンドウやオブジェクトの配置・整理整頓を自律的に行うシステム，AIDE (Autonomous Interactive Desk Environment) を提案する．通常のディスプレイに加え，机の部分にもタッチパネルを配置し，更にユーザに対して物理的支援を行うためのロボットアーム 2 基にて構成され，各ウィンドウとオブジェクトをそれぞれ自律エージェント化し，マルチエージェント協調メカニズムに基づいてユーザに対する適切なインタラクションを実行する．各エージェントの移動指標を設定するため，ユーザの作業ログから定型的な作業パターンを抽出し，これらのデータを収集することによりエージェントの移動先の学習を行う．

1　はじめに

オフィス等における PC を用いたデスクワークにおいては，各ウィンドウの配置や，机上のノート，ペン，マグカップといったオブジェクトの配置に煩わしさを感じることがある．しかし現状では，配置作業は手動で行われている．

このような人手の作業をロボット等のエージェントにより支援する研究がある．Duan[1] らは，デスクワークの 1 つであるセル生産組立を対象とし，ロボットが部品を作業者のもとまで配膳する，組立情報を机上に表示する等，作業者とシステムとの協調による組立作業支援システムを提案している．しかしこのシステムでは，動作タイミングをユーザが指示しなければならず，システムがユーザの作業状況を見て自律的に動作するという事はできない．より人の負荷を減らすには，ユーザの動作からユーザが次に何を求めるか予測し，自律的に支援することが必要である．

そこで本研究では，PC を用いたデスクワークを対象とし，デスクワーク支援システムとして，ウィンドウやデスク上にある物の配置･整理整頓をユーザの代わりに自律的に行う机型デバイス，AIDE (Autonomous Interactive Desk Environment) を構築する（図 1）．

2　AIDE の構成

AIDE では，ディスプレイを正面に 1 台置き，手前に 3 台寝かせ，寝かせた 3 台のディスプレイを情報表示可能な机として利用する．本研究では，机として利用するディスプレイを「水平ディスプレイ」と呼ぶ．水平ディスプレイを用いることにより，本や書類，メモ書き等の手元にある物とディスプレイ上の情報を見比べやすくする．また，手前の 3 台のモニターはタッチモニターとなっており，直感的な操作を可能とする．



図 1. AIDE

また，机上の状態を認識するため，ウェブカメラを設置する．また，机の奥，左右にロボットアームを設置し，ウェブカメラとロボットアームを用いてデスク上の物の配置や整理整頓といった物理的支援を行う．さらに，ユーザの手の動きや姿勢に対応してインタラクションを実行させるために，Kinect を設置し，ユーザの動作を検出する．

3　提案システムの概要

3.1　ウィンドウの自律エージェント化

ディスプレイ上に表示されている情報を自動的に配置･整理整頓を行うため，ウィンドウ等を自律エージェント化する．本研究では，自律エージェント化

Copyright is held by the author(s).

* 電気通信大学 大学院情報システム学研究科 社会知能情報学専攻

したウィンドウを“ウィンドウエージェント”と呼ぶ．

ウィンドウエージェントの目的は“ユーザに自分を見てもらう”ことである．この目的を設定することより，ウィンドウエージェントは自身の存在価値が最も高くなる位置へと自ら移動する．

ウィンドウエージェントの移動先候補は，過去に自身が配置された位置やユーザがよく視線を送る位置等から設定し，各ウィンドウエージェントはなるべく希望順位の高い移動先候補へ移動しようとする．本研究では，ディスプレイ上のある位置へユーザが視線を送る頻度を，その位置における“視線頻度”と呼ぶ．

しかし，上記の設定だけでは複数のウィンドウエージェントが特定の箇所に集中するという問題が起こる．そこで個々のウィンドウエージェントに“優先度”を設定し，優先度が高いウィンドウエージェントほど，視線頻度が高く，自身が頻繁に配置される位置へと移動できるようにする．優先度は，ユーザのウィンドウ操作履歴から，アクティブウィンドウ状態の時間等を用いて算出する．例えば，ブラウザやPDF ファイル等，資料として用いることが多いアプリケーションよりも，Microsoft Office Word や Latex 等，仕事として用いる事が多いアプリケーションの方がより優先度が高くなるようにする．個々のウィンドウエージェントに優先度を設定する事により，第 1 移動候補が同じであるウィンドウエージェントが複数存在した場合，最も高い優先度を持つウィンドウエージェントが自身の第 1 移動候補へ移動でき，優先度が低いウィンドウエージェント達は自身より高い優先度を持つウィンドウエージェントと重なる事を避けるため，自身の移動先候補の中から次の候補の位置へと移動する．ウィンドウエージェントに与えられる優先度は，ウィンドウエージェントが起動したときに設定される．

また，ユーザのウィンドウ操作記録を取得することで，ウィンドウ同士の定型的な配置パターンを発見し，ウィンドウエージェントの累計の優先度だけでなく，定型的な配置パターンの時だけの優先度を計算する．これによりパターンに当てはまるウィンドウエージェントが起動したとき，定型的な配置パターンの時の優先度を適用し，ウィンドウエージェントが定型的な配置位置へ移動することを可能とする．

3.2 机上のオブジェクトの自律エージェント化

ユーザのデスク上には様々な物が存在している．本研究では，デスク上にある物を“机上オブジェクト”と呼ぶ．AIDE システムでは，ディスプレイをデスクとして利用する事から，ウィンドウと机上オブジェクトが重なってしまうという状況が起こり得る．ウィンドウとオブジェクトが重なってしまった場合，ユーザは仕事の手を止めてウィンドウと机上オブジェクトの重なりを解決しなければならない．そこで，机上オブジェクトの移動手段にロボットアームを用いることで，机上オブジェクトを自律エージェント化する．本研究では，エージェント化した机上オブジェクトを，“オブジェクトエージェント”と呼ぶ．

オブジェクトエージェントにもウィンドウエージェント同様に，ユーザの使用頻度から優先度を計算し，自身の配置履歴とディスプレイ上の視線頻度の状況から，移動先候補を設定する．

3.3 エージェントの移動指標の設定方法

ウィンドウエージェントの優先度を設定するために，ユーザが使用した各アプリケーションのウィンドウサイズやウィンドウ配置等，ウィンドウ操作情報を記録する．また，視線頻度を算出するために，Kinect を用いてユーザの姿勢や動作，顔の位置，視線等も記録する．さらに，オブジェクトエージェントの優先度を設定するために，ウェブカメラを用いてオブジェクトの配置や使用頻度等，水平ディスプレイ上の状況を記録する．

ユーザがウィンドウを配置した位置は，必ずユーザにとって見やすい位置である．また，Kinect を用いてユーザの様子を観測する事により，ユーザがディスプレイ上のどの位置をどの程度見ていたか把握できる．これらのデータを収集することにより，ディスプレイ上の x-y 座標における視線頻度が定量化できる．ディスプレイ上の x-y 座標における視線頻度のマップを作成し，エージェントの移動指標とする．

そして，操作ログから，メール作成や文章作成，ネット検索といった，タスク単位での定型的なウィンドウ配置やデスク上の文房具の配置等のパターンを抽出し，定型的パターン適用時の優先度を設定するための指標とする．

4 おわりに

現在，図 1 の装置の組み立てが完成し，これから操作ログの取得を開始する．また，インタラクションのためのマルチエージェント協調メカニズムの構築を開始し，AIDE への提案システムの実装を目指す．提案システムを AIDE に実装した後，ユーザと AIDE との協調作業に関する被験者実験を行う計画である．

参考文献

[1] Feng Duan, Masahiro Morioka, Jeffrey Too Chuan Tan, Ye Zhang, Kei Watanabe, Nuttapol Pongthanya, Masao Sugi, Hiroshi Yokoi, Ryou Nihei, Shinsuke Sakakibara and Tamio Arai, ” Multimedia based Assembly Supporting System for Cell Production”, 2008

BrightFace: プロジェクションによる表情拡張システム

辻田 眸　暦本 純一*

概要. 対人コミュニケーションにおいて，重要な役割を果たしているのは表情や身振り，声の調子などの非言語情報であると言われている．なかでも表情は自分の感情を表出し，また他者の感情を理解するために効果的な手段である．しかしながら人によっては自分の感情をうまく表現できなかったり，他者には伝わりにくいときもあるだろう．メールなどの文字メッセージでのコミュニケーションの場合，顔文字や絵文字などを利用し，自分の感情をより効果的に伝えることができる．そこで実世界の対面コミュニケーションにおいても，文字メッセージで行われているようなアイコンを利用し，自分の表情を補足することができれば，より効果的に自分の感情を他者に伝達でき，コミュニケーションに役立つのではないかと考えた．本研究ではユーザの表情に応じて，アイコンをユーザの顔に投影し，表情を拡張表現することで，対人コミュニケーション支援を行うシステム「BrightFace」を提案・実装した．

1　はじめに

Mehrabian 氏らの研究 [1] によると，対人コミュニケーションにおいて，表情やしぐさが他者の印象や感情を判断，理解するのに重要な役割を占めているとされている．しかしながら人によっては自分の感情をうまく表現できなかったり，また他者の表情を読み取れず，感情の読み間違いをしてしまうこともあるだろう．一方，メールやチャットなど，文字によるコミュニケーションにおいては，顔文字や絵文字など，自分の表情を表すアイコンを利用することで，相手に自分の感情や雰囲気を伝達する役割を果たしている．さらに漫画やアニメーションの世界では，キャラクターの顔にアイコンを書き加えたり，頭上に線を書き足し強調して表現することで，絵に対する印象を変化させる表現技法が知られている．

そこで，日常生活での対面会話のときに，このような文字コミュニケーションで行われているアイコンの付加や漫画等で用いられている表現技法を利用し，ユーザの表現能力を拡張することで，お互いの感情理解を深め，コミュニケーション円滑化の支援になるのではないかと考えた．本研究ではユーザの表情に応じて，アイコンをユーザの顔に投影し，表現能力を拡張することで，コミュニケーション支援を行うシステム「BrightFace」の提案を行う．

2　BrightFace

BrightFace はユーザの表情に応じて顔にアイコンを投影させることで，ユーザの感情をよりわかりやすく他人に伝達できるシステムである．例えば図 1 のように，ユーザが笑顔になったときにはその表情を拡張表現できるような幸せそうなハートアイコンを，またユーザがやや悲しい表情をしたときには，涙マークなどのアイコンを顔にプロジェクションする．ユーザの表情に応じて，リアルタイムに表情理解が容易にできるようなアイコンをユーザの顔に投影してあげることで，ユーザの表現能力を拡張し，他者の表情・感情理解を深め，コミュニケーション円滑化につながるのではないかと考えた．



図 1. **BrightFace** システム．ユーザの表情を検出し，それに応じたアイコンをプロジェクションする．

本論文では基本的感情表出の中の，「喜び・怒り・悲しみ」に着目した．プロジェクションするアイコンについては，漫画でその表情を拡張表現するときに用いられているような一般的なアイコンを採用した（図 2）．ただし喜び（笑顔）のアイコンについては，様々なアイコンを投影し検討した結果，より表情理解が容易にできるような，図 2 のようなハートアイコンを投影することにした．表情拡張する感情の種類，提示するアイコンの適正については今後評価実験を通じてさらに検討していきたい．

Copyright is held by the author(s).

* 東京大学大学院情報学環　株式会社ソニーコンピュータサイエンス研究所



図 2. プロジェクションするアイコンの例.

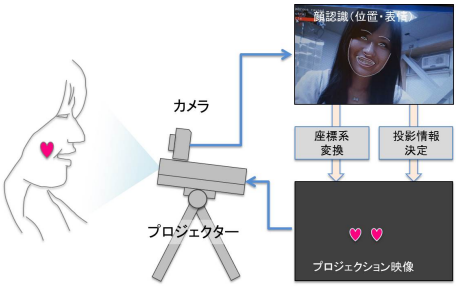


図 3. BrightFace のシステム構成.

2.1 プロジェクターでのアイコン投影

ユーザの表情に対してフィードバックを与える方法はいくつか考えられる．例えば，ユーザの表情や会話の内容に関するフィードバックをディスプレイに投影したり，テーブルとディスプレイが一体となっているデバイスに，アニメーションを映し出したりすることも考えられる．しかしながら，そういった場合，相手の顔から目線をはずすために，相手の表情とアイコンの両方を見ることは容易ではない．そこでユーザの目線をそらすことなく，また会話を阻害しないような形で，他者の表情理解の支援，また自分自身の表現能力の拡張を行い，感情を容易に相手に伝えることができるような表現方法を模索した．その結果，顔に直接プロジェクションすることで，相手の表情とアイコンを同時に見ることができるので，ユーザの目線をそらすことなく，またプロジェクターによりアイコンだけでなく，様々な表現が可能になるため，会話を阻害せず，コミュニケーション支援に役立つのではないかと考えた．

3 システム構成

BrightFace のシステム構成について述べる（図 3）．USB カメラと小型プロジェクターが一体になっている．プロジェクターについては，小型でかつ，明るい室内においても，プロジェクションされたアイコンや効果がわかるように，サンワサプライ株式会社の DLP 方式採用のモバイルプロジェクターを使用している．テーブル上においてもあまり圧迫された感じがなく，会話自体を阻害しない小型なものを採用した．ユーザの表情検出には openframeworks の ofxFaceTracker を利用した．またあらかじめユーザの表情を登録しておくことで，笑顔だけでなく，ユーザの表情に応じてアイコンを表出・プロジェクションし，ビジュアルフィードバックを与える．アイコンはユーザの顔の傾き，顔の大きさに応じてサイズを変え，ユーザの顔に位置にあうように工夫している．

4 関連研究

これまでにも会話の可視化や，プロジェクションマッピングによって人が受ける印象を変化させる研究は多くなされている．埴淵らは [2] 2 者間の対面会話環境において，笑顔アイコンをテーブル上のディスプレイに提示することで，会話や顔面動作にどのような影響を与えるのかを検証した．IllumiRoom[?] は，プロジェクターと Kinect を組み合わせ，テレビの周囲に画面の延長となる映像や，画面と連動する映像を映し出すことで，ゲームの没入感を高めている．扇情的な鏡 [3] ではディスプレーに映し出されたユーザの顔をシステムが自動で目や頬の位置を変えることで，ユーザの表情とは関係なく，笑い顔や悲しい顔に変形するシステムである．我々は人の顔に直接プロジェクションを行い，表情の表現能力を拡張することで，人の印象がどのように変化するのかを検証したり，コミュニケーション支援や会話の活性化を目指している．

5 まとめと今後の予定

ユーザの表情に応じて，アイコンをユーザの顔に投影し，表情を拡張表現することで，対人コミュニケーション支援を行うシステム「BrightFace」を提案・実装した．USB カメラと小型プロジェクターを利用し，基本的感情表出の中の，「喜び・怒り・悲しみ」に着目し，それに応じたアイコンをユーザの顔に直接プロジェクションを行った．現在は卓上のシステムになっているが，将来的にはモバイルな

参考文献

[1] Mehrabian, A. Communication without words, Psychology Today, 1968, 2 (4), 53-55.

[2] 埴淵俊平, 伊藤京子, 西田正吾. ソーシャルインタフェースの設計に向けた笑顔アイコン利用方法の検討, ヒューマンインタフェース学会論文誌, 2011-11-25, 13, 4, 365-375.

[3] S.Yoshida, S.Sakurai, T.Narumi, T.Tanikawa and M.Hirose: Manipulation of an Emotional Experience by Real-time Deformed Facial Feedback, AH2013, Stuttgart, Germany, Mar, 2013.

ボケる対話型システムの検討

藤本 浩介　　岡部 誠　　尾内 理紀夫*

概要．本研究では，ボケる日本語の対話型システムを提案する．駄洒落やなぞかけ等，笑いに関するシステムの研究は数多く存在する．しかし，駄洒落やなぞかけ等は出力される文そのものが面白いということに対し，ボケやツッコミは対話そのものが面白いという違いがある．そこで，本研究では人間とコンピュータとの既存の対話型システムを参考にし，ボケる対話型システムを提案する．本システムは，テキスト情報を入力として受け取る．入力を解析した後，あらかじめ用意された辞書から類似した文字列を検索し，ボケを出力する．また，出力されたボケに対する入力からユーザがボケを理解できているのか，ボケに対しツッコミを入れているのか，ボケに便乗してボケているのかを判定し，それぞれの判定結果に適した出力を行う．さらに，音声合成や効果音，文字サイズの変更を用いて分かりやすく提示する．

1　はじめに

近年，Apple の Siri[1]や，NTT ドコモのしゃべってコンシェル[2]等の対話システムの普及により，人間とコンピュータが対話を行う機会が増えている．人間とコンピュータが対話を行うにあたって，ユーザの要求したタスクをこなすことは実用性という面で重要であるが，親しみやすさという面では面白い対話を行うことが重要であると考える．しかし，面白い対話を行うことを目的としたシステムは存在しない．そこで，我々は面白い対話を行う方法の１つとしてボケる対話型システムを提案する．なお，ボケやツッコミには様々な種類が存在するが、本研究では「ボケ」とはわざと間違った発言をすること，「ツッコミ」とはボケの間違いを指摘することと定義する．

2　関連研究

駄洒落やなぞかけ等の笑いに関するシステムの研究は数多く行われている [1, 2]．これらのシステムは入力に対して駄洒落やなぞかけ等を出力するという一往復でユーザとのやりとりが完結されている．駄洒落やなぞかけ等は出力される文そのものが面白さを含んでいるが，ボケやツッコミは対話そのものが面白さを含んでいるという違いがあるため，ユーザとの対話を行う必要性がある．

人間とコンピュータとの対話型システムに関する研究は多数存在する [3, 4]．対話型システムを実現する主流な方法の１つにルールベースがある．ルールベースとは，入力として与えられた文に代名詞が入っていた場合はこのような文を出力するといったようなルールをあらかじめ設定しておき，対話を実現する方法である．本システムはボケることに限定しているので，ルール記述のコストが小さいため，ルールベースで対話を行う．

3　システム概要

本システムは入力としてテキスト情報を受け取る．入力を解析し，ボケることが可能な場合にボケを出力する．その後，出力されたボケに対する入力からユーザがボケを理解できているのか，ボケに対しツッコミを入れているのか，ボケに便乗してボケているのかを判定し，それぞれの判定結果に適した出力を行う．対話の出力には音声合成を用い，システムがボケを出力した場合やユーザがツッコミを入力した際には効果音を出力する．システムとの対話例を表１に示す．

表 1. システムの対話例

ユーザ	:やっぱり徳川家の家紋ってかっこいい
システム	:チャラリ〜鼻から牛乳〜
ユーザ	:嘉門達夫じゃねえよ．つかベタなぼけやなぁ．お前はお笑い芸人の新人かよ
システム	:（耳が）でっかくなっちゃった！
ユーザ	:それはお笑い芸人のマギー審司だろ
システム	:良いツッコミです．話の続きをどうぞ

4　実装

ユーザインタフェースを図１に示す．システムがボケたり，ユーザがツッコミを入れた部分は対話のログの部分で文字サイズを大きくすることにより，

Copyright is held by the author(s).

* Kosuke Fujimoto，電気通信大学，Makoto Okabe，電気通信大学/JST CREST，Rikio Onai，電気通信大学

[1] Siri: http://www.apple.com/jp/ios/siri/?cid=MAR-JP-GOOG-IPHONE

[2] しゃべってコンシェル: http://www.nttdocomo.co.jp/service/information/shabette_concier/

どこでボケやツッコミが行われたのかを分かりやすく提示する．

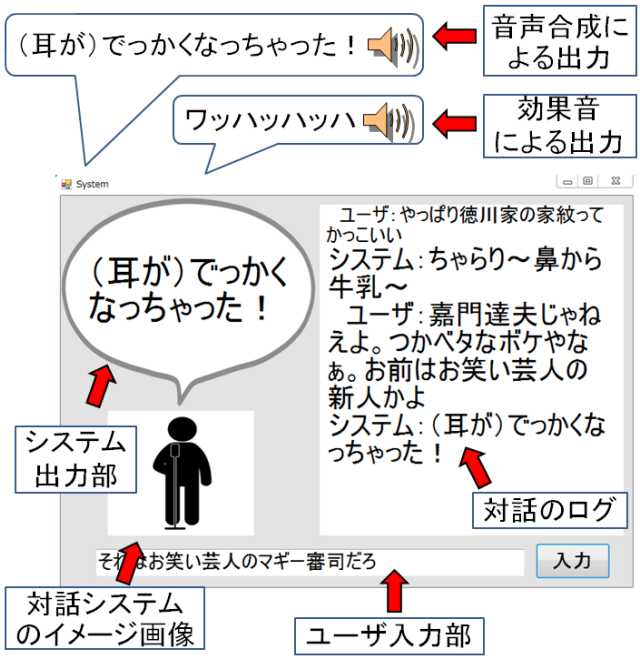


図 1. ユーザインタフェース

本システムはボケとして、わざと聞き間違いによるボケを出力する．その理由は，大島の研究により日本人が面白いと感じる話の傾向として「言い間違い・聞き間違い」による話の割合が一番高いということが判明しているからである [5]．

ボケを出力する際にはまず，入力されたテキスト情報に対して形態素解析を行い，形態素ごとに分割する．それらをあらかじめ用意された辞書と比較し，辞書から文字列が類似している項目を検索する．辞書には，芸人や歌手，それらに関連するギャグや曲名等のテキスト情報を Web サイト[3] [4]を参考にし，登録している．

文字列の類似度はレーベンシュタイン距離を用いて判定している [6]．レーベンシュタイン距離は対象の文字列が，比較対象の文字列に変形するまでに文字の削除や挿入等の手順がどれほど必要かを求め，文字列の類似度を判定する手法である．類似した文字列の項目が存在した場合，その項目を用いて聞き間違いによるボケを出力する．

本システムは，ボケに対する入力からユーザの状態を判定する．ボケに対してユーザは，「疑問」，「ツッコミ」，「ボケ」の 3 つの状態のいずれかになると仮定して判定する．「疑問」とは出力されたボケが理解できていない状態であり，「ツッコミ」とはボケに対してツッコミを入れている状態であり，「ボケ」とは出力されたボケに便乗してボケている状態である．

「疑問」状態の判定はユーザ入力の中に「誰」や「なに」等疑問を表す言葉が出現する場合にユーザは「疑問」状態であると判定し，何故そのようなボケをしたのかの説明を出力する．

「ツッコミ」は入力の語尾から判定する．あらかじめ漫才で用いられたツッコミを収集したテキストデータから頻出の語尾を特定し，それらの語尾が出現していた場合，ユーザは「ツッコミ」状態だと判定し，入力に対してさらにボケることが可能ならばボケを出力する．

「疑問」でも「ツッコミ」でもないものは「ボケ」状態だと判定し，ユーザに話を続けるよう促す．

本システムは，音声合成と効果音を用いる．音声合成はシステムによるボケ等の出力の際に用いる．効果音は，システムがボケた時やユーザがツッコミを入れた時に拍手や笑い声等を出力する．

5 まとめ

本研究ではレーベンシュタイン距離により類似した文字列をあらかじめ用意された辞書から検索することで，わざと聞き間違いによるボケを出力し，さらに，音声合成や効果音，文字サイズの変更を用いてユーザに分かりやすく提示するボケる対話型システムを提案した．

参考文献

[1] 田辺公一朗． 駄洒落のコンピュータによる処理: 駄洒落生成システムの基本設計． 産能大学紀要, Vol. 26, No. 1, pp. 65–74, 2005.

[2] 前田実香, 鬼沢武久． 単語の関連性とおもしろさを取り入れたなぞかけ生成．感性工学研究論文集, Vol. 5, No. 3, pp. 17–22, 2005.

[3] 稲葉通将, 平井尚樹, 鳥海不二夫, 石井健一郎. 非タスク指向型対話エージェントのための統計的応答手法. 電子情報通信学会論文誌 D, Vol. 95, No. 6, pp. 1390–1400, 2012.

[4] 岡本昌之, 山中信敏． Wizard of oz 法を用いた対話型 web エージェントの構築．人工知能学会論文誌, Vol. 17, No. 3, pp. 293–300, 2002.

[5] 大島希巳江． 日本人がおもしろいと感じる話の傾向: 日本一おもしろい話プロジェクト (2010 年 4 月 ~ 2011 年 3 月) の結果と分析．笑い学研究, No. 18, pp. 14–24, 2011.

[6] Vladimir I Levenshtein． Binary codes capable of correcting deletions, insertions and reversals. In *Soviet physics doklady*, Vol. 10, pp. 707–710, 1966.

[3] 音域.com: http://www.music-key.com/
[4] Yourpedia: http://ja.yourpedia.org/wiki/

紙窓: そこに置くだけで操作可能なカードインタフェース

加藤 邦拓　宮下 芳明 *

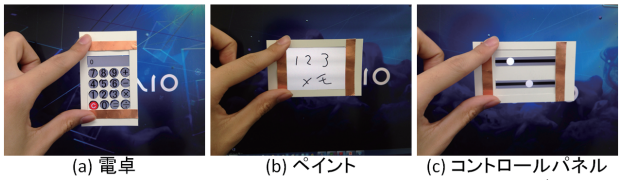
概要. コンピュータ上で作業を行う際，複数のアプリケーションを同時に使用することがある．しかし，それらの全てが作業全体を通して必要なものではなく，使用したらすぐに閉じてしまうようなアプリケーションも少なくない．こうした簡単なアプリケーションは深い階層に置かれてしまいがちであり，起動に煩雑な手順が必要となってしまう場合がある．本研究では,「今ちょっとだけ使いたい」簡単なアプリケーションを即時に起動する，置くだけで操作可能なカードインタフェースを提案する．複数の導電部を配置したカードをタッチパネルディスプレイ上に乗せることで任意のアプリケーションが表示され，ユーザはカードを通してアプリケーションを使用することができる．

1　はじめに

コンピュータ上での作業を行う際，作業の中心となって使用されるアプリケーションの他に，複数のアプリケーションを同時に使用することがある．しかし，それら全てのアプリケーションが作業全体を通して必要なものではなく，使用したらすぐに閉じられてしまうような，簡単な機能しか持たないアプリケーションも少なくない．

現在普及しているデスクトップメタファを取り入れた PC では，情報をディスプレイ内に並べて整理を行う．しかし，椎尾らも言うように，多くのユーザはデスクトップや浅い階層のディレクトリを，場所に基づいた情報整理が便利であるという理由で好んで使われるが，コンピュータ画面は有限であり，一等地はすぐに使い尽くされてしまう [1]．そのため，機能の少なく，使用頻度の低い簡単なアプリケーションは深い階層に置かれがちになってしまう．こうした簡単なアプリケーションであっても，一貫して同じ手順で起動を行う必要がある．場合によってはスタートメニューを開き，アプリケーションを探してから起動をするような，煩雑な手順が必要になり，ユーザに余計な負担がかかってしまう．

本研究では，作業の本筋ではないが「今ちょっとだけ使いたい」アプリケーションを即座に起動し，使用可能なカードインタフェースを提案する．カードに複数の導電部を配置し，タッチパネルディスプレイに乗せることでカードに対応したアプリケーションが表示される (図 1)．カード自体が入出力を持つインタフェースとなり，カードをディスプレイに置いている間のみアプリケーションを使用可能なため，カードを置く・離すといった動作によってシームレスな作業の切り替えが可能となる．また，カードという物理オブジェクトを用いることで，PC のディスプレイのような限られた環境でなく，広々とした現実の環境でのアプリケーション管理が可能となる．



(a) 電卓　(b) ペイント　(c) コントロールパネル スライダ

図 1. アプリケーション例

2　関連研究

実物体からコンピュータ内の情報へアクセスする研究として，椎尾らは IconSticker を提案している [1]．バーコードを印刷したラベルを実世界の物に貼り付け，リーダで読み取ることで，それに対応させたソフトウェアやファイルを呼び出すことができる．

実物体を介したインタフェースの研究として，Rekimoto らの DataTiles がある [2]．平面ディスプレイ上に無線タグを埋め込んだ透明なタイルを配置し，それぞれのタイルが異なったインタフェースとして使用できる．これらのタイルには操作ガイドとして溝が掘られたものがあり，ユーザは溝に沿ってペンデバイスを動かし操作ができる．本研究では DataTiles と同様に，カードインタフェースに操作ガイドとなる穴を空け，それに沿ってタッチを行い操作する．青木らは複数導電部を持つ円形の物理オブジェクトを用い，それを回す操作によって情報の拡大・縮小などの操作を行うくるみるを提案している [3]．Manuela らは AR マーカを取り付けた透明なタイルを用いたテーブルトップインタフェース Tangible Tiles を提案している [4]．Malte らは半透明なシリコン製のキーボードやスライダなどの物理オブジェクトを

Copyright is held by the author(s).

* Kunihiro Kato, 明治大学大学院理工学研究科新領域創造専攻ディジタルコンテンツ系, Homei Miyashita, 明治大学総合数理学部先端メディアサイエンス学科, 独立行政法人科学技術振興機構, CREST

タッチスクリーン上に置くことで，それぞれに応じた操作が可能な SLAP を提案している [5]．Yvonne らは，タブレット端末上に物理オブジェクトを乗せることによる大型ディスプレイ用インタフェースとして Tangible Remote Controllers を提案している [6]．これらの研究では，テーブルトップ環境や，タブレット端末上での作業に物理オブジェクトを用いているが，本研究では一般的な PC 環境での使用を目的としている．

3　提案手法

提案手法ではカードをタッチパネルディスプレイに乗せることで，カード毎に対応したアプリケーションを表示する．各カードはそれぞれ異なったデザインの穴が切り抜かれており，ユーザはカードを置いたまま穴からディスプレイをタッチして操作を行う．

タッチパネルディスプレイに乗せたカードの位置の検出と識別を行うため，図 2 のように複数の導電部をカードに配置し，それら全てを接続する．ユーザがカードをディスプレイに置いたまま，表面の導電部に触れることでカードの位置を検出する．アプリケーションは，カード裏面に配置された 4 点の導電部全てがディスプレイに接触している状態の時のみ表示される．また，この 4 点の導電部の配置パターンによってカードの識別を行い，それぞれのカードに対応したアプリケーションが表示される．

ユーザのタッチ操作部として，カードの内側を一部切り抜いた．この際，各カードに割り当てられたアプリケーションのイメージに合わせたデザインでカードを切り抜く．これによりユーザはカードの見た目から割り当てられたアプリケーションをある程度判断することができる．

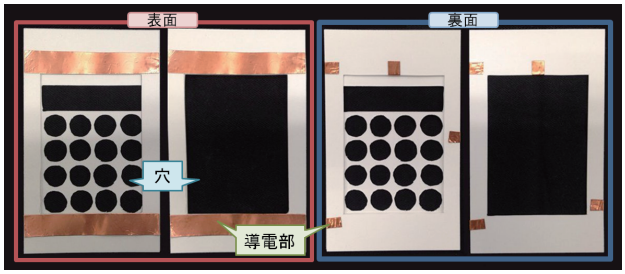


図 2. アプリケーションカード

4　アプリケーション例

プロトタイプとして電卓，ペイント，コントロールパネルスライダの 3 種類のアプリケーションを実装した．

電卓アプリ (図 1a) では，ボタン部として 4*4 マスに穴を空け，数値表示部として長方形の穴を空けたカードを用いる．カードをディスプレイに乗せることで，ボタン部の穴に数値や演算子などのボタンが表示され，タッチにより入力ができる．また，アプリ終了時に計算結果をクリップボードに記録する．

ペイントアプリ (図 1b) では，長方形に穴を空けた枠型のカードを用いる．ユーザはカードの枠内に自由にストロークを描くことができ，簡単な手書きメモアプリとして使用できる．カードをディスプレイから離し，アプリケーションを終了する際，描画した内容を画像として保存する．

コントロールパネルスライダ (図 1c) では，数カ所の細長い穴を空けたカードを用いる．このカードに空けた穴がそれぞれスライダとなっており，カード内の穴に沿ってディスプレイをタッチしスライダを動かす．このスライダにより，PC の音量，ディスプレイ輝度の調整を行うことができる．

5　導電性インクの導入

今回試作したアプリケーションカードは，紙製のカードに導電部を貼り付けて実装したが，現状の問題点としてカードを作成する手間がかかることが挙げられる．そこで，カードの作成に導電性インクとインクジェットプリンタの導入を検討している．導電性インクを用いたプリンタにより出力されたカードを指示通りに切り抜いたり，折り曲げたりするだけで容易にカードの作成ができると考えられる．

参考文献

[1] 椎尾 一郎, 美馬 義亮: IconSticker : 紙アイコンによる情報整理, コンピュータソフトウェア Vol.16, No.6, pp.24-32, 1999.

[2] Jun Rekimoto, Brygg Ullner, Haruo Oba: DataTiles: AModular Platform for Mixed Physical and Graphical Interactions, Proceedings of CHI'01, pp.269-276, 2001.

[3] 青木 良輔, 宮下 広夢, 井原 雅行, 大野 健彦, 千明 裕, 小林 稔, 鏡 慎吾: くるみる: 複数導電部をもつ枠型物理を用いたタブレット操作, 報処理学会研究報告 HCI, ヒューマンコンピュータインタラクション研究会報告 HCI-144, pp.1-8, 2011.

[4] Manuela Waldner, Jorg Hauber, Jurgen Zauner, Michael Haller, Mark Billinghurst: Tangible Tiles: Design and Evaluation of a Tangible User Interface in a Collaborative Tabletop Setup, Proceedings of OZCHI'06, pp.151-158, 2006.

[5] Malte Weiss, Julie Wagner, Roger Jennings, Yvonne Jansen, Ramsin Khoshabeh, James D.Hollan, and Jan Borchers: SLAP widgets: bridging the gap between virtual and physical controls on tabletops, Proceedings of CHI'09, pp.481-490, 2009.

[6] Yvonne Jansen, Pierre Dragicevic, Jean-Daniel Fekete: Tangible Remote Controllers for Wall-Size Displays, Proceedings of CHI'12, pp.2865-2847, 2012.

擬音語アニメーションによる動画音響の可視化手法

王 方舟　　柏野 邦夫　　永野 秀尚　　五十嵐 健夫*

概要.

動画に含まれる音の情報は，動画の視聴体験を構成する重要な要素である．一方で，実際に視聴者が動画を視聴する際は，状況によっては必ずしも音を伴って視聴できない場合が存在する．従来，このような状況下で音の情報を視覚的に補う手段としては字幕が用いられてきたが，そのほとんどが人の発した声を文字に書き起こしたものであり，声以外の一般的な音の情報に関する表現力は非常に乏しいのが現状である．本稿では，動画中に生起する音の種類を自動で判別し，擬音語（オノマトペ）を用いて可視化する手法を提案する．生成された擬音語は音の変化に合わせてアニメーションされる．これにより，動画中に含まれる一般的な音の種類およびそのダイナミクスを自然な形で可視化し，視聴者に伝えることが可能になる．

1　はじめに

多くの動画コンテンツにおいて，音は重要な役割を担う．人の話すセリフやナレーションなどの音声情報はそれ自体が不可欠な情報であるし，それ以外の音であっても，例えばカーチェイスのシーンでのエンジンの唸り，サッカーの試合のシーンでの人々の歓声など，様々な場面において音は視聴者の感情に影響を及ぼし，映像と合わせて一つの視聴体験を構築する．一方で，視聴者が動画を視聴する状況は様々であり，必ずしも常に音が聞き取れるとは限らない．そのような状況下では，動画本来の視聴体験が大きく損なわれることがある．

従来，このような状況下で音の情報を伝達する手段としては，字幕が使われてきた．一般的に使われている字幕は，静的なテキストをシーンの進行に合わせて画面下に表示するものであるが，1) 人手による事前の準備が必要であり，2) 人の声以外の効果音や環境音に対する表現力が低く，3) 音の生起のタイミングや変化を表現できない，などの問題点がある．特に，効果音や環境音については，単に「ガラスの割れる音」など音の状況を説明する文字列を配置するものが多く，音が本来持つ視聴体験を高める働きは失われてしまっている．

この問題を解決するため，本研究では，動画中に生起する音の種類を認識して擬音語に変換することによって可視化し，さらに音の変化に合わせて適切にアニメーションする手法を提案する．提案手法は入力された動画から音の情報を自動で認識するために人手の介在が不要である．さらに音を自然な視覚表現である擬音語に変換した上で，音の変化に合わせて適切にアニメーションすることにより，音のない状況下でも豊かな視聴体験を提供することを可能にする．



図 1. 擬音語アニメーションによる動画音響の可視化例

2　関連研究

音の種類を識別するための研究は従来より多くなされてきた [2] これに対し，石原ら [1] は擬音語の表現力に着目し，環境音を直接擬音語に変換する手法を提案している．山本ら [3] は，環境音を直接擬音語に変換し，音の大きさや質に応じた適切なフォントを用いて可視化する手法を提案している．一方で山本らの提案手法は極めて限定された実験環境を対象としており，単発の短音のみを対象とし，一度生成された擬音語表示は音の変化に対して静的であるなど，実在の動画に対して適用し得るものではない．

3　システム設計

本研究で実装したシステムは，自動車のレース動画を対象とし，動画中のエンジン音，およびドリフトやブレーキの際に発生するスキール音の 2 種類の音を自動認識して擬音語アニメーションを付与するものである．生成する擬音語の種類およびアニメー

Copyright is held by the author(s).

* Houshu Oh and Takeo Igarashi, 東京大学大学院 情報理工学系研究科 コンピュータ科学専攻, Kunio Kashino and Hidehisa Nagano, NTT コミュニケーション科学基礎研究所.

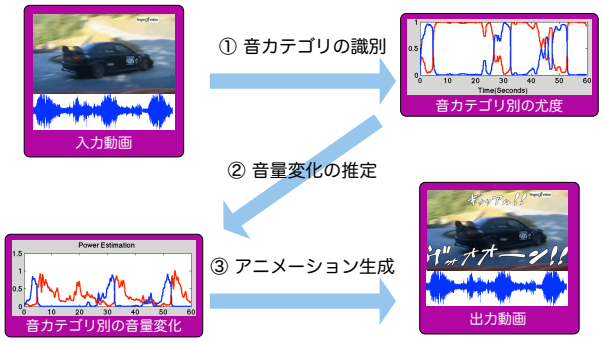


図 2. アルゴリズムの概要

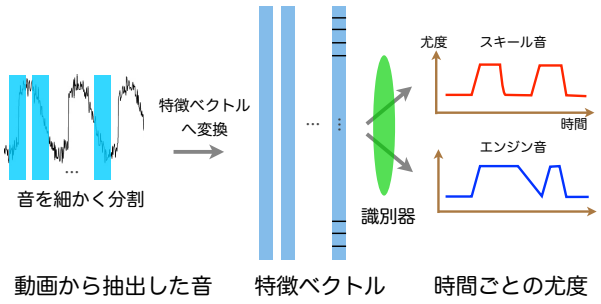


図 3. 音カテゴリの識別方法

ションのデザインは，考えうる最もシンプルな形として，1) 基本的な音の種類 (エンジン音またはスキール音) に応じてあらかじめ用意された擬音語画像を，2) 映像の固定位置に表示し，3) 音の大きさに合わせた拡大および縮小アニメーションを行うものとする．

4 アルゴリズム

提案するアルゴリズムは，音付きの動画を入力とし，擬音語アニメーションが付与された動画を出力するものである．アルゴリズムは大きく分けて，1) 音カテゴリの識別，2) 音量変化の推定，3) 擬音語アニメーションの生成，の 3 つのステップをたどる (図 2)．

アルゴリズムはまず，入力された動画の音の種類を識別する．この際に用いる識別器は，ラベル付きの音データを用いて SVM で機械学習を行うことで事前に構築しておく．構築する識別器は，0.75 秒の音の断片から生成された特徴ベクトルを入力とし，その音の断片がどの音カテゴリに属するかの尤度を計算するものである．入力された動画の音データを時間方向に細かく分割し，各断片を特徴ベクトルに変換して，構築した識別器に入力として与えることにより，動画の時系列に沿って音のカテゴリ別の尤度が計算される (図 3)．

アルゴリズムは次に，音のカテゴリ別に音量変化を推定する．推定は，前のステップで計算された音カテゴリ別の尤度に，元の音の音量を掛け合わせることによって行われる．

アルゴリズムは最後に，推定されたカテゴリ別の音量変化をもとに擬音語アニメーションを生成する．アルゴリズムは入力動画を 1 フレームずつ見ていき，それぞれのフレームに対応する音のカテゴリ別音量を前ステップの結果から得る．次に，音量が一定の値以上を超えた音のカテゴリについて，擬音語画像を生成する．現実装では，システムが予め音の種類の数だけ準備した擬音語画像をそのまま用いる．最後に，生成した擬音語を対応する音カテゴリの音量の大きさに合わせて拡大・縮小し，映像に埋め込む．

5 議論と今後の課題

環境音から擬音語を生成する研究は以前よりなされてきたが，その多くは入力となる音を日本語に直接変換することを目的としている [1]．このような手法は擬音語表現が豊富な日本語に対しては有効であるが，一方で他の言語に適用しにくいというスケーラビリティの問題や，音の混合に弱いという問題を抱えている．本稿で提案した手法は音を言葉に直接変換するのではなく，音の「種類」を認識し，対応する擬音語に変換するという手法を採用している．これにより，言語的な表現力を多少犠牲にしつつも，多言語拡張へのスケーラビリティと音の混合に対する一定のロバスト性を確保している．

今後の課題としては，音源物体の位置に合わせた擬音語の適切な配置がある．これは，一般物体認識の手法や Saliency 解析などを用いて行うことが考えられる．音の変化をより自然に捉えたアニメーションの生成も課題である．例えば，音量の起伏を複数の立ち上がりと減衰に分割することで，同一カテゴリの音についても複数の音がかぶさりあって生起している様子を捉えることが可能になると考えられる．さらにはアニメーションの要素として，拡大・縮小だけではなく，ゆるやかな減衰部分についてはゆっくりフェードアウトさせるなどといったより自然なアニメーションも実現できるようになると考えられる．

参考文献

[1] K. Ishihara, T. Nakatani, T. Ogata, and H. G. Okuno. Automatic Sound-Imitation Word Recognition from Environmental Sounds focusing on Ambiguity Problem in Determining Phonemes. *PRICAI 2004: Trends in Artificial Intelligence. Lecture Notes in Artificial Intelligence*, 3157:909–918, 2004.

[2] D. Li, I. K. Sethi, N. Dimitrova, and T. McGee. Classification of general audio data for content-based retrieval. *Pattern Recognition Letters*, 22(5):533–544, 2001.

[3] 山本貴史, 松原正樹, 斎藤博昭. 擬音語と書体表現を用いた環境音の可視化. 芸術科学会論文誌, 11(1):1–11, 2012.

ダンスパフォーマンスにおける 2 軸動作可能な電動カーテンを用いた隊形練習支援

土田 修平　　寺田 努　　塚本 昌彦*

概要. 近年，ダンススタジオの増加やキッズダンスの流行など，ダンスパフォーマンスによる身体表現能力やリズム感の向上に注目が集まっている．ダンスの初心者でも大人数でダンスパフォーマンスを披露する機会が増えているが，複数人でのダンスパフォーマンスでは音楽と振り付けが上手くリンクしているものであっても，隊形が崩れているとダンスパフォーマンスとしての質が大幅に低下してしまう．そのためダンスパフォーマンスにおいて，隊形の練習は振り付けの習得と同様に重要な要素となっている．隊形を上手く魅せるためには，バラつきなく人との間隔を保つことと，同じタイミングで動くことを意識しなければならない．しかし練習の際に一人でもダンサーが欠ければ，隊形の適当な間隔を掴むことが困難となる．そこで先行研究 [1] において，隊形練習における一部のダンサーが欠けてもスムーズに練習を行うための自走型スクリーンを用いたダンス練習支援システムを提案した．しかし，自走型スクリーンは不規則な移動による過剰な存在感，スクリーンとの衝突の恐れなどの要因により人の移動を必要以上に制限したため，人と踊る際の位置の再現性が低かった．そこで本研究では，衝突の恐れが少なくよりスクリーンを正確に移動させるために，スクリーンの素材が柔らかくても良く，移動方向がずれないカーテンレールに着目し，映像が投影された 2 軸動作可能な電動カーテンと踊る手法を提案する．

1　はじめに

ダンスパフォーマンスを上手く魅せるためには，バラつきなく人との間隔を保つことと，次の隊形へスムーズに移行することを意識しなければならない．しかし練習の際に一人でもダンサーが欠ければ，隊形の適当な間隔をとることが困難となる．これまでにダンスパフォーマンスを支援する様々な研究が行われている．モーションキャプチャ技術を用いた手法 [2] や慣性センサを用いた手法など振り付けの習得を支援するものが挙げられる．これまでにダンスパフォーマンスを支援する様々な研究が行われている．モーションキャプチャ技術を用いた手法や慣性センサを用いた手法 [3] など振り付けの習得を支援するものが挙げられる．しかし，隊形練習に着目した研究は行われていない．そこで先行研究 [1] では，隊形練習における一部のダンサーが欠けてもスムーズに練習を行うための自走型スクリーンを用いたダンス練習支援システムを提案した．しかし，自走型スクリーンは不規則な移動による過剰な存在感，スクリーンとの衝突の恐れなどの要因により人の移動を必要以上に制限したため，人と踊る際の位置の再現性が低かった．そこで本研究では，衝突の恐れが少なくよりスクリーンを正確に移動させるために，スクリーンの素材が柔らかくても良く，移動方向がずれないカーテンレールに着目し，映像が投影された 2 軸動作可能な電動カーテンと踊る手法を提案する．以降，2 章では 2 軸動作可能な電動カーテンの設計について述べ，3 章で実装について述べ，最後に 4 章でまとめを行う．

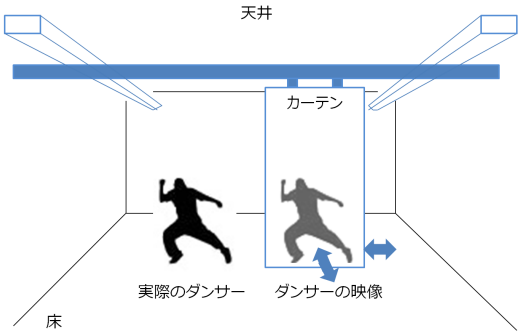


図 1. 2 軸動作可能な電動カーテンの配置，移動方向

表 1. 自走型スクリーンとの違い

	利点	欠点
2 軸移動可能な電動カーテン	制御が簡単 高速移動可能	大仕掛け 複数人での利用は難しい

2　2 軸動作可能な電動カーテンの設計

2 軸動作可能な電動カーテンの配置，移動方向についてに図 1 に示す．2 本のカーテンレールを平行

Copyright is held by the author(s).

* Shuhei Tsuchida and Masahiko Tsukamoto, 神戸大学大学院工学研究科, Tsutomu Terada, 神戸大学大学院工学研究科 / 科学技術振興機構さきがけ

表 2. 被験者 3 人の距離差の平均値 (mm)

実験番号	1 人で踊る	自走ロボットと踊る	プロジェクタ映像と踊る	自走型スクリーンと踊る	カーテンと踊る
平行移動	56	150	83	204	150
接近移動	74	90	188	109	63

に天井に取り付け，それらに対し垂直に新たに 1 本のカーテンレールを先程の 2 本のカーテンレールで吊るしている．これによりカーテンを 2 軸に移動させることができる．自走型スクリーンとの違いを表 1 に示す．カーテンレールが固定されており移動方向がずれないため不規則な動きを抑えられること，上からスクリーンを吊るすためスクリーンを固定する枠を必要とせず衝突の際に怪我の恐れが軽減できること，また床を走る自走ロボットを必要としないためより近い距離ですれ違うような練習ができる．さらにスクリーンの重量が減ったために自走型スクリーン以上のスピードを出すことができ，重量があるモータごと移動させる必要がないため高出力のものが利用でき更なるスピードの上昇が期待できるなどの利点が挙げられる．しかし，自走型スクリーンと比べカーテンレールを天井等に設置するなど仕掛けが大きくなること，複数人での利用方法が限定されてしまうことなどの欠点が挙げられる．これらの理由より，複数人での利用については限定されるものの 2 軸動作可能な電動カーテンは自走型スクリーンの問題を解決できるだけではなく，自走型スクリーンでは利用しにくい練習もできると考えられる．また，カーテンによる影響を調べるために考えられるそれぞれの隊形練習手法において，実際にダンサーと二人で踊る際の感覚に近いかどうか評価した．結果を表 2 に示す．数値が小さい程二人で踊る際の移動に近いといえる．

3 2 軸動作可能な電動カーテンの実装

実装した 2 軸動作可能な電動カーテンの外観を図 2 に示す．横方向の移動についてはナビオ社の電動カーテンレール TV24 の回転部分を加工したカーテンレールを利用した．回転部分はモータとカーテンが動作する部分とのキア比が低かったために，モータの回転が直接カーテンレールを動作する部分と繋がるよう加工した．これにより人の歩行速度程度のスピードで移動させることができる．前後方向については横移動に使用した電動カーテンレールの両端に垂直に 2 つのカーテンを取り付けることによって，電動カーテンレール自体を移動させることによって可能にしている．前後方向については現在手動での移動のみとなっており，今後横方向と同様にモータを取り付け自動化を目指す．スクリーンは風が抜けやすいようレースカーテンを使用している．



図 2. 2 軸動作可能な電動カーテンの外観

4 おわりに

本論文では，先行研究より得られた知見から考えられる 2 軸動作可能な電動カーテンの利点について述べ，自走型スクリーンの実装を行った．今後システムの更なる改善を目指し，複雑で長いに振りでの実験を行い，2 軸動作可能な電動カーテンによる影響ついて調査する．また，隊形練習において提示する存在感による振りなどへの影響を細かく調査することを目指す．

参考文献

[1] S. Tsuchida, et al.: A System for Practicing Formations in Dance Performance Supported by Self-Propelled Screen, *Proc. of tthe 4th Augmented Human International Conference (AHI '13)*, pp. 178–185 (Mar. 2013).

[2] L. Deng, et al.: Real-Time Mocap Dance Recognition for an Interactive Dancing Game, *2011 IEEE Computer Society Conference on Computer Animation and Virtual Worlds (CASA '11)* Vol. 22, pp. 229–237 (Apr. 2011).

[3] 林 貴宏, 尾内理紀夫: モーションキャプチャと加速度センサを用いた振りの練習支援, 電気学会論文誌 E, Vol. 129, No. 6, pp. 173–180 (June 2009).

モチベーションの維持を考慮したピアノ学習支援システム

福家 悠人　　竹川 佳成　　柳 英克*

概要.　楽器の演奏技術の向上には多大な時間や労力を必要とするため，敷居の高さに利用を断念したり，習熟効率の低さから挫折してしまう演奏者が多い．鍵盤演奏の敷居を下げるために，学習支援システムが提案されてきたが，学習者のミスに対して厳格で，打鍵ミスが続くと次に進めないためフラストレーションがたまり練習へのモチベーションが下がってしまう．そこで，本研究では，モチベーションの維持を考慮したピアノ学習支援システムの構築をめざす.

1　はじめに

ピアノ演奏では，譜読み，指示されている鍵への正確な打鍵，適切な運指 (指使い)，リズム感覚，打鍵の強弱，テンポなど，さまざまな技術が求められ，それらの修得には長期間の基礎的な練習を必要とする．ピアノ演奏には多大な時間と労力を必要とするため，敷居の高さに利用を断念したり，習熟効率の低さから挫折してしまう演奏者が後を絶たない．従来システムでは学習者がミスをしないように細心の注意を払わなければならず，打鍵ミスが重なると次の音符に進めない状態が続きフラストレーションがたまり学習者のモチベーション (熟達に向けて練習を続けたいと思う気持ち) が下がってしまう.

そこで，本研究ではモチベーションの維持を考慮したピアノ学習支援システムの構築をめざす.

2　関連研究

これまでピアノ学習の支援につながる試みはいくつか行われている．ビデオや音声による模範演奏の提示する，Piano Tutor[2] や打鍵すべき鍵，運指，手本映像を表示するキーボードやソフトウェア [1, 3] がある．これらはいずれも打鍵情報から演奏を評価し学習目的に必要な情報を提示しているが，本研究で提案するミスの許容度は考慮されていない.

演奏の敷居を下げる試みとして竹内らの Two Finger Piano[4] は，どの鍵を弾いても常に正しい音が出力される．これは，本研究で提案するミス許容度をすでに導入している事例であり，あたかも演奏しているように見せることができる．しかし，本研究のように段階的にミス許容度を変化させ演奏を学習するといった学習は考慮していない.

3　設計

1 章で述べたように，本研究ではピアノ初心者を対象としており，五線譜やシステムが生成する補助情報を活用しながら学習者はある楽曲を一から練習し，できるだけモチベーションを維持しながら習熟し，最終的にシステムの補助なしで演奏できるようになることをめざす.

3.1　システム構成

提案する学習支援システムのシステム構成を図 1 に示す．演奏者の前面にディスプレイを設置し，ディスプレイに楽譜や仮想鍵盤を提示する．また，MIDI 情報 (打鍵位置や打鍵強度) を入力とする．ディスプレイに表示された仮想鍵盤と実際の MIDI キーボードとのサイズは一致しているため，仮想鍵盤の鍵上に提示された情報が MIDI キーボードのどの鍵と対応しているか直観的に理解できる.

3.2　学習方法

提案する学習方法は許容範囲内の打鍵ミスであれば，学習者が本来弾くべき位置と異なる鍵を打鍵したとしても，練習している楽曲の音高データベースから正しい音高データを取得し，ベロシティ・打鍵タイミング・離鍵タイミングを残したままで，音高のみ正しい音高に差し替えて出力する．これにより，学習者が正確に打鍵できない低い熟達度であっても，完成度の高い演奏を行えるためモチベーションを維持しながら練習に取り組める．打鍵位置がミスの許容範囲外であった場合，誤りであることを意味する効果音を出力することで学習者が誤りを認識できるようにする．なお，楽曲として単音の旋律を想定している.

学習モードの許容度が低くなるにつれ，学習者の打鍵ミスを許容する範囲が狭くなる．以下で説明する学習モードは初心者に向けて，順に許容範囲を狭くした学習モードである.

エニーキーモード　エニーキーモードでは，学習者はどの鍵を打鍵したとしても正しい音高の音が出力されるモードである．ベロシティ，打鍵タイミング，離鍵タイミングは学習者の操作がそのまま反映

Copyright is held by the author(s).

* 公立はこだて未来大学

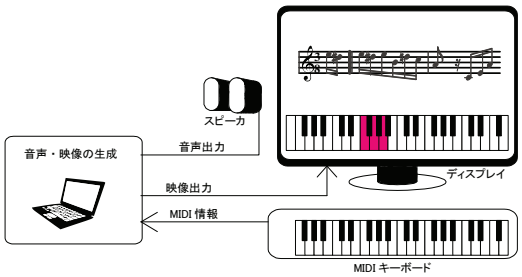


図 1. システム構成

される．すでに聞き慣れた楽曲を練習している場合，学習者はベロシティ，打鍵タイミング，離鍵タイミングを十分に理解しており，単音の旋律であるため，少しの練習で音長を制御しながら演奏できるようになる．これにより学習者は完成系をイメージでき，練習するモチベーションが向上する．ピアノ初心者が楽譜をみながら，楽譜通りの打鍵位置を正確に打鍵することが大変困難であることから，最も許容度の低いモードとして打鍵タイミングだけを意識できるエニーキーモードを提案する．

カラーモード　カラーモードは，学習者が黒鍵の演奏に慣れるために有効なモードであり，鍵の色 (黒鍵あるいは白鍵) が正しければ，システムは打鍵ミスと判定せず正しい音高を出力する．鍵盤楽器の他の楽器と異なる特徴として白鍵と黒鍵の存在がある．鍵盤をはじめて見た人でも分かる色の違いを楽譜上の記号から判断できるように慣れることを狙いとしている．本モードはエニーキーモードの次に許容度が低いモードとして設計している．

ディレクションモード　ディレクションモードでは，打鍵位置の方向が正しければ，システムは打鍵ミスと判定せず正しい音高に差し替えた音を出力する．次に打鍵する鍵と，現在演奏している鍵が同じ高さの場合，同じ鍵を弾く必要がある．初心者であっても，次に弾く音が現在の音よりも相対的に高いか低いかは認識しやすいため，楽譜における音符の高さが，鍵盤上の横の位置の対応関係にあることを理解することを狙いとしてディレクションモードを提案する．

ディレクションカラーモード　ディレクションカラーモードは，前述のディレクションモードとカラーモードを統合したモードである．次の打鍵位置の方向に加え，鍵の色が正しければシステムは打鍵ミスと判定せずに正しい音高を出力する．本モードはディレクションモードとカラーモードの両方を学習した後に利用されることを想定している．

鍵幅可変モード　鍵幅可変モードは本来打鍵するべき打鍵位置および鍵の色 (白鍵か黒鍵か) を基準に指定した鍵数範囲内であれば正しい打鍵とみなし，正しい音高を出力する．例えば，許容範囲が ± 1 であれば，本来打鍵すべき位置の両隣りまでを許容範囲とする．学習者は熟達するにつれ許容範囲を下げる．許容範囲 0 が，光る鍵盤 [1] と同じ難度になる．学習者が前述したモードの練習後，鍵幅可変モードで練習することを想定しており，前述したモードはいずれも，現在の打鍵位置を基準に正誤判定を行っていたが，鍵幅可変モードは，楽譜上の音符に対応する鍵を基準に正誤判定を行う．

打鍵位置提示無モード　これは，打鍵ミスを許容せず仮想鍵盤上に打鍵位置を提示しないモードで，楽譜上に現在の打鍵位置を示す緑の実線のみを提示する．楽譜を見ながら練習する状況とほぼ同じである．

4　実装

3 章で述べた学習支援システムのプロトタイプを実装した．PC は TOSHIBA 社の dynabook CX/47E を使用し，MIDI 鍵盤として M-AUDIO 社の eKeys を使用した．PC 上のソフトウェアの開発は，Windows Vista 上で Processing を用いて行った．

5　まとめ

本研究では，学習者のミスの許容度に注目しモチベーションの維持を考慮したピアノ学習支援システムを構築した．段階的に難度を変更し，その都度，無理なく少しの練習で高い成功体験を得られるように学習方法を設計した．

今後は，両手演奏や和音を含む場合における学習方法の拡張や，さまざまな世代を対象とした評価実験を行う予定である．

参考文献

[1] CASIO：光ナビゲーションキーボード：http://casio.jp/emi/key_lighting/.

[2] R. B. Dannenberg, M. Sanchez, A. Joseph, P. Capell, R. Joseph, and R. Saul: A Computer-Based Multi-Media Tutor for Beginning Piano Students, Journal of New Music Research, 19 (2-3), pp. 155–173 (1990).

[3] 竹川佳成, 寺田努, 塚本昌彦：リズム学習を考慮したピアノ演奏学習支援システムの構築, 情報処理学会インタラクション 2012, pp. 73–80 (2012 年).

[4] 竹内好宏, 片寄晴弘：Two Finger Piano による曲想の表現, 情報処理学会研究報告 (音楽情報科学), pp. 37–44 (2001 年).

Bezel Check: ベゼル上におけるダブルクロッシングジェスチャ

黒澤 敏文　　志築 文太郎　　田中 二郎*

概要．本稿では携帯型タッチスクリーン端末における新たなジェスチャである Bezel Check を示す．Bezel Check によりユーザは 1 ストロークの動作にてデータを空間的に配置する事が可能となる．Bezel Check はベゼルを開始点として画面端を通過するダブルクロッシングジェスチャであり，以下に示す利点を持つ．1) 素早い操作を行う事が出来る．2) 従来のタッチジェスチャと共存出来る．3) データを空間的に配置出来る．我々は Bezel Check によりテキストのコピーや Web ページのブックマークを行い，コピーされたデータを素早く使用出来るシステムデーモンを作成した．

1　はじめに

スマートフォンやタブレット端末等の携帯型タッチスクリーン端末はタッチやスワイプ等のシンプルなジェスチャの組合せにより用いられている．しかしながら，コピーやペースト等，頻繁に用いられるコマンドには操作に多くの時間を要するものがある．例えば，Windows 8 タブレット上においてユーザがテキストをコピーするためには，テキストを選択し，ロングタップによりポップアップメニューを表示し，メニューから「コピー」を選ぶ必要がある．

本稿では，携帯型タッチスクリーン端末において 1 ストロークの動作にて素早くデータを空間的に配置する事が可能となるジェスチャである Bezel Check を示す．Bezel Check はピンチやスワイプ等の従来のタッチジェスチャと競合せず，これらと共存する事が出来る．

2　関連研究

本手法にて用いられるジェスチャと同様にベゼルを開始点とするジェスチャが多く研究されてきた [2, 3, 5, 4]．Bragdon らは，ベゼルを開始点とするジェスチャはモバイル環境において最も速いジェスチャであり，ユーザに最も好まれる事を示した [1]．Roth らは，ベゼルからのスワイプにより従来のマルチタッチジェスチャと競合せずに操作出来る手法 Bezel Swipe を提案した [4]．我々は，ベゼルをジェスチャの開始点としてのみならず終了点としても用いる事により，物理的空間へデータの配置を行う事を可能にした．

3　Bezel Check

Bezel Check は携帯型タッチスクリーン端末のベゼル上におけるダブルクロッシングジェスチャである．ユーザはまず，テキストまたはウェブページ等のオブジェクトを通常通りの方法を用いて選択する．続いて，ユーザはベゼル上にてダブルクロッシングジェスチャをする (図 1 左)．即ち，ユーザはチェックマークを描くようにタッチスクリーンのベゼルから画面内部へスワイプし，またベゼルへ指を戻す動作をする．すると，Bezel Check を行ったベゼル付近の画面端に仮想クリップボード (図 1b 上部の矩形) が作成され，選択されたオブジェクトのデータがそこに格納される．仮想クリップボードに格納されたデータを参照するには，ユーザはまずそのデータを用いるアプリケーションにフォーカスを移す．続いて，ベゼルから参照したいデータが格納されている仮想クリップボードをクロッシングするようにスワイプする．すると仮想クリップボードを中心にマーキングメニューが展開されるので，ユーザはその中から実行したいコマンドを選択し，指を離す（図 1 c）．



(a)　(b)　(c)

図 1. **a) Bezel Check** の実行 **b)** 仮想クリップボード（上部の矩形）の表示 **c)** ベゼルからのシングルクロッシングにより表示されるマーキングメニュー

また，ユーザは画面端に配置された仮想クリップボードに対して，移動/他の仮想クリップボードとの統合/消去を行う事が出来る．ユーザが統合された仮想クリップボードにタッチした際，その内容がその仮想クリップボードを中心に展開される（図 2 左）．ユーザは統合された仮想クリップボード内のデータを用いたい場合，展開されたデータをベゼル部分からスワイプする事により選択し，表示されたマーキングメニューからコマンドを選択し，指を離す（図 2 右）．これにより，ユーザは容易に仮想クリップ

Copyright is held by the author(s).

* Toshifumi Kurosawa，筑波大学大学院 システム情報工学研究科，Buntarou Shizuki and Jiro Tanaka，筑波大学 システム情報系

ボードの位置関係の変更や，データのグルーピングを行う事が可能となる．

3.1 利点

Bezel Check はベゼルからのダブルクロッシング動作のみにより用いる事が出来るため，次に挙げる利点を持つ．

1. 素早い操作：Bezel Check は 1 ストロークにて操作できるため，ユーザはコピー等の動作を通常の方法よりも素早く行う事が可能となる．
2. 従来のマルチタッチジェスチャとの共存：ベゼルからのスワイプ動作には，ピンチや通常のスワイプ等の従来のタッチジェスチャと競合せず，これらと共存出来るという利点を持つ [4]．
3. 空間的なデータ配置：ユーザは画面端の任意の位置にデータを配置する事が出来る．そのため，ユーザはデータをタッチスクリーンの四辺に分類して配置する事が出来る．また，データは画面端に表示されるので画面の表示領域をあまり減少させない．

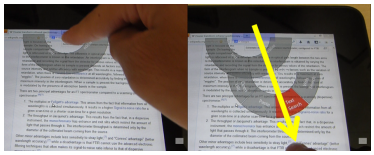


図 **2.** (左) ベゼルからのスワイプによる，複数データが格納された仮想クリップボードの展開 (右) コマンドの選択

3.2 アプリケーション

我々は携帯型タッチスクリーン端末において Bezel Check を提供するシステムデーモンを開発した．本手法により展開されるマーキングメニューに表示されるコマンドは，端末上においてアクティブになっているアプリケーションに依存して決定される．今回の実装では，テキストエディタ・Web ブラウザ・PDF ビューアの 3 種類のアプリケーションを対象にした．表 1 にそれぞれのアプリケーションがアクティブである場合に展開されるマーキングメニューの内容を示す．

4 実装

我々は 3.2 節にて述べたシステムデーモンを Windows 8 タブレット上に C# .NET を用いて実装した．ベゼル上におけるダブルクロッシングジェスチャを検出するために，我々は画面端にタッチ検出を行う半透明ウィンドウを配置した．このシステムデーモンは，常にアクティブアプリケーションを監視する．

表 **1.** 各アプリケーションがアクティブである場合に展開されるマーキングメニューの内容

アプリケーション	マーキングメニューの内容
テキストエディタ	ペースト/テキスト検索
Web ブラウザ	ペースト/テキスト検索 URL を開く/Web 検索
PDF ビューア	テキスト検索

アクティブアプリケーションと仮想クリップボードとの間においてデータを送受信するために，システムデーモンはキーイベントを送り， OS から提供されるクリップボードを中継地点として使用する．例えば，ユーザがテキストエディタをアクティブにし Bezel Check を実行すると，システムデーモンは選択されたテキストを OS から提供されるクリップボードにコピーするために Ctrl+C のキーイベントを送る．その後，システムデーモンは OS から提供されるクリップボードから Windows API を用いてコピーされたテキストを取得する．また，ベゼル部分に保持されたテキストをエディタ上にペーストする際には，システムデーモンは Windows API を用いてそのテキストを OS から提供されるクリップボードにコピーし，その後 Ctrl+V キーイベントを送る．

5 まとめと今後の予定

本稿では，携帯型タッチスクリーン端末において 1 ストロークの動作にて素早くデータを空間的に配置する事が可能となるジェスチャである Bezel Check を示した．Bezel Check はピンチやスワイプ等の従来のタッチジェスチャと競合せず，これらと共存する事が出来る．

今後は，画像や一般ファイル等のテキスト以外のデータへの対応と，操作速度やエラー率等の定量的評価，また使用感等の定性的評価を行う．

参考文献

[1] A. Bragdon, E. Nelson, Y. Li, and K. Hinckley. Experimental analysis of touch-screen gesture designs in mobile environments. CHI ’11, pp. 403–412.

[2] K. Hinckley, J. Pierce, M. Sinclair, and E. Horvitz. Sensing techniques for mobile interaction. UIST ’00, pp. 91–100.

[3] S. Kim, J. Yu, and G. Lee. Interaction techniques for unreachable objects on the touchscreen. OzCHI ’12, pp. 295–298.

[4] V. Roth and T. Turner. Bezel swipe: conflict-free scrolling and multiple selection on mobile touch screen devices. CHI ’09, pp. 1523–1526.

[5] M. Serrano, E. Lecolinet, and Y. Guiard. Bezel-Tap gestures: quick activation of commands from sleep mode on tablets. CHI ’13, pp. 3027–3036.

スマートフォンと MAD-SS を用いた長距離センサーノードの開発

内田 泰広　瀬川 典久　澤本 潤　杉野 栄二　矢澤 正人　後閑 政昭*

概要.　センサネットワークで広域なネットワークを構築するには、一般的に利用されているセンサでは通信距離が短いため多くのセンサノードの設置が必要となり長距離通信可能なノードが必要である。また、山間部でのセンサネットワークインフラの構築では、周囲の環境に強く、電源が大容量バッテリーを持つノードが必要不可欠となる。本研究では広域なセンサネットワークを構築することを目的とし、数理設計研究所が開発した長距離通信可能なスペクトラム拡散通信技術を耐久性のあるスマートフォンを組み込み機として実装を行った。

1　はじめに

センサネットワークの研究が進化するにあたって、屋内などの限られたエリアから、屋外たとえば里山などの広域なエリアでのセンサネットワークの活用が考えられてきている。センサネットワークで屋外、広範囲での情報取得の需要として、線路の落石情報や、動物の行動パターンを監視したいというものがある。一般的なセンサネットワークを里山などで利用する場合、通信距離が 100m、長くても 1km 程度であり、多くのセンサノードが必要となる
また、地形の影響により通信距離が短くなるという問題点があげられる [1]。

里山でのセンサノードの設置は電源の確保が困難であったり、金銭的コストや時間が多くかかってしまう。これより里山では少ないセンサノード数で、広範囲な情報を取得できることが理想である [2]。

本研究では数理設計研究所が開発した MAD-SS を活用し、長距離通信可能なセンサノードを開発することを提案する。過去に MAD-SS は、DSP を用いた実装、ARM Cortex-M4F での実装が行われてきた。受信側では計算能力が求められるため、PC、ARM マイコンで行われる。今後、さらなる広範囲での情報取得のため、地形の影響を受けない高い高度に一時的に簡易な気球などに乗せる時、耐久性のあるセンサノードが必要となる。そこで、スマートフォンを利用した長距離通信可能なセンサノードの提案を行う。

Copyright is held by the author(s).

* Yasuhiro Uchida, 岩手県立大学 ソフトウェア情報学部 ソフトウェア情報学科, Norihisa Segawa, Jun Sawamoto, and Eiji Sugino, 岩手県立大学,Masato Yazawa,Masaaki Gokan

2　MAD-SS

2.1　MAD-SS の特徴

MAD-SS は、数理設計研究所が開発した「スペクトラム拡散通信の高速同期法」による長距離通信可能な技術である。送信方法は一般的なスペクトラム拡散通信の一つである直接拡散方式であり、受信側で高速で受信するために特殊な方法を用いている。

通信速度は (10bit/sec) であるが、超低出力で長距離通信可能である。微弱電波を使用しているため、免許不要で運用することができる。通信速度は低速であるがセンサネットワークで扱う情報量は潮位や温度、湿度などで問題にはならない。

免許不要時の超低出力 (50nW) で MAD-SS を用いた場合の通信距離と、免許を受けた場合に運用できる MAD-SS の通信距離を表 1、表 2 に示す。里山で十分な電波の受信を行えるために高度でも動作可能なスマートフォンに MAD-SS を実装することを試みている [3]。

表 **1. MAD-SS** を用いた場合の通信距離 **(50nW)**

送信アンテナ高さ	受信アンテナ高さ	周波数	距離
1.5m	1.5m	150MHz	0.5km
6.0m	1.5m	150MHz	1.2km

表 **2. MAD-SS** を用いた場合の通信距離 **(10mW)**

送信アンテナ高さ	受信アンテナ高さ	周波数	距離
1.5m	1.5m	150MHz	7km
6.0m	1.5m	150MHz	14km

2.2 MAD-SS の受信手続き

従来のスペクトラム拡散通信の受信方法であるスライド法を用いた場合、同期を確立するために時間がかかってしまう。MAD-SS では受信を高速に行えるように工夫がなされている。以下に MAD-SS の受信方法（高速同期法）の手順を示す [5]。

1. 変調符号が変化する位置（トグル点）を検出
2. 拡散コードを一周期分のトグル点をあらかじめ生成しておき、1 と相互相関させて拡散コードのズレ量を推定
3. 2. で求めたズレ量を用いて復調（逆拡散）
4. 逆拡散した搬送波を確認し、受信したスペクトラムが有効なら取得

従来のスライド法では、拡散符号をスライドさせつつ逆拡散を行う。一方、MAD-SS の高速同期法では同期位置の推定をしてから、逆拡散を行う。MAD-SS の受信アルゴリズムの一つとして高速フーリエ変換 (FFT) があり、負荷がかかる処理は DSP を用いて、省電力、小型化を目的とされた研究がされてきた。

3 スマートフォンを用いた長距離センサノード

3.1 Android NDK

Android NDK は、Android SDK と利用される開発キットである。Android アプリケーションの一部をネイティブコードで作成が可能で既存の C/C++のコードを再利用することができる。Android NDK のライブラリを使用することで Android の Application framework を介さずに、Android の内部構成の最下層である Linux のレイヤーの機能を使えるため、Java で実行するより高速となる [4]。

3.2 システムについて

本システムでは、処理の高速化、既に開発された (C/C++) のソースコードの一部を用いることが可能である Android NDK を使い JNI(Java Native Interface) を利用した。Cygwin を用いることで (C/C++) のソースコードはビルドを行い共有モジュール (*.so) を生成し、Java 側で作成したネイティブのライブラリをロードする。これにより、android アプリからネイティブコードを通常のメソッドとして利用可能である。

実装においては、電波の受信部を Java、MAD-SS の高速同期法はネイティブコード (C/C++) を用いる。MAD-SS の中核となる高速フーリエ (FFT) はネイティブコードで行う。

4 まとめとの課題について

本研究では MAD-SS の受信部をスマートフォンに移植を行い、自然環境に強く、大容量バッテリーを搭載しているセンサノードの開発について提案している。今後として、里山や大気圏まで飛ぶ簡易な気球上 (図 1) などで正常に受信を行えるかを実験を行い、センサノードとしての有用性を調べる必要がある。長距離通信可能なセンサノードを開発することで里山などで必要とされる情報を幅広く取得できることを目的としている。

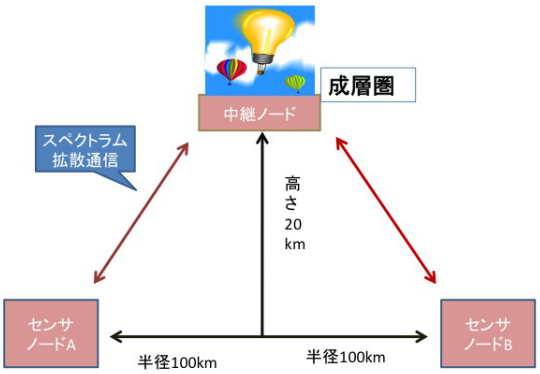


図 1. MAD-SS の受信機を気球に乗せたイメージ例

参考文献

[1] 瀬川典久, 澤本潤, 松原和衛, 出口善隆, 大石明広, 山本信次, 東淳樹, 青井俊樹:里山での活動を支援するセンサネットワーク環境構築の提案, 電子情報通信学会技術研究報告. USN, ユビキタス・センサネットワーク, pp.141-146 , (2009)

[2] 柏田師宏, 瀬川典久, 澤本潤, 玉置晴朗, 矢澤正人:MAD-SS と DSP を用いた長距離通信可能なセンサーノードの開発, ソフトウェア情報学研究科修論,(2012)

[3] 澤田直哉, 瀬川典久, 澤本潤, 杉野栄二, 玉置晴朗, 矢澤正人, 後閑政昭: Coretex-M4F と MAD-SS を使用した長距離センサーノードの開発,WISS2012,(2012)

[4] AndroidNDK. http://developer.android.com/tools/sdk/ndk/index.html

[5] 数理設計研究所. http://www.madlabo.com/ss/.

暗記学習のための文節とモーラ数に基づいた替え歌自動生成システム

伊藤 悠真　　寺田 努　　塚本 昌彦 *

概要. 情報処理技術が進展し，さまざまな学習支援システムが開発されている．特に，学習者の暗記を支援するための一般的な手法としては，元の文字列を他の意味のある文字列に置き換える語呂合わせがよく用いられ，その語呂合わせを自動で生成するシステムが存在する．しかしこれらのシステムには文字数や暗記対象に制限があり，汎用的に暗記学習に使えるとはいえない．一方で，暗記学習のために替え歌を用いるという手法は，語呂合わせと同様に広く知れ渡っている手法であるが，替え歌を生成する一般的な手法は筆者らの知る限り確立されておらず，学習者が暗記したい項目に対し自由に替え歌を生成することができないという問題がある．したがって本研究では暗記学習のための替え歌自動生成システムの構築を目的とする．提案手法では，学習者自身の知っている楽曲からなる楽曲データベースに対し，学習者が暗記したい項目を入力すると，楽曲データベース中のいくつかの曲の歌詞に暗記したい項目が割り当てられて替え歌として出力される．本研究では，この割り当てに楽曲のリズムと暗記項目のモーラ数を用いる手法を提案した．

1　はじめに

情報処理技術が進展し，さまざまな学習支援システムが開発されている．特に，学習者の暗記学習を支援するための研究例として元の単語の羅列を頭文字を繋げるなどして他の文字列に置換する語呂合わせを自動で生成するというものがある [1]．これらの研究では，4 個から 8 個の単語に対し語呂合わせを生成しているが，多数の単語で語呂合わせを生成した場合，生成結果が不自然になるという問題がある．

一方，暗記したい項目を既存の楽曲のメロディに乗せて暗記する替え歌による手法は語呂合わせと同様に広く知れ渡っている手法であり，有用性が示されている [2]．しかし，暗記用の替え歌を生成する手法は筆者らの知る限り確立されておらず，学習者が暗記したい項目に対し自由に替え歌を作成できないという問題がある．

そこで本研究では暗記用の替え歌自動生成システムを構築する．提案手法では，替え歌の元となる楽曲データベースを保持し，学習者が暗記したい単語のリスト（以下暗記リストとよぶ）を入力すると，楽曲データベース中のいくつかの曲の歌詞に暗記リストが割り当てられて出力される．本研究では，歌詞の文節と歌詞のリズムに着目し割り当ておよび割り当て時の変更度合いを決定している．

2　設計

図 1 に提案システムの処理の流れを示す．まず，図 1 に示すようにシステムは楽曲が登録された楽曲データベースを保持しており，データベース中の楽曲から替え歌が作られる．文献 [2] より，既に記憶している楽曲に対し替え歌を生成することが重要であるため，本研究では楽曲データベースはユーザ個人が記憶している楽曲や好きな楽曲で構成されている．学習者が暗記リストをシステムに入力すると，システムはデータベース中の各楽曲と暗記リストで音符数などの条件を元に替え歌が生成可能かどうかを判断する．生成可能であれば替え歌リストに生成結果を追加し替え歌として出力する．さらに，文献 [2] より，生成された替え歌が既に記憶している楽曲という条件を満たすために，替え歌の元となる楽曲と生成された替え歌の間に大きな差異がでないように設計する必要がある．本研究では各替え歌に対し「変更度」という言葉を定義し，暗記リストを楽曲の歌詞に割り当てる際に元の楽曲をどれだけ変更したかの値を定量化することで，小さいものから順に元楽曲からの変化が小さい自然な替え歌として出力する．

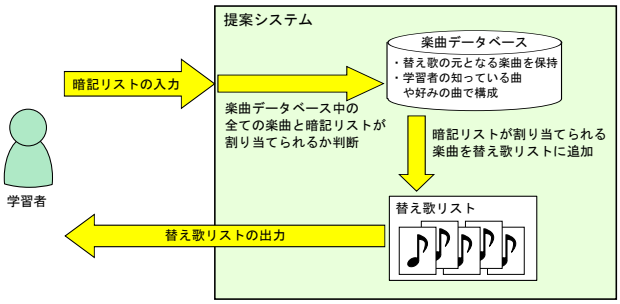


図 1. 提案システムの処理の流れ

本研究では楽曲の歌詞の文節ごとに暗記リストの割当てを行う．原則として，各文節に含まれる音符

Copyright is held by the author(s).

* Yuma Ito and Masahiko Tsukamoto, 神戸大学大学院工学研究科, Tsutomu Terada , 神戸大学大学院工学研究科 / 科学技術振興機構さきがけ

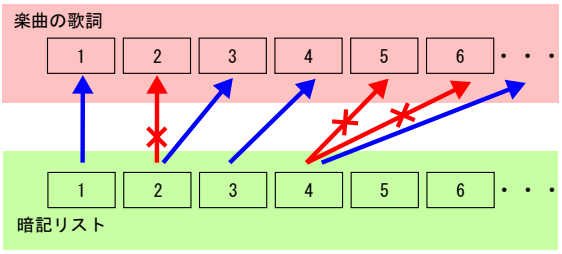


図 2. 暗記リストの割当て例

列に対し暗記リストの 1 項目のモーラを均等に割り当て，音符数とモーラ数の差によって変更度を定義する．ここで，モーラとは日本語の拍と呼ばれるもので，通常 1 対の子音と母音から成る．例として「チョコレート」をモーラで分けると「チョ，コ，レ，ー，ト」となり，モーラ数は 5 となる．また，音符数が足りない場合，音符列中で最長の音符を分割することで音符数を増やす．音符を分割する場合，16 分音符より短い音符では割り当てられたモーラが聞き取りにくくなるので，これ以上は分割しないとする．よって全ての音符が 16 分音符になっても割り当てる音符が足らない場合を「割当て不可能」とし，それ以外を「割当て可能」とする．

次に，楽曲全体での文節および暗記リストの割り当て方を図 2 を用いて説明する．図 2 上部の矩形は楽曲の歌詞の文節を表し，下部の矩形は暗記リストの各項目を表す．さらに，青色矢印は暗記項目がその矢印の指す文節に対し割当て可能であることを表し，赤色矢印は割当て不可能であることを表す．

図 2 に示すように，提案システムは楽曲の歌詞の最初の文節から順に暗記リストの各項目が割当て可能かどうかを判断する．割当て可能の場合は，次の文節および次の暗記リストで割当て可能かを判断し，割当て不可能の場合は，次の文節で割当て可能かどうかを判断する．このようにして，最後の文節もしくは最後の暗記リストでの割当て可能かどうかが判断されるまで処理を続け，最終的にこの楽曲を替え歌リストに加えるかどうかを以下のように判断する．

- 暗記リストが全て歌詞の文節に割り当てられた場合
  割当てられなかった文節は，それらの全ての音符に対し「ラ」を割り当てることとし，この楽曲を替え歌リストに追加する．このとき，楽曲全体の変更度を文節ごとでの変更度の総和で定義する．
- 暗記リストが全て歌詞の文節に割り当てられなかった場合
  この場合はこの暗記リストと楽曲では替え歌は生成できないとし，現在の曲を替え歌リストには追加しない．



図 3. プロトタイプシステムの概観

3 実装

2 章で述べた替え歌自動生成システムのプロトタイプを実装した．図 3 にプロトタイプシステムの概観を示す．図 3 に示すウインドウでは指定した暗記リストがどのように楽曲の歌詞に割り当てられているかを見ることができる．さらに，上部のコンボボックスには替え歌リストが格納されており，この選択を切り替えることで，ユーザはどの楽曲での替え歌を暗記に用いるか選ぶことができる．また，左上の Export ボタンにより musicXML 形式のファイルを出力することができ，このファイルを楽譜作成ソフトや歌声合成ソフトウェアに入力することで生成結果を楽譜や音声に変換できる．

4 まとめ

本研究では暗記学習のための替え歌自動生成システムの設計と実装を行った．提案システムは楽曲中の歌詞の文節ごとに，歌詞のリズムと暗記リストのモーラ数を考慮して暗記項目を割り当てることで，替え歌生成を行っている．

今後の課題としては，暗記リストを替え歌の有無それぞれで実際に被験者に暗記させ，暗記学習に対して提案手法が有効であるかの評価実験が考えられる．また，モーラ数や文節の音符数によっては複数の文節にまたがって単語を割り当てる手法も検討する必要がある．

参考文献

[1] 岡安優弥ほか: 品詞による文評価を用いた日本語語呂自動生成手法, 情報処理学会創立 50 周年記念 (第 72 回) 全国大会, pp. 511 – 512 (2010).

[2] R. Hale-Evans: Mind Performance Hacks: Tips & Tools for Overclocking Your Brain, *O'Reilly Media* (2006). (夏目 大訳: MIND パフォーマンス HACKS-脳と心のユーザーマニュアル, 株式会社オライリー・ジャパン (2007)).

軟らかくてよく弾むセンサ内蔵ボールの開発

佐藤 俊樹 *

概要. 本研究では、ドッジボール等の球技の拡張に用いることが可能な軟らかくてよく弾むセンサ内蔵ボールを提案する。ボールはボール外皮、内皮、センサユニット、センサコア及び固定用バルーンから成り、各種センサが格納されたセンサコアは、コア上に埋め込まれた固定用バルーンによりボール中心に柔らかく固定される。これにより、ボールの軟らかさと弾力性を維持しつつ、落下時の衝撃や外からの加重に対して高い耐久性を持つセンサ内蔵ボールを実現することが可能になる。本論文ではこれらの機構と実際に開発したプロトタイプについて述べ、現時点の課題と今後の展望について述べる。

1 はじめに

近年デジタル技術を用いたスポーツ拡張に注目が集まっている。従来のスポーツに各種センサや無線技術を組み合わせることで既存スポーツのルール拡張や全く新しいスポーツやエクササイズ、身体的要素を用いたゲームの創造が活発に行われるようになってきた。スポーツ拡張技術の中でも、特にボールの拡張はこれまで最も活発に行われてきており、企業の製品やコンセプトモデルを含め様々なセンサ内蔵ボールが存在する [1][3]。

実際にスポーツやエクササイズ、遊戯で実用可能なセンサボールを開発するためにはいくつかの技術的課題がある。まず、内蔵するセンサやマイコン等がボールに加わる強い衝撃で壊れないようにする必要がある。次に、センサ等を内蔵することでボールが必要以上に重くなったり、弾まなくなったりすることも避けなければならない。また、内蔵するデバイスによる重心の偏りや、形状の変形も抑えなければならない。そこで本研究では、これらの問題を考慮した新しいセンサ内蔵ボールを提案する。

1.1 提案と実装

本研究では図 1 のようなバルーン機構を用いたセンサ内蔵ボールを開発した。この機構の大きな特徴は、固定機構が軽量であること、衝撃や加重に対する耐久性が高いことである。さらにボール内の空気による弾力性を損なわず、安価に実現可能であることも挙げられる。以下では、実際に開発したプロトタイプシステムについて述べながら本機構の詳細を説明し、最後に今後の展望について述べる。

1.2 ハードウェア

提案するボールは、ボール外皮、ボール内皮、センサコア、コア固定用バルーン、及びセンサユニットから成る。センサユニットが格納されるセンサコアには、4 つの固定用バルーンユニットが固定されている (図 2(右))。バルーンユニットはコア上に均等に分散配置されており、これらに均等に圧縮空気を送り込むことで膨らまし、ボール中心にコアを固定する。これと同時に、一般的な空気式ボールと同様に、コアを包むボールの内皮にも空気を送り込みボール全体を膨らませる。

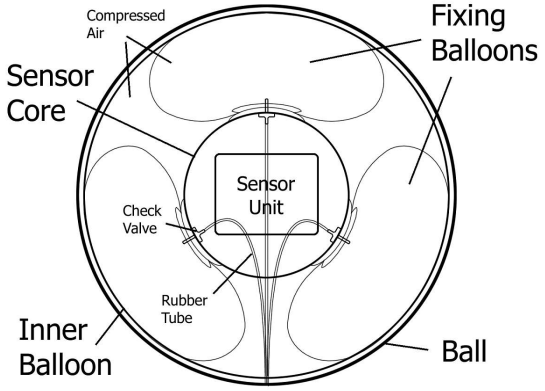


図 1. ボール内機構

開発したプロトタイプボールを図 2(左) に示す。今回のプロトタイプ実装においては、外皮として 7 インチの無色透明ゴムボール、内皮として 16 インチの風船、コア固定用バルーンとして 9 インチの風船を用いた。また、センサコアは 3D プリンタを用いて ABS 樹脂製で作成した。固定用バルーンに空気を入れるためには、それぞれのバルーンに接続した逆支弁付きのチューブを用いた。このチューブは、コア内部を通り内皮用バルーンの口に集中させ、内皮用バルーンと同時に空気を送り込めるようにした。プロトタイプシステムでは入手製、加工性を考慮し市販の風船を用いたが、ボールのバウンド、壁への投げつけ等による衝撃を加えても割れることは無く、

Copyright is held by the author(s).

* 電気通信大学

内蔵しているセンサが壊れることもなかった。また弾み方も一般的なゴムボールとほぼ同等であった。外皮にはコアを収納する際の切れ込みを入れる必要があったが、内皮さえ閉じられていれば、ビニールテープ等での簡単な切れ込みの補強のみで投げて遊ぶだけの十分な強度が得られた。



図 2. プロトタイプボール (左) 及びコア (右)

1.3 内蔵センサユニット

これらの詳細な構成を図 1 に示す。次に、ボールに内蔵するセンサユニットの開発を行った。センサユニットはマイクロコントローラ、無線モジュール、センサ及びバッテリから構成される (図 3)。今回の実装では、マイクロコントローラ及び無線モジュールにケイツー電子製 WCU-C2543 μ を用いた。またセンサは加速度センサ (ADXL345)、ジャイロセンサ (ITG3200) に加え、圧力センサ (MPL115A2) も内蔵した。提案ボールはセンサが完全に密閉されているため、圧力センサを搭載することでボール内の空気圧の変化を測定可能である。これにより加速度センサでは計測が困難なボールに加わる強い衝撃を検出したり、ボールを「触る」・「握る」等のボールに外力を加える要素を持つプレイヤーの身体動作を検出することも可能になると思われる。これらのボール内センサを用いた検出技術の開発については引き続き行っていく予定である。

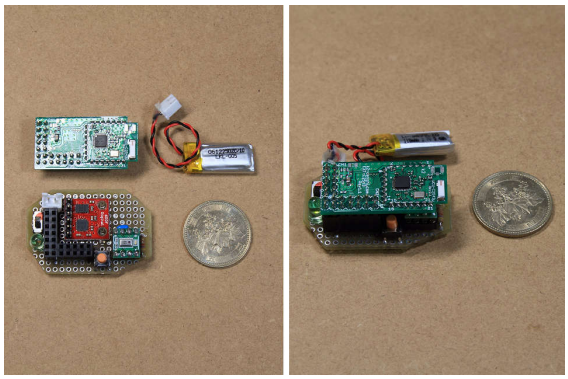


図 3. センサユニット (右) とその構成 (左)

2 課題と展望

現時点の課題としては、衝撃によりコアが中心から若干ずれてしまう問題がある。そこでコアと固定用バルーンをより強く固定できるような固定方法を開発したいと考えている。次に、コアをボール中心に固定するためには、固定用バルーンの大きさを均一にする必要があり、また固定用バルーンの空気圧が高過ぎるとボールが変形してしまう。しかし現在は固定用バルーンに逆止弁を装着しているため空気の調節が困難であるという問題がある。そこでバルーンの空気圧調節を容易にするためのバルブ機構の開発も行いたいと考えている。またバルーンの形状や素材についても実験、調査を行い、より適したものに改良していきたいと考えている。

今後の展望としては、実際に実用し改良及び性能の評価を行うと同時に、ボールを用いたスポーツ格調技術についても開発を行いたいと考えている。

3 関連研究

センサ内蔵ボール研究の代表的なものとして、児玉らの跳ね星がある [2]。跳ね星はセンサ類を透明ゴムで覆うことで保護しているが、ゴム層が厚くなるとボール自体が重くなるという欠点がある。次に、センサ内蔵のトレーニング用バスケットボール 94Fifty[3] が開発されている。94Fifty では、センサモジュールは非常に小型でありバルブ部分に装着されているが、加速度センサ等がボール回転時に発生する遠心力の影響を受けるため、センサはボール中心に位置した方が理想的と考える。また、Adidas の Smart Ball[1] のように柔軟性のある紐 (タイ) 状構造物を用いた固定方法も存在する。本研究で用いているコア装着式バルーンを用いた手法は、ボールの中心にセンサを固定するより安価で容易な手法であり、ボール内蔵センサ開発のためのプロトタイピングにも適していると考えている。最も手軽な手法として発砲ウレタン製ボールを切り開いてセンサを内臓する手法もあるが、空気式のボールと比較して弾力性に劣り、重さや手触りも独特であるという問題がある。

参考文献

[1] addidas. addidas Smart Ball, May 2013.

[2] O. Izuta, J. Nakamura, T. Sato, S. Kodama, H. Koike, K. Fukuchi, K. Shibasaki, and H. Mamiya. Digital sports using the "Bouncing Star" rubber ball comprising IR and full-color LEDs and an acceleration sensor. In *ACM SIGGRAPH 2008 new tech demos*, SIGGRAPH '08, pp. 13:1–13:1, New York, NY, USA, 2008. ACM.

[3] I. S. Technologies. 94Fifty Basketball, March 2013.

Overview Scrollbar: ページ全体を概観として表示するスクロールバー

溝口 晃　　坂本 大介　　五十嵐 健夫*

概要. 画面に収まり切らない縦長のドキュメントを表示する際，多くのアプリケーションは画面の端にスクロールバーを表示する．しかし，現在使用されているスクロールバーの殆どは古典的なものであり，改善の余地が多分に存在する．本研究ではひとつの改善策として，スクロールバー上にページ全体の概観を表示する手法，Overview Scrollbar を提案する．このような手法を取った時，縦長のページを閲覧する場合に概観をどのように縮小して表示するのかが問題となってくる．なぜなら縦長ドキュメントのアスペクト比を維持しつつ概観を表示しようとすると，その横幅が非常に細くなってしまうからである．我々はドキュメントの長さに関わらず概観の横幅を一定に保って可読性を維持するために，部分的に縮小率を変化させて概観を描画する手法を新たに提案し，Overview Scrollbar の性能や問題点を調査するためにユーザスタディを実施した．



図 1. ドキュメントと Overview Scrollbar



図 2. (a)Uniform Scaling (b)Vertical Compression (c)Fisheye (d)Perspective

1 はじめに

現在使用されているスクロールバーの多くは，ドキュメントの表示領域を狭めないようにウィンドウの端に細長く表示されるようになっている．しかし，近年ディスプレイの解像度が上昇し，また横長のディスプレイが増えてきているため、必ずしもスクロールバーを細長く表示する必要は無くなってきている．そこで我々は，ページ全体の概観を縮小して表示する幅の広いスクロールバー，Overview Scrollbar を提案してきている (図 1)[2].

Reader's Helper[1] や Popout Prism[3] では我々の Overview Scrollbar と似たインタフェースが実装されているが，これらはドキュメントが縦長の場合にどう概観を縮小するのかという問題について深く言及していない．そこで我々は，単純な 2 種（Uniform Scaling, Vertical Compression）に加え，新たな縮小方法を取り入れた 2 種のスクロールバー（Fisheye, Perspective）を実装し，計 4 種の Overview Scrollbar の性能や問題点を調査するためにユーザスタディを実施した．

2 提案手法

我々は，概観の縮小方法別に 4 種類のスクロールバー（Uniform Scaling, Vertical Compression, Fisheye, Perspective）を設計した (図 2). 以下に，それぞれのインタフェースがどのようにドキュメントを縮小して概観を表示するか述べる．

(a)Uniform Scaling: ドキュメントのアスペクト比を維持して縮小する．そのため，ドキュメントが縦長の場合にスクロールバーが細くなってしまう事がある．

(b)Vertical Compression: 基本的にはドキュ

Copyright is held by the author(s).

* Ko Mizoguchi, Daisuke Sakamoto, and Takeo Igarashi
東京大学大学院情報理工学系研究科

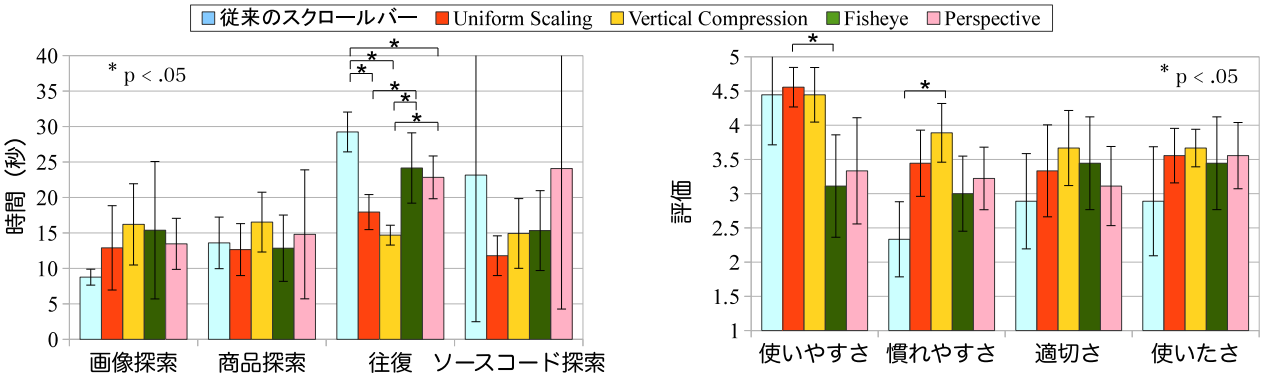


図 3. 実験結果: (左) タスクにかかった時間の平均 (右) アンケート結果の平均

メントのアスペクト比を維持するが，横幅はウィンドウの 20%に固定されており，ドキュメントが一定以上縦長の場合，縦方向にのみ縮小する．

(c)Fisheye: スクロールバー上に焦点領域と背景領域が存在する．焦点領域ではアスペクト比を維持して概観を比較的大きく表示する一方で，背景領域では縮小率を徐々に上げて行き，魚眼レンズのような視覚効果を出す．ユーザはスクロールバー上で右クリックすることで，焦点領域を自由に動かすことが出来る．

(d) Perspective: 基本的には Fisheye と同様のインタフェースとなっているが，Perspective では全体を通して横幅が一定で，ウィンドウの横幅の 20%に固定している．

3 ユーザスタディ

Overview Scrollbar の性能や問題点を調査するためにユーザスタディを実施した．参加者は 12 名で全員従来のスクロールバーを使用した経験があった．スタディを通して，参加者がそれぞれのタスクにかかった時間を計測した．

3.1 タスク

大きなドキュメントの中から，ある対象を探してクリックするタスクを 4 種類設計した．

(1) 画像探索: 参加者は初めに探す対象となる画像を提示され，その画像をドキュメント内から探してクリックする．

(2) 商品探索: 参加者は初めに探す対象となる商品を指定され，その商品をドキュメント内から探してクリックする．商品は色，形などの外見上の特徴を文章によって指定される．

(3) 往復: 参加者は大きなドキュメント上に設置された 2 つのリンクを，上下にスクロールしながら交互に合計 6 回クリックする．

(4) ソースコード探索: 1 つの変数がソースコード中の全ての箇所でハイライトされており，参加者はその中から指定された条件に合う変数を探してクリックする．

3.2 アンケート

アンケートを通じて参加者の好みを収集した．使いやすさ・使い慣れやすさ・適切さ・使いたさをそれぞれ 5 段階で評価してもらい，感想・印象などの任意のコメントも収集した．

4 結果とまとめ

実験結果を図 3 に示す．タスクにかかった時間では往復タスクに有意な差が確認され，従来のスクロールバーが最も遅く，Vertical Compression が一番早かった．またアンケート結果では使いやすさでは Fisheye，Perspective が低く評価され，慣れやすさでは Vertical Compression が高く評価される傾向にあった．

全体を通して高性能だった Vertical Compression は既に実用的であると考えられる．Fisheye や Perspective の性能はそれほど高くなかったが，焦点領域を移動するインタフェースをユーザにとってより直感的にすることで，性能が上昇する余地は存在すると考える．

参考文献

[1] J. Graham. The reader's helper: a personalized document reading environment. In *Proc. of CHI '99*, pp. 481–488. ACM, 1999.

[2] K. Mizoguchi, D. Sakamoto, and T. Igarashi. Overview Scrollbar: A Scrollbar Showing an Entire Document as an Overview. In *Proc. of INTERACT '13*, pp. 603–610. Springer Berlin Heidelberg, 2013.

[3] B. Suh, A. Woodruff, R. Rosenholtz, and A. Glass. Popout prism: adding perceptual principles to overview+detail document interfaces. In *Proc. of CHI '02*, pp. 251–258. ACM, 2002.

行動ログの入力を支援するユーザインタフェースの設計と実装

鈴木 孝宏* 美馬 義亮†

概要. 本稿では，スマートフォンを利用した行動ログの入力作業を支援するユーザインタフェースを提案し，それを用いたプロトタイプシステムのアイディアを示す．ここで提案するユーザインタフェースは，現実世界に存在する様々な状況を一般化して二次元平面に構成したものであり，その平面をユーザがタップすることによって行動ログの入力を可能にするものである．このユーザインタフェースを用いたプロトタイプシステムでは，ユーザのタップ操作をトリガーに行動を報告するテキストを自動的に生成し，行動を適切に言語化して記録可能にする．これにより，生成されたテキストを日記に使用したり，Twitter に投稿してライフログとするなど，行動記録を目的とした文章作成・入力にかかる負担を軽減できる可能性がある．

1 はじめに

本稿では，スマートフォンを利用した行動ログの入力作業を支援するユーザインタフェースを提案する．ここで提案するものは，現実世界に存在する様々な状況を二次元平面に構成したものであり，それをユーザがタップすることによって行動ログの入力を可能にするものである．このユーザインタフェースを用いて実装したプロトタイプシステムでは，ユーザのタップ操作をトリガーに行動を報告するテキストを自動的に生成し，行動を適切に言語化して記録可能にする．これにより，日記の文章作成が容易になったり，Twitter に投稿してライフログとするなど，行動記録を目的とした文章作成・入力にかかる負担を軽減できると予想している．

1.1 スマートフォンを利用したライフログ

ライフログは，日常生活のすべてをデジタルデータで記録することである．デジタルで記録することにより，いつでも過去に遡って記憶を取り出せるメリットがある．記録のテーマは行動を書き留めた日記にとどまらず，日々の食事，訪れた場所，体調管理，育児記録などと多岐にわたる．記録のためのツールは，カメラや GPS 機能の付いたスマートフォンであることが多い．スマートフォンではライフログを目的としたアプリケーションも多く存在し，人気の具合を知ることができる．

本稿で扱う行動ログは，日々の行動のすべてを記録するライフログであり，日記に近い．しかしながら，ログを記録しようとする度にテキスト入力や位置情報の入力などユーザの操作が必要で，記録にかかる負担が大きい．センサ情報を用いてログを自動記録するアプリケーションや，記録するテーマに合わせた独自のユーザインタフェースを備えたものも存在するが，記録がセンサ情報に限定されたり，言語化して記録できるものはまだ少なかったりと課題がある [1]．このことから，行動記録に必要な操作ステップを減らし，テキストも自動生成することにより，記録にかかる負担を軽減できると考える．

本稿で提案するプロトタイプシステムでは，スーパーマーケットでの行動記録を対象とした．行動記録を支援するためのユーザインタフェースを用い，少ない操作ステップでユーザの行動を適切に言語化する．このことにより，行動ログの記録にかかる負担の軽減，テキストとして記録する際の文章作成の負担を軽減する可能性がある．

2 行動ログの入力を支援するインタフェース

本稿で提案するインタフェースは，現実世界に存在する状況を二次元平面に構成し，タップ操作によって行動ログを記録できるようにしたものである．図 1 左の例では，スーパーマーケットに買い物に訪れた際に，どんな買い物をしたかを記録する用途に特化させたインタフェースを用いたプロトタイプシステムである．



図 1. スーパーマーケット店内のイラストによるインタフェース（左）と，パラメータ設定画面（右）

Copyright is held by the author(s).

* 公立はこだて未来大学大学院システム情報科学研究科
† 公立はこだて未来大学システム情報科学部

2.1 インタフェースの設計

図 1 左で示したインタフェースは，スーパーマーケットの店舗全体の配置を表現している．これは，特定のスーパーマーケット店内を意図したものではなく，一般の店舗に共通して存在するコーナーを配置したものである．このことは，ユーザの持っている一般的な店舗イメージを反映していると言え，また特定の店舗に依存しない行動ログの記録を可能にしている．

画面最下部には 3 つのテキスト (品名表示領域) を配置した。ここに表示された 3 つの品名から 1 つをユーザが選択することで，購入する品物の数や大きさを決める画面（図 1 右）に遷移する．その画面で各パラメータを設定することで，行動ログとなる文書が自動的に生成される．

2.2 文章の生成

行動ログとなる文章は，予め用意したテンプレートに，品名・大きさ・数量を挿入することで生成している．今回のプロトタイプで生成する文章のテンプレートは,「(大きさ)(品名) を (数量) 買った。」という形式で用意した．これによって,「大きいじゃがいもを 4 つ買った。」などの文章を自動で生成することが可能である．

品名については，ユーザに提示する店内全体のイラストと対応する縦 10 ×横 10 のセルを持った商品名リストに用意してある（図 2）．店内イラストは拡大縮小して表示することが可能である．このため，品名表示領域に表示する 3 つのテキストは，画面に表示されているイラストの範囲に対応するリストからランダムにピックアップすることで表示している．

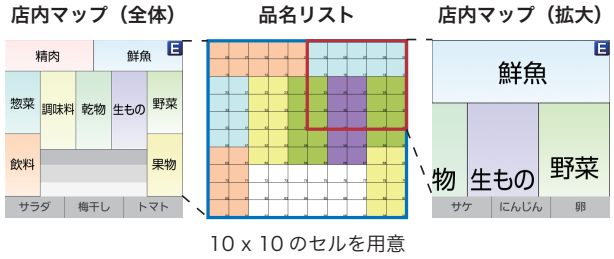


図 2. 店内のイラストマップと商品名リストの対応

3 今後の展望と課題

3.1 システムの特徴と可能性

本稿で提案した行動記録を支援するユーザインタフェースは，現実世界に存在する状況を二次元平面に構成するために，状況を一般化したイラストをインタフェースに用いた．これにより，特定の状況に依存せず，類似する様々な状況に対応できるものを設計できる可能性がある．このことは，ある状況に対してユーザの持っているイメージを崩さぬまま操作できる工夫であると言え，システム操作にかかる負担軽減にも繋がると考えている．

3.2 行動ログの負担軽減と記録できる行動の追加

ユーザの負担軽減のためには，ユーザの行動パターンを学習し，ユーザの次の行動を予測してユーザインタフェースを提示する必要がある．そのためには，スーパーマーケットだけでなく日常生活のあらゆる状況を二次元平面に構成し，感情も記録できるようにするなど，より多くの行動記録に対応することが課題である．記録できる行動が増えることで，ユーザの行動パターンを学習でき，その履歴から次に使用するインタフェースの予測が可能となる．

3.3 文章生成のアルゴリズムの検討

このシステムでは，行動記録となる文章のテンプレートと状況を関連づけているが，数多くの文章表現に対応するには工夫が必要である．予め用意する文章のテンプレートに多数の表現を持たせるには，正規表現で記述したテンプレートを用意する方法を用いることで，似たような意味や構造を持つ文章を大量に生成することができる．さらに，それらの文章をセンサ情報等をもとにフィルタリングすることで，より適切な文章生成が可能になると考えている．この考え方は，展開ヘルプ [2] で用いられた Generate and Filter という手法を参考にしている．

4 まとめ

本稿では，スマートフォンを利用した行動ログの入力作業を支援するユーザインタフェースの設計と，それを用いて実装したシステムのアイディアを提案した．このシステムは，ユーザの行動記録にかかるタップ操作を減らし，行動記録のテキストを自動的に生成する．システムが自動的に言語化することは，行動ログのための文章作成・入力にかかるユーザの負担軽減にも繋がる．記録できる状況・行動の種類が増えていけば，終日の行動記録をもとに日記を自動生成することも可能になり，日記執筆の負担軽減も可能になるのではないかと考えている．

謝辞

本稿における研究は IPA2013 年度未踏 IT 人材発掘・育成事業の支援のもとで行われている。

参考文献

[1] C. G. Bell, J. Gemmell, and 飯泉恵美子. ライフログのすすめ: 人生の「すべて」をデジタルに記録する!, pp.257-311, 早川書房, 2010.

[2] 増井俊之. 展開ヘルプ. インタラクション 2012 論文集, pp.89-96, March 2012.

ペイントアプリケーションにおけるタッチアクションを用いたエフェクトの提案

武藤 寛　井川 裕貴　瀬川 典久　杉野 栄二　澤本 潤*

概要. 近年タブレット端末が普及してきているとともに、タブレット端末用のアプリケーションを用いて絵を書く人が増えてきた。従来のペイントアプリケーションでは、画面上のキャンバスの上部や側部にペンの種類や色、太さを変更するためのツールバーが表示されている。しかし、これらの機能性が操作を複雑にし、自由な発想を妨げていた。そこで本研究では、シンプルなインターフェースを持つ enchantMOON を用いて、自由な発想を実現する操作方法として手を使った鉛筆エフェクトを提案する。

1 はじめに

本格的なイラスト描画を行うデジタルデバイスとしてペンタブレットや液晶タブレットが挙げられるが、近年タブレット端末が普及してきているとともに、タブレット端末用のアプリケーションを用いて絵を書く人が増えてきた。タブレット端末で絵が描ければ、改めて機材を買い揃える必要がなく、趣味や遊び目的で絵を描く人にも容易に環境が整う。

従来のペイントアプリケーションでは、インターフェースとしてウィンドウの中央に描画用のキャンバスがあり、上部や側部にツールバーというレイアウト構成がほとんどである。それらのツールを使って描いたものに対してエフェクトをかけるが、その操作は直感的なものではなく、使い方を覚える必要がある。これでは単純に絵を書くことが目的であるのにも関わらず、手間が増えてしまった。そして、この手間がユーザにストレスを感じさせていた。デッサンやアイディアスケッチの場合このようなストレスを感じることは発想の妨げになってしまう。

本研究ではシンプルなインターフェースを持つ enchantMOON を用いて、自由な発想を実現する操作方法として手を使った鉛筆エフェクトを提案する。具体的にはツールバーの代わりに手によるタッチイベントを用いる手法である。ここで手を使う意味として、現実世界で絵を描く際に紙と鉛筆を使うというシチュエーションを想定している。筆跡の周りをぼやかしたいときは紙を指で擦るようにデジタルデバイスも画面を擦ることでぼやけるのが理想だと考えた。

今回の研究で用いた enchantMOON は iPad や Android タブレットとは異なる斬新なインターフェースで話題を呼んだ。iPad や Android についても従来のパソコンと比べて直感的な操作が可能ではあるが、まだまだ使いこなせないという声も多く耳にする。そこで本研究の延長で鉛筆に限らずに様々なエフェクトについても現実世界に近いアクションで行うことにより、デジタルデバイスをアナログと同様な操作ができればユーザへの負担を解消できると考えられる。

Copyright is held by the author(s).

* 岩手県立大学ソフトウェア情報学部

2 関連研究

ペイントアプリケーションのツール選択のインターフェースとしてユーザが自由にレイアウトを変更できるシステム [1] や音声入力を用いる研究 [2] がある。音声入力を用いる手法ではオノマトペを用いることで、感覚的にツールを選択することができ、絵を描く事が楽しくなる。音声入力では場所を選んでしまうという欠点があった。本研究では手の動作を用いることでこの欠点を改良することができた。

3 enchantMOON

本研究で用いる enchantMOON[3]（図 1）とは、ユビキタスエンターテイメント社が開発したタブレット端末である。ほとんどの操作をデジタイザーペンによる手書きと指のタッチ（図 2）のみで行うシンプルなインターフェースである。筆跡ストロークのデータ構造は、X,Y の座標と筆圧の 3 つで構成されていて JSON 形式で保存される。ストロークはオブジェクトとして扱うことができ、さらに、JavaScript で記述したプログラムを機能として追加することができる。本研究の実装にあたってこのインターフェースとストロークに対する動作への機能追加の容易性から、enchantMOON が最適であると考えた。、

4 提案するシステム構成

4.1 システム概要

本システムは enchantMOON のシールを用いて機能を追加するので、JavaScript で実装される。以



図 1. enchantMOON 全体図



図 2. enchantMOON における指での指示例

下に動作フローを示す。図 3 に、本システムの動作を示す。

1. ペンで絵を描く。
2. タッチしてシールを実行。
3. 指で筆跡を擦る。
4. 擦った部分の筆跡がぼやける。

本システムは、ペンで書くという動作と擦るという動作を区別する。それぞれペンイベントとタッチイベントで異なる処理を施す。ペンイベントでは描画を行う。タッチイベントでは指が画面に接したタッチダウンイベント、そこから移動するタッチムーブイベント、指を離すまでのタッチアップイベントの 3 段階に分けられる。エフェクト処理はムーブイベントに対して行う。

エフェクト処理はプロトタイプのため筆跡ストロークデータの筆圧を下げることで、筆跡を薄くすることができた。

5 まとめと今後の課題

本研究では enchantMOON を用いて、自由な発想を実現する直感的な操作方法として手を使った鉛筆エフェクトを提案・試作を行った。これによりユーザは現実世界に近い紙と鉛筆を用いる感覚でデッサンやアイディアスケッチが可能となる。

今後はエフェクト処理について調査を進めて、より鉛筆に近い表現ができるように改良を加えていきたい。

参考文献

[1] 飛田博章 磯 大輔 暦本淳一 VelvetPath:スケッチとペイント操作によるレイアウトデザインシステム, 情報処理学会論文誌 44(11), 2528-2537, 2003-11-15

[2] 神原 啓介　塚田 浩二 オノマトペン 第 16 回インタラクティブシステムとソフトウェアに関するワークショップ (WISS2008), http://www.wiss.org/WISS2008Proceedings/papers/paper0023.pdf

[3] enchantMOON ; The Hypertext Authoring Tablet http://enchantmoon.com/

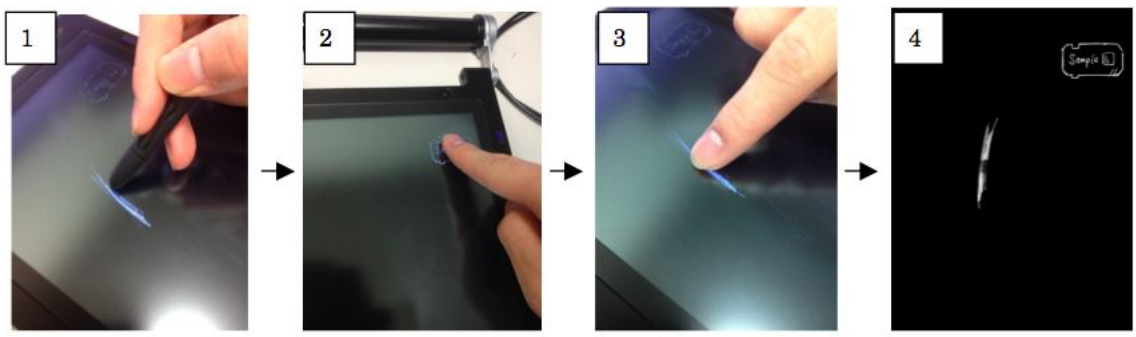


図 3. 本システムの動作フロー

デジタルコンテンツ生成のためのクラウドソーシングによるパラメタ最適化

小山 裕己　　坂本 大介　　五十嵐 健夫*

概要．写真補正，スライドの作成，CG 技術による画像生成など，多くのデジタルコンテンツ制作の過程には，パラメタ最適化，すなわち最適なパラメタセットを探索するという作業が含まれている．この作業は，人間の美的感覚に基づいて実行されるため，機械のみによって自動化することが困難である．本研究では，クラウドソーシング技術によるヒューマンコンピュテーションを用いることで，このようなパラメタ最適化を自動化する手法を提案する．我々の手法は，最適化のために必要なタスクを，2 つのパラメタセット間の相対的なランク付けという単純なマイクロタスクに分割するという点が特徴的である．

1　はじめに

CG 技術を用いて画像や映像を制作する際，制作者は各種パラメタの微調整に多くの時間を取られることが問題となっており，多くの研究者がこの問題に取り組んでいる．また CG 技術に限らず，一般のデジタルコンテンツ生成においてもこのようなパラメタ最適化の作業が多く見られ，専門家，一般ユーザ問わずこのような作業には多大な時間をとられてしまう．こうしたパラメタ最適化は機械による自動化が困難であることが多い．その理由は，パラメタセットの良し悪しの評価が人間の美的感覚に基づいていることが多いためであると考えられる．

そこで本研究では，特に近年注目されているマイクロタスク型のクラウドソーシングを用いたヒューマンコンピュテーションをシステムの内部で活用することにより，従来は困難であった美的感覚に基づくパラメタ最適化の自動化を行う手法を提案する．

2　関連研究

ヒューマンコンピュテーションとは，人間を明確に計算資源と見做して問題解決の手段として用いることである．またクラウドソーシングとは，不特定多数の人からなる群衆 (crowd) に対し，主にインターネットを通して業務を外部委託 (outsourcing) することである．近年，Amazon Mechanical Turk[1] などのサービスの登場により，例えば数秒から数分程度で終えることができるような非常に小さなタスク，すなわちマイクロタスクを，数百や数千という大量のワーカに安価に行ってもらうなどといった，大規模かつ流動的な雇用が可能となっている．

このようなマイクロタスク型クラウドソーシングを用いてヒューマンコンピュテーションを実現するという研究が近年多く提案されている [3]．これらの研究の多くは，コンピュータビジョン，自然言語処理，データベースにおけるタグ付け処理，機械学習における教師データの生成などにクラウドソーシングを活用するという，主にコンテンツの理解のための研究に分類される．本研究の目的はコンテンツの理解ではなく生成であり，これらの研究とは目的が異なる．

3　提案手法

3.1　想定するパラメタの種類

本研究で対象とするパラメタとは，写真補正における明るさやコントラスト，CG におけるマテリアルの設定，カメラの姿勢，要素の配置，手続き的モデリングのパラメタなど、コンテンツ生成のためにユーザが指定する連続的な値の集合である．特に数百次元や数千次元といった，パラメタ空間の次元が高い場合も想定する．

3.2　マイクロタスク

クラウドソーシングでは不特定多数の人間にタスクを委託するため，タスクそのものは特別なスキルを必要としないできるだけ単純なタスク，すなわちマイクロタスクとすると効果的である [2]．そこで本研究でもマイクロタスクを採用する．

まず，対象とするパラメタ空間からランダムに 2 点選出し，それぞれのパラメタセットに対応する画像を生成しておく．それぞれのマイクロタスクでは，これら 2 枚の画像を同時に見せ，相対的にどちらの方が「好み」の度合いが大きいかを点数付けしてもらう．Gingold ら [2] や Chaudhuri ら [1] の手法と同様に，度合いの相対値を点数付けしてもらうことによって，1 枚の画像に対して度合いの絶対値を点数付けしてもらう場合よりも，より妥当な結果が得られることが期待される．

Copyright is held by the author(s).

* Yuki Koyama, Daisuke Sakamoto, and Takeo Igarashi, 東京大学

[1] http://www.mturk.com/

3.3 パラメタ空間の解析

上述のマイクロタスクによって得られるデータは，パラメタ空間における 2 点における相対的な「好み」の度合いの差の集合である．このようなデータから，パラメタ空間全体の「好み」の分布を推定するために，本手法では適切なエネルギーを定義し，エネルギーの最小化問題として定式化するというアプローチをとる．エネルギーの定義は，連続的なパラメタ空間を扱うという仮定から近傍のパラメタセットの「好み」の度合いをスムージングする項，及び「好み」の度合いの相対的な差を保つような制約を表す項の 2 つを足し合わせたものとする．

4 実験

本研究のアプローチの妥当性を確認するため，写真補正を例に簡単な実験を行った．今回の実験では，明るさ，コントラスト，彩度の 3 つのパラメタからなるパラメタ空間を対象とした．

クラウドソーシングのプラットフォームとしては，CrowdFlower[2]で利用可能な全てのサービス，地域を対象とした．また，クラウドソーシングにおける回答の品質管理については，本実験では duplication と呼ばれる手法 [2] を使用した．

条件： 今回の実験ではパラメタ空間中から 2000 点をランダムにサンプリングし，その中からランダムに 2 点ずつペアにして，合計で 1000 ペアについての比較結果を集めた．図 1 に比較，及び評価を行うために実際に使われた画面の例を示す．

結果： 図 2 (a) に示す写真を入力として実験を行った結果として推定された最適なパラメタセットを元に，図 2 (b) に示す写真が得られた．本実験では合計で 4.00 USD の報酬がクラウドワーカに支払われた．

5 考察

我々は実験結果としてある程度期待通りのものが得られたと考えているが，コンテンツ生成には厳密な正解があるわけではないため，本手法の妥当性の評価方法は本研究の大きな課題であると考えている．また本手法の有用性の評価としては，機械による自動最適化との比較，及び同額の報酬という条件下でクラウドによって生成された結果と専門家によって生成された結果との比較という 2 つの方法を考えている．

手法の改善課題としては，パラメタ空間において最適解が存在すると思われる領域を重点的にサンプリングすることによって，より少ない比較回数で，結果としてより安価に同等の結果を得ることができると考えている．そのために，前のクラウドワーカの回答結果を元に，次のクラウドワーカのタスクを動的に生成するような手法を考案していく予定である．

[2] http://crowdflower.com/



図 1. 実際のタスクの様子．



図 2. (a) 元の写真．(b) 実験で得られた最適なパラメタセットに基づいて生成された写真．

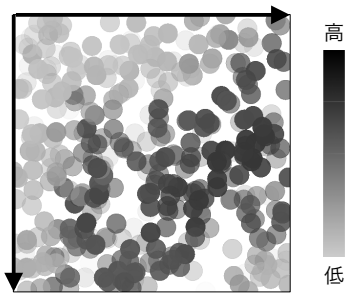


図 3. 推定した「好み」の度合いを可視化した様子．横軸は明るさ，縦軸はコントラストを表している．ここでは二次元画像として可視化するため，彩度に関して適切な断面を選び，その付近を表示している．

参考文献

[1] S. Chaudhuri, E. Kalogerakis, S. Giguere, and T. Funkhouser. Attribit: content creation with semantic attributes. In *Proc. of UIST '13*, pp. 193–202, 2013.

[2] Y. Gingold, A. Shamir, and D. Cohen-Or. Micro perceptual human computation for visual tasks. *ACM Trans. Graph.*, 31(5):119:1–119:12, 2012.

[3] A. J. Quinn and B. B. Bederson. Human computation: a survey and taxonomy of a growing field. In *Proc. of CHI '11*, pp. 1403–1412, 2011.

UnReC：文字入力に基づいた Undo/Redo

大江 龍人　　志築 文太郎　　田中 二郎*

概要．テキストエディタの Undo/Redo では，ボタンやショートカットを用いて履歴を一度ずつ操作する手間があり，Undo/Redo 操作後に文章を上書きした際 Undo/Redo が不可能な場合がある．この問題を解決するために，我々は履歴から任意の時点を選択し Undo/Redo を行うことが可能であるインタフェースの有効性を調査している．現在，この目的のために文字入力に基づいた Undo/Redo インタフェースである UnReC を開発している．UnReC を用いることにより，ユーザは過去に入力した文字列をキーとして履歴内を検索し，その時点まで Undo/Redo を行うことが可能である．本論文では UnReC を用いた操作手法と実装を示す．

1　序論

テキストエディタにおいて Undo/Redo を行う場合，ユーザは図 1a に示す様に線形の履歴から一度ずつ操作する．この Undo/Redo 操作において目的の時点へ Undo/Redo を行うためには，ユーザは複数回の操作を用いて履歴を探索する必要がある．またこの環境では，履歴から任意の時点を選択することが出来ないため，複数回の操作を用いて履歴を探索する手間がある．さらに，ユーザが Undo/Redo 操作後に文章を上書きした際 Undo/Redo が不可能な場合がある．

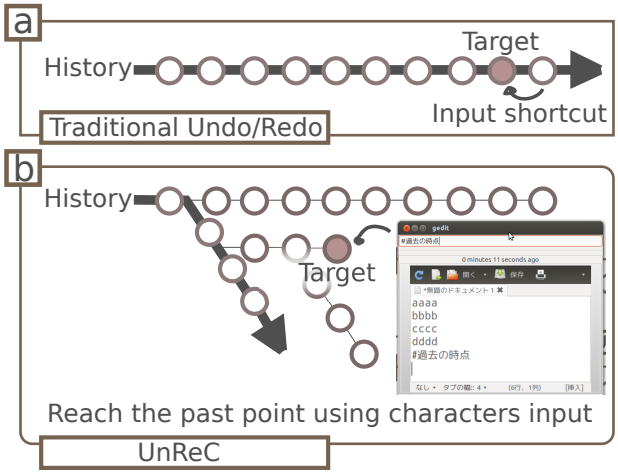


図 1. a）通常の **Undo/Redo**，b）**UnReC** を用いた **Undo/Redo**.

Copyright is held by the author(s).

* Tatsuhito Oe, 筑波大学大学院システム情報工学研究科, Buntarou Shizuki and Jiro Tanaka, 筑波大学システム情報系

[1] http://www.emacswiki.org/emacs/UndoTree

[2] http://www.vim.org/scripts/script.php?script_id=3304

履歴から任意の時点を選択する Undo/Redo を可能にした Selective Undo [1] は，選択した時点のみの Undo/Redo を可能にしたが，一般的な線形の履歴に対応していない．また Regional Undo [2] は，一部の範囲内のみに Undo/Redo の対象を絞ることを可能にしたが，任意の時点に戻ることが出来ない．任意の時点に戻ることが可能である木構造の履歴に対応したものとして，Emacs の Undo Tree[1]や Vim の Gundo[2]が挙げられる．しかし，これらの環境においてユーザは，ある時点へ到達するために木を探索する必要がある．

我々は履歴から任意の時点を選択し Undo/Redo を行うことが可能であるインタフェースの有効性を調査している．現在，この目的のために文字入力に基づいた Undo/Redo インタフェースである UnReC を開発している．図 1b に示す様に UnReC を用いることにより，ユーザは過去に入力した文字列をキーとして履歴内を検索し，その時点まで Undo/Redo を行うことが可能である．なお，我々は UnReC の名前を Undo/Redo by Characters Input（文字入力に基づいた Undo/Redo）の頭文字から命名した．

2　UnReC

UnReC を実際に使用している例を図 2 に示す．図 2 に示す様に，ユーザはテキストエディタと共に UnReC を使用する．UnReC の文字入力フォームに文字列を入力すると，過去の文字入力時点のスクリーンショットが閲覧出来る．そのスクリーンショットをマウスクリックすることにより，その時点まで Undo/Redo することが可能である．

3　UnReC の応用

UnReC の応用として「簡易なバージョン管理」と「過去の文章の再利用」が挙げられる．それぞれ以下に説明する．



図 2. **UnReC** の使用例．この例において，ユーザは「コミット」と文字入力した時点を履歴から検索し，そこまで **Undo** を実行しようとしている．

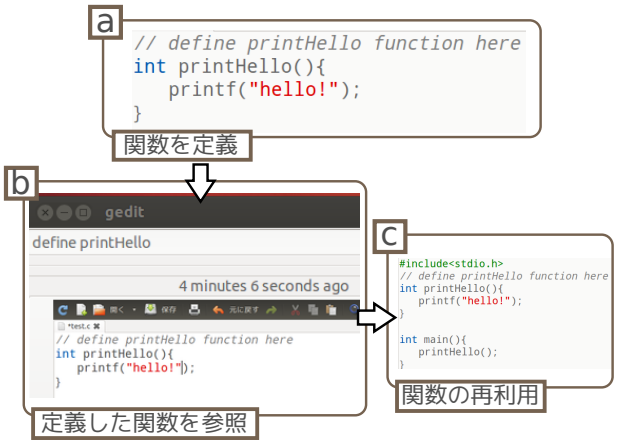


図 3. 過去の文章の再利用例．**a**）コーディング中に関数を定義する，**b**）必要な際に定義した関数を参照する，**c**）過去の関数を現在のコードに再利用する．

UnReC を用いることにより，あらかじめメタな文字列を入力し，その時点まで Undo/Redo を行う簡易なバージョン管理が実現可能である．例えば図 2 において，ユーザはコーディング中に戻りたい時点を定義するために「//ここでコミット 1」や「//コミットしとく 2」等の文字列を入力している．その後 UnReC の文字入力フォームに「コミット」と入力し時点を検索する．最後に，出現したスクリーンショットをマウスクリックすることにより，コミットした時点まで Undo 可能である．

UnReC の別の応用として，図 3 に示す過去の文章の再利用がある．この例においてユーザはコーディング中に関数を定義し（図 3a），その後，関数が不要だと考え削除し，数時間コーディングを続けているものとする．ここで改めて関数が必要になった場合，ユーザは UnReC の文字入力フォームに関数名や定義時のコメントを入力し（図 3b），UnReC 上に表示されるスクリーンショットをクリックしさえすれば，過去の関数の参照及び再利用が可能である（図 3c）．

4 実装

UnReC の実装は保存，検索，Undo/Redo 処理から構成される．今回，我々は Ubuntu 上の Gedit テキストエディタのプラグインとして UnReC を実装した[3]．

保存：エディタにおいて 1 行分の文字列が入力されるごとに，システムはその時点における状態を保存する．保存する内容は，「検索用の 1 行文字列」，「全文章」，「スクリーンショット」から構成される．この保存処理は，通常の Undo/Redo では削除される文章も全て保存するため，任意の時点における Undo/Redo が可能になる．

検索：UnReC の入力フォームに文字列が入力されると，システムはその文字列が検索用の 1 行文字列に含まれているか否かの判定を行う．含まれている場合，保存したスクリーンショットを表示する．この際スクリーンショットの順序は，1 行文字列と入力フォームの類似度順にて表示される．ユーザがスクリーンショットを選択すると，システムは Undo/Redo 処理を実行する．

Undo/Redo 処理：現在の全文章を全て削除し，過去の全文章を貼り付けることにより，Undo/Redo を行う．

[3] svn checkout `https://subversion.assembla.com/svn/oe_public_projects/Gedit_UnReC`

5 結論と今後の発展

本論文では，履歴から任意の時点を選択し Undo/Redo を行うことが可能であるインタフェースである UnReC を示し，さらに UnReC を用いた操作手法と実装を示した．

今後は UnReC を様々なユーザに使用してもらう．実際に使用したユーザの意見を反映させることにより，UnReC をより実用的なシステムとする．

参考文献

[1] T. Berlage. A Selective Undo Mechanism for Graphical User Interfaces Based on Command Objects. *ACM Trans. Comput.-Hum. Interact.*, 1(3):269–294, 1994.

[2] Y. Kawasaki and T. Igarashi. Regional Undo for Spreadsheets. In *Proc. UIST 2004 Demostration Abstract*, pp. 1–2. ACM, 2004.

どこでも「いいね！」ボタン:位置評価情報の提示によるユーザの興味に関する研究

佐々木 佳祐　　寺田 実*

概要. ニュースや新聞から得られる情報には速さと正確さが求められるが, 新製品やイベントなどの情報を公開する側は, ユーザの興味を引きつけておくためにあえて少ない情報を公開するという方法を選ぶことも多い. そこで本稿では,「入手するまでの速度は速いが内容がある程度隠れている情報」はユーザの興味を引くであろういう仮説を立てた. さらに, 仮説の検証のために適していると思われる情報として緯度や経度といった位置情報に着目し, スマートフォン上に現在地点に関する好感/嫌悪感の二値評価を投稿, 集約, 表示するアプリケーションを作成, 運用することで, 仮説の妥当性について考察を行った.

1 はじめに

昨今の情報通信技術と Twitter を始めとするソーシャルネットワーキングサービスの普及により,「今, どこで何が起きているのか」という情報を限りなくリアルタイムに取得することが可能となってきている. ネットワークを通じて友人の行動や催し物の情報を把握し, その日の予定を立てるといったことも珍しくない.

一方, 新製品やイベントの情報などを公開する側の情報提供者は, 情報の一部を隠しつつ小出しに公開する, という手法を用いる場合がある. 代表的な例として「ティザー広告」という手法が挙げられる. 広告は一般的に, 商品の名称や機能, 外観, 価格などのセールスポイントとなる部分を打ち出すものだが, ティザー広告では具体的な情報を敢えて隠し, 消費者が能動的に注意を向けることを目的としている. 大手企業やイベントなど, ブランド力の高いベンダーによるティザー広告は, それ自体が大きな話題となる場合も多い.

以上 2 つの要素を組み合わせると,「リアルタイムでありながら一部が隠れている情報」は, 正確な情報の全貌が明らかになるまでは消費されず残るために情報の寿命が長く, よりユーザの興味を引けるのではないかと考えられる.

そこで本稿では, 情報を制限することによるユーザの興味について以下の 2 つの仮説を立案し, 仮説を基にした「最もユーザ興味を引くことのできる情報の隠し方」について検討するため, 我々の生活において身近な情報である緯度, 経度の情報を利用した実世界におけるリアルタイムな位置評価情報共有システムを提案, 手始めに情報の隠蔽度を最大限に高めたシステムを実装し, 運用実験を行った.

Copyright is held by the author(s).

* Keisuke Sasaki and Minoru Terada, 電気通信大学大学院 情報理工学研究科

仮説 1. 実世界においてユーザは, 自分以外の誰かが「良い」「悪い」といった評価を下したある地点について何らかの興味を持つはずである

仮説 2. ある地点で起きている事象全ての情報と, 仮説 1 における二値の情報の間に, ユーザの興味を最も換気するポイントが存在する

2 提案システム

2.1 概要

本研究における仮説を検討するために,「実世界における位置評価情報共有システム」を提案する (図 1). システムは,GPS 機能を持った携帯端末から, ある緯度および経度に対して, ユーザがその位置に対して抱いた「好感」または「嫌悪感」の二値の情報を収集する. また, 収集された二値の情報は, システムが表示する地図上に散布される (図 2(a), 2(b)). 本システムでは, 仮説 1 における「リアルタイム性」と「制限された情報」の両者を満たしていると考えられる. リアルタイム性においては,

- 地図を開いた際にリアルタイムな評価の情報がすぐに取得できる

情報の制限という面では,

- 「誰が」評価したのかわからない
- 「何故」評価されたのかわからない
- そこに「何があるのか」は実際に行ってみないとわからない

という点が挙げられる. ユーザに提示される情報はあくまで「誰かが好感, または嫌悪感を抱いた位置」のみである. 情報を敢えて二値に設定することで, 実際にその場で何が起こっているのかをユーザに想起させる狙いがある. この「制限された情報」が提示された際にユーザがどのような行動を起こすかを, 実際にシステムを運用して明らかにしていく.

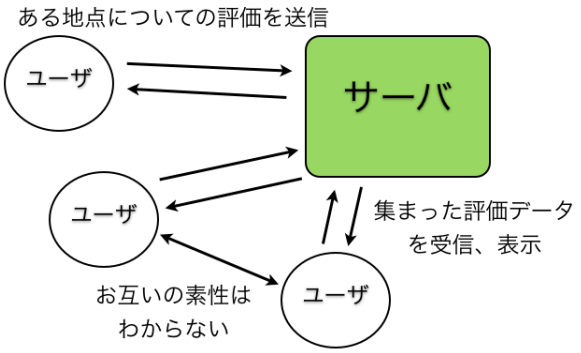


図 1. システムとユーザの振る舞い



(a) 地図及び評価アイコン (b) 評価送信後のフィードバック画面

図 2. 実装したアプリケーションの画面

2.2 スマートフォン用アプリケーション

近年爆発的に普及しているスマートフォンはそのほとんどが GPS 機能, 地図表示機能を搭載しているため, システムはスマートフォン用アプリケーションとして実装するのが望ましいと思われる. そこで提案システムは, スマートフォンに備えている地図表示機能を拡張し, 地図および評価情報の散布表示機能, またユーザの現在地の評価情報を送信する機能を備えたアプリケーションとして設計を行った.

アプリケーションは以下の機能を備えており, 今度も機能追加を行う.

- 評価データの送受信
- 自身の評価データ送信履歴
- 評価データに対するリアクション機能

リアクション機能とは, 散布された評価データに対し「同意」のリアクションを送信することが可能となっている. 自身が行った評価にフィードバックがあることで, 評価データの送信を促進する狙いがある. リアクションの形式についても今後検討していく.

2.3 運用

アプリケーションを実際に公開し, 不特定多数のユーザにアンケート調査を行ったところ, 他人の評価について興味を持ったユーザ数は 4 割に上った. しかしながら実験に参加したユーザ数は 17 名と統計的なデータとするには不十分な値であったため, 実験方法の見直しやアプリケーションの機能の洗練が今後の課題となった.

3 今後の展望

本システムが目指す地点は, 「ユーザがアプリケーションを起動すると常にどこかの位置に評価データが集中しており, 近場であれば足を運び, 距離がある場合はその地点で何が起きているのか能動的に調査を行う」という状況を発生させるところにある. そのためには莫大なユーザ数が必要となるが, twitter や Foursquare など既に普及している SNS のデータを用いるなどして規模を拡大していく.

また, ユーザの興味を最も喚起させられる適度な情報の隠し方を探るために, さまざまな手法を考案, 実装, 実験を行う. 現在, 「実際に足を運ぶ」ことが容易なシチュエーションである学園祭での実験を計画しており, そのために位置精度の向上, 低解像度に加工された写真の添付といった機能追加を行っている.

4 まとめ

本研究では, 実世界におけるある地点について, 敢えて情報を制限してユーザに提示することで興味を引けるのではないか, という仮説に着目し, ある地点について誰かが好感, または嫌悪感を感じているという情報のみを提示する手法を提案し, Android OS 用のアプリケーションとして実装した. 実装したシステムを運用し, 可能性, 有効性, 問題点などを明らかにした. 今後は被験者を増やし, 大人数によるユーザテストを実施する予定である. また, 情報の制限量を変更した際にユーザが抱く興味にどのような相関が現れるか, といった調査も引き続き検証していく.

SmarTrackBall：スマートフォンのアウトカメラを用いたトラックボール

深津 佳智　　志築 文太郎　　田中 二郎*

概要.　スマートフォンのアウトカメラを用いてトラックボールと同様の操作を可能とするシステム SmarTrackBall について示す．SmarTrackBall では，ユーザはスマートフォンの背面で操作を行うため，スマートフォンの画面を塞ぐことなく操作を行うことができる．加えて，ハードウェアを着脱できる，また，片手で操作できるという利点を持つ．本システムのプロトタイプを作成し，本システムを用いた 2 種類のアプリケーションを作成した．

1　はじめに

従来のスマートフォン操作においては，主に，タッチスクリーンに指を触れることによるタッチ操作が行われる．これに対し，スマートフォンにシンプルなハードウェアを追加するのみにより，タッチ操作以外の多様な操作を可能にする研究が行われている．我々の提案システム SmarTrackBall では，スマートフォンのアウトカメラ上にマーカ付のアクリルボールを持つハードウェアを取り付けることにより，トラックボールと同様の操作（以下，トラックボール操作）を可能にする．

本稿では，SmarTrackBall の概要，プロトタイプの設計，アプリケーションについて述べる．

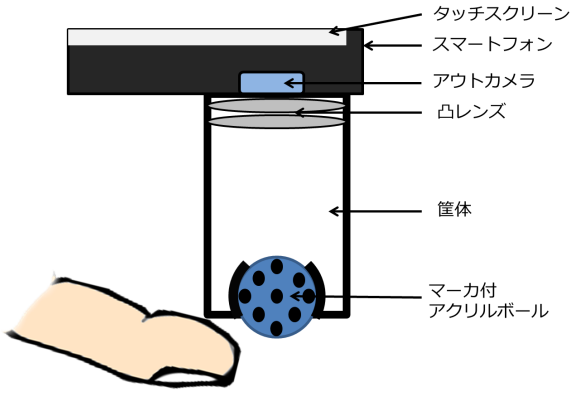


図 1. SmarTrackBall

2　関連研究

スマートフォンにシンプルなハードウェアを追加するのみにより，タッチ操作以外の多様な操作を可能とする研究が行われてきた．Yu ら [4] は，タッチスクリーンの端に導電性ゴムでできた物理ボタン Clip-on Gadgets を取り付けることにより，物理ボタンを押す操作の認識を実現した．渡部らは [5]，マーカを埋め込んだ弾性体をスマートフォンの内蔵カメラ上に取り付け，ユーザが弾性体に力を加えた際のマーカの移動量を光学式力分布測定手法 [2] により測定することにより，ポインティングや押し込み操作を実現した．Spelmezan ら [1] は，圧力センサと近接センサを組み合わせた物理ボタンをスマートフォン側面に取り付けることにより，スマートフォンを片手操作するのための 6 種類の親指ジェスチャの認識を実現した．Yang ら [3] は，スマートフォンのインカメラに全方位カメラを取り付け，周辺視覚を検出することにより，周辺環境の認識，周辺物の認識，ハンドジェスチャ認識，ユーザの動きの認識などを実現した．我々の提案システム SmarTrackBall では，スマートフォンのアウトカメラ上にマーカ付アクリルボールを取り付け，ユーザが指でボールを転がしたことをアウトカメラを用いて検出することにより，トラックボール操作を実現した．

3　SmarTrackBall

SmarTrackBall は，スマートフォンのアウトカメラを用いてトラックボール操作を可能とするシステムである（図 1）．本システムの利点を以下に示す．

- ユーザはスマートフォンの背面で操作を行うため，スマートフォンの画面を塞ぐことなく操作を行うことができる．
- ハードウェアを着脱可能なため，システムを使用しない時にハードウェアを外しておけば，通常の操作の妨げにならない．
- 一方の手が塞がっている場合でも，片手で操作できる．

Copyright is held by the author(s).

* Yoshitomo Fukatsu, 筑波大学大学院 システム情報工学研究科, Buntarou Shizuki and Jiro Tanaka, 筑波大学 システム情報系

4 プロトタイプシステム

プロトタイプシステムは，アウトカメラ上に取り付けるハードウェア，及び，トラックボールの動きを検出するソフトウェアから構成される．

4.1 ハードウェア

本システムのハードウェアは，凸レンズ（直径 21 mm，拡大率 3 倍，2 枚），マーカ付アクリルボール（直径 10 mm），それらを固定する筐体からなる（図 1）．このハードウェアをスマートフォン（Apple iPhone4S）のアウトカメラ上に取り付けた．

4.2 ソフトウェア

本システムのソフトウェアは，アウトカメラから取得した画像から，アクリルボールの動きを検出することにより，ユーザの操作を認識する．具体的には，まず，アウトカメラから取得した RGB 画像（図 2a）をグレースケール画像（図 2b）に変換する．毎フレーム毎に，前のフレームと現在のフレームにおけるグレースケール画像のオプティカルフローを計算することにより，アクリルボールの動きを検出する．オプティカルフローの計算には，オープンソースのコンピュータビジョンライブラリ OpenCV[1] を用いた．

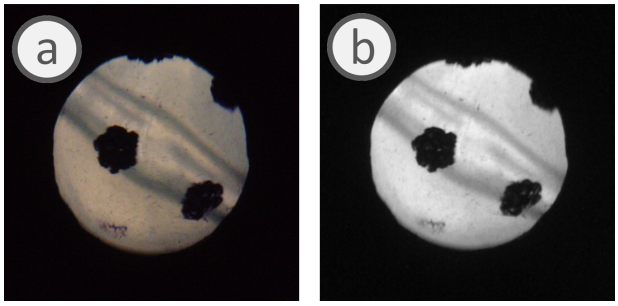


図 2. アウトカメラから取得した画像．a）RGB 画像，b）グレースケール画像

5 アプリケーション

SmarTrackBall を用いた 2 種類のアプリケーションを作成した．

5.1 スマートフォン背面のトラックボールによるスクロール操作

ユーザは，スマートフォンを用いてウェブブラウザなどのスクロール画面を閲覧する際に，スクロール操作を行う．従来のタッチ操作では，ユーザの見ている画面をタッチ操作を行う指で塞いでしまう問題が生じる．提案システムでは，スマートフォン背面のトラックボールを転がしてスクロール操作を行うため，ユーザは画面を塞ぐことなく操作を行うことができる．

5.2 マウス操作

SmarTrackBall をマウスパッド上で操作することにより，従来のマウスと同様の操作を行うことができる．本操作は，従来のマウス操作と比べ，スマートフォンのタッチスクリーンを用いて，マルチタッチジェスチャを行うことができる，また，視覚的フィードバックを与えることができるという利点がある．

6 まとめと今後の課題

スマートフォン背面においてトラックボール操作を可能とするシステム SmarTrackBall について述べた．SmarTrackBall を用いることにより，ユーザは，スマートフォンの画面を塞ぐことなくトラックボール操作を行うことができる．我々は，プロトタイプシステムを作成し，トラックボール操作を用いた 2 種類のアプリケーションを示した．

今後は，ハードウェア及びソフトウェアの改良による認識精度の向上，また，アプリケーションの評価を行う．

参考文献

[1] D. Spelmezan, C. Appert, O. Chapuis, and E. Pietriga. Controlling widgets with one power-up button. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, UIST '13, pp. 71–74, New York, NY, USA, 2013. ACM.

[2] K. Vlack, T. Mizota, N. Kawakami, K. Kamiyama, H. Kajimoto, and S. Tachi. GelForce: a vision-based traction field computer interface. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pp. 1154–1155, New York, NY, USA, 2005. ACM.

[3] X.-D. Yang, K. Hasan, N. Bruce, and P. Irani. Surround-see: enabling peripheral vision on smartphones during active use. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, UIST '13, pp. 291–300, New York, NY, USA, 2013. ACM.

[4] N.-H. Yu, S.-S. Tsai, I.-C. Hsiao, D.-J. Tsai, M.-H. Lee, M. Y. Chen, and Y.-P. Hung. Clip-on gadgets: expanding multi-touch interaction area with unpowered tactile controls. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, UIST '11, pp. 367–372, New York, NY, USA, 2011. ACM.

[5] 渡部 陽一, 佐藤 克成, 牧野 泰才, 前野 隆司. 光学式力測定手法を用いた携帯型タッチパネル端末用入力デバイスの提案. インタラクション 2012, pp. 521–526. 一般社団法人情報処理学会, 2012.

[1] OpenCV: `http://opencv.org/`

人間の表情を活用するライフログ閲覧

高井 啓　志築 文太郎　田中 二郎*

概要．　ライフログの閲覧に人間の表情を活用するシステムを提案する．リアルタイムに生成されるライフログデータは膨大なものとなる．我々はこの問題を解決する為に，ライフログの閲覧に表情を活用するシステムを実装した．提案システムはライフログデータと同期してユーザの表情を記録する．これにより，表情をキーとしたライフログデータの検索を可能とする．我々はケーススタディとして，PC の画面を記録し続けたライフログデータを，提案システムを使って閲覧した．本稿ではその結果を元に，表情を活用したライフログの閲覧の有効性について考察する．

1　背景と目的

人間の行動を記録し続けることによって生成されたライフログデータを利用し，ユーザの行動の振り返りや記憶支援に役立てる研究が存在する．その中でも，コンピュータの操作ログやユーザの視界を撮り続けるログ等，リアルタイムに生成されるライフログデータは膨大なものとなるため，効率的にデータを閲覧する為には，大量のデータの中からユーザが必要としている部分を抽出する必要がある．

ユーザが必要としている部分を抽出する研究の例として，Datchakorn ら [2] による研究がある．Datchakorn らは，動画，GPS，心拍数等，様々なセンサを利用して作られたライフログから，コンテキストに注目してサマリを作成した．これにより，会話等，ユーザにとって重要な意味を持つ部分を抜き出すことが可能になった．また，福本ら [1] は，目元の筋肉の動きから笑顔を認識するメガネ状のデバイスを作成し，その情報を元に，ウェアラブルカメラによるライフログから，ユーザが笑顔の時に見ていた部分を取得し，提示する方法を提案した．

本研究はライフログの閲覧に人間の表情を活用するシステムを提案する．表情は感情を表す最も基本的な要素であり，記憶と感情は結びついている．その為表情が動いた時，即ち感情が動いた時のライフログデータは他のライフログデータに比べ，ユーザにとってより有益な情報を含んでいると考えられる．そこで，提案システムはライフログデータと同期してユーザの表情を記録する．これにより表情をキーとしたライフログデータの検索を可能とする．この検索を用いることにより，ユーザは自分の行動を見返したり，隠れた好みを見つけ出すことが出来ると考える．

Copyright is held by the author(s).

* Hiromu Takai, 筑波大学大学院システム情報工学研究科, Buntarou Shizuki and Jiro Tanaka, 筑波大学システム情報系

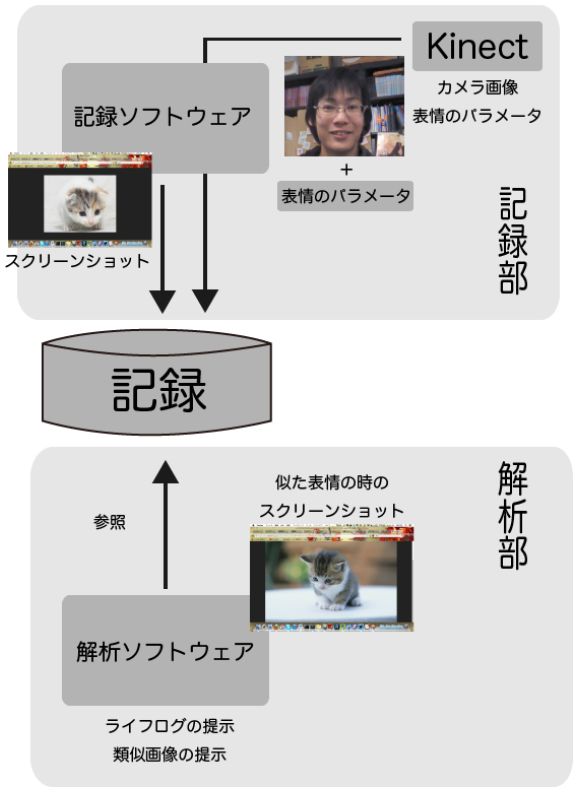


図 1．システム構成

2　システム

システムは記録部と解析部によって構成される．図 1 にシステム構成を示す．

記録部は表情とライフログデータの記録を行う．我々は表情を記録する為に，Kinect とそれに付随するライブラリを利用した．なお今回我々はライフログとして，PC の画面のスクリーンショットを記録し続けるライフログ（以降，画面ログ）を採用した．これはユーザが何を見ていたかが直ぐに解る為，実験やデバッグに適していると考えたからである．

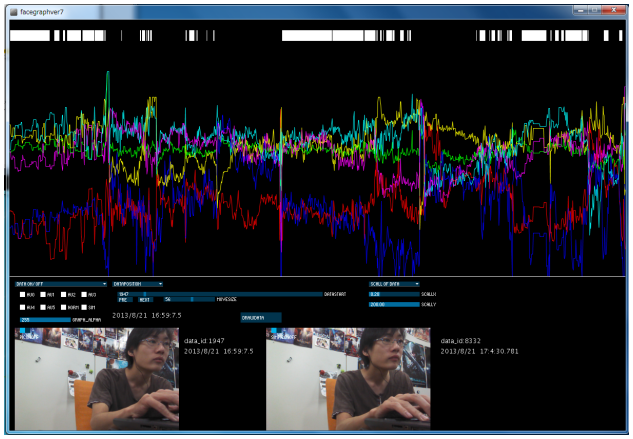


図 2. 解析ソフトウェア

解析部は記録した表情を解析し，ライフログデータを提示する．ユーザはこの解析部を利用してライフログを閲覧する，これにより自分の行動を見返す，または自分の隠れた好みを見つけ出すことが出来る．解析部のうち，ユーザにライフログのデータを提示する解析ソフトウェアのスクリーンショットを，図 2 に示す．表情を数値で表したデータがグラフで表示され（図 2 上）またその下に，スライダで指定された時点での表情（検索キー）（図 2 左下）と，その表情に似た表情（以降，類似画像）が表示される（図 2 中央下），類似画像を提示した理由はユーザが指定した時点によく似ている時点を探すことが出来るようにする為である．これは似ている表情の瞬間は似ている状況にあるという仮説に基づいている．そして，似ている状況を探すことができれば，行動の分析に便利であると考える．そこで表情を表すパラメータ同士の標準ユークリッド距離を類似度として類似画像の提示を行った．

3　ケーススタディ

このシステムを利用し，画面ログの閲覧を行った．保存した画面ログは 1 秒間につき 1 枚のスクリーンショットから構成される．利用したデータは 2012 年 10 月 15 日 20 時から 2013 年 8 月 23 日 16 時までであり，その内容は，解析部のプログラミングをしている時や，PC を使用して動画又はウェブサイトを閲覧している時，また論文のサーベイを行っている時の情報を含んでいる．閲覧した動画には，映画の DVD の他，ニコニコ動画で配信されていたコメント付き動画も含む．

類似画像を用いてライフログデータを検索できることの有効性を確かめるために，ケーススタディとして，様々な閲覧作業を行った．

結果

ユーザは解析ソフトウェアを利用して画面ログの中から手作業で表情が笑顔になっている時点を一箇所選択した所，解析ソフトウェアによってその顔と似ていると判断された時点の表情が表示された．そこで，ユーザはその時点を選択し，さらにそれと似ている表情となっている時点を探してみた．結果として，一度笑った表情を見つけた後は，類似画像を利用して似た表情をしている時点を簡単に抽出することができた．

また，画面ログ中，映画を見ている間の画面ログにおいて同様の作業を行った所，画面ログ自体は静止画のみにも関わらず，その映画のどのシーンで笑ったか，さらに映画のそのシーンの内容まで思い出すことが出来た．

一方，プログラミングを行っている間の画面ログにおいては，映画を見ている時のような内容の抽出を行うことは出来なかった．この原因を調べるためにプログラミングを行っている間の表情を観察した所，顔は笑顔を浮かべていなかった．これが抽出に失敗した原因だと考えられる．また，映画のスクリーンショットとは異なり，プログラミングを行っている時のスクリーンショットを見ても，この表情の時にどんな作業を行っていたかということは，映画とは異なり思い出すことが難しかった．

4　まとめと今後の課題

ある表情に似た表情をしていた時点をライフログデータから検索することが可能なシステムを作成した．ケーススタディを行った所，映画（動画）鑑賞時に於いて，ユーザによって印象的なシーンを抽出することができた．今後は提案システムのライフログ閲覧に於ける有用性を調べて行きたい．

参考文献

[1] K. Fukumoto, T. Terada, and M. Tsukamoto. A smile/laughter recognition mechanism for smile-based life logging. In *Proceedings of the 4th Augmented Human International Conference*, AH '13, pp. 213–220, New York, NY, USA, 2013. ACM.

[2] D. Tancharoen, T. Yamasaki, and K. Aizawa. Practical experience recording and indexing of Life Log video. In *Proceedings of the 2nd ACM workshop on Continuous archival and retrieval of personal experiences*, CARPE '05, pp. 61–66, New York, NY, USA, 2005. ACM.

音楽の構造情報を反映したミュージックビデオを簡単に制作・共有できるWebサービス

中嶋 誠　　五十嵐 健夫 *

概要．　本稿では，音楽の構造情報を反映したミュージックビデオを簡単に制作できるWebシステムを提案する．音楽の区間・小節・拍に同期したエフェクトや映像切り替え効果には，手動で煩雑な作業が必要だが，先行研究のsongleシステムと連携することで簡単に実現可能となった．またWebシステムの特性を生かすことで，素材映像の収集・管理と，制作物の共有が容易である．加えて，音楽構造に対する映像の連動方法を抽象化して定義することにより，自分の好きな音楽に好きな映像セットを組み合わせた視聴が可能な，新しい種類のコンテンツである．

1　はじめに

近年コンテンツのリッチ化が進み，多くのユーザが，音楽に加えて映像の付いたミュージックビデオを楽しむようになっている．また作曲人口の増加や，動画共有サイトなど発表の場の拡大も，音楽に付随する映像の重要性を高める要因になっている．しかし，このような需要の増加に対し，供給が追いついていないという問題がある．なぜなら通常ミュージックビデオは，一つの楽曲に対して一つの映像が職人的な手作業で制作されるため，労力とコストの問題で大量生産が難しいからである．

そこで我々は新たに，ユーザが簡単に音楽と連動する映像を制作して，視聴・共有できるWebシステムを提案する．提案システムには以下の利点がある．

- songleシステム [1] を利用する事で，音楽の構造情報を反映した映像をWeb上で簡単に制作できる
- システムをWeb上に構築する事で，ユーザは映像素材の管理や制作物の共有にかかる手間を大幅に減らす事ができる
- 音楽と映像の一対一対応ではなく，音楽構造に対する映像の連動方法を抽象化して定義するので，自分の好きな音楽に好きな映像セットを組み合わせて視聴できる

2　提案システムが目指す音楽映像

提案システムが目指すのは，音楽の構造情報（区間・小節・拍）と連動した映像である．これらの情報は，ユーザに音楽と映像の高い連動性を感じさせる重要な特徴量である．岩宮 [2] は，各種音楽特徴量と映像要素を連動させて提示し，マッチしている

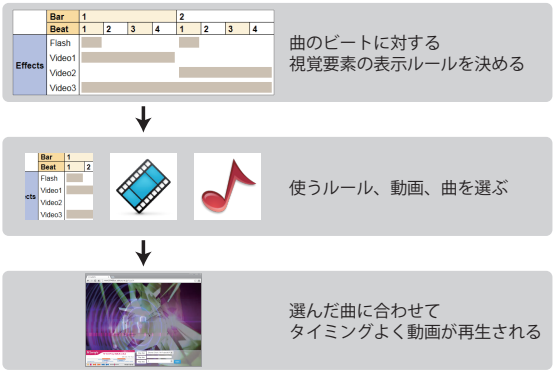


図 1. 提案システムのコンセプト図

と感じる度合いを被験者に尋ねる心理学実験を行った．その結果，音楽の拍情報と，映像の切れ目が同期しているとき，ユーザは高い調和度を感じることが報告されている．また既存のミュージックビデオでも，小節の切れ目や，拍に合わせて映像を切り替える演出がしばしば見られる．既存システムとして，音楽の拍情報と連動した図形アニメーションを表示するsongle外部埋め込みプレイヤーがある [1]。提案システムは，この埋め込みプレイヤーのエフェクトを動画に拡張し，かつユーザがWeb上で動的にエフェクトを定義できるようにしたものである．

音楽映像表現の一つに，音響信号をリアルタイムで変換し映像として表示するビジュアライザがある．提案システムはそれと異なり，あらかじめ用意した動画素材を組み合わせて表示する．これはビジュアライザほど音響信号に直結した表現はできない代わりに，レンダリングに時間のかかる3Dモーショングラフィクス素材や，あらかじめ加工した実写素材など，多様な映像素材を利用可能である．

Copyright is held by the author(s).

* 東京大学

[1] http://staff.aist.go.jp/m.goto/Songle/songle_embedded_player_demo.html

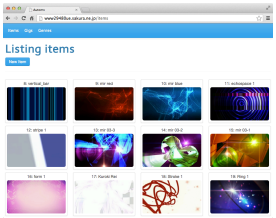


図 2. 動画素材一覧画面

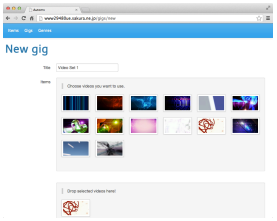


図 3. 動画セット作成画面

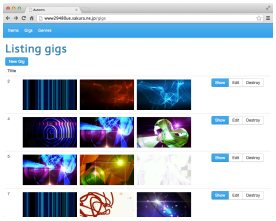


図 4. 動画セット一覧画面

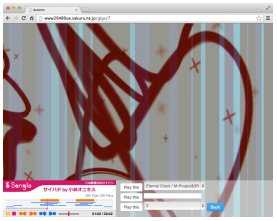


図 5. プレイヤー画面

3 システムの詳細

3.1 ユーザインタラクション

まずユーザは，提案システムの Web インタフェースにより，動画素材をアップロードする．アップロードされた動画素材は，動画素材一覧画面 (図 2) に表示される．将来的には，他のユーザがアップロードした動画を利用する仕組みも考えられる．次にユーザは，アップロードされた動画の中から好きなものを選び，動画セットを作る (図 3)．作成された動画セットは，動画セット一覧画面に表示される (図 4)．動画セット一覧画面で，各動画セットの横にある show ボタンを押すと，その動画セットを再生するプレイヤー画面に遷移する (図 5)．プレイヤー画面には songle 埋め込みプレイヤーが表示され，再生ボタンを押すと音楽が始まり，同時に映像が連動して再生される．またデフォルト曲以外にも，songle の曲 URL を右側フォームに入力し，Play this ボタンを押す事で，songle システムで解析済みのものであれば任意の音楽に，映像セットを連動して再生できる．

ユーザが作成した動画セットは，それぞれが固有の URL を持つので，その URL を共有する事で，他のユーザも視聴できる．またアドレスの末尾に songle の曲 URL を追加して，再生する音楽を指定できる．つまり使用する動画セットと音楽の組み合わせを，一つの URL として表現する事ができる．この URL を他のユーザと共有し，この曲にはこの映像がマッチすると教え合う使い方も可能である．

3.2 実装

提案システムの実装には，Ruby on Rails フレームワークを使用した．動画ファイルのミックス再生には，HTML5 Canvas を利用している．Javascript で songle システムの埋め込みプレイヤーから，音楽構造情報と，音楽の再生状態を取得し，描画する動画素材を制御している．

プレイヤー画面での映像表示について，具体的には，小節ごとに表示する映像素材の切り替え，不透明度を利用した重ね合わせ，リピート区間で拍に連動したズーム効果，サビ区間で小節頭に連動したフラッシュ効果を適用している．

4 まとめと今後の課題

本稿では，音楽の構造情報に連動したミュージックビデオを制作する Web システムを提案した．songle システムを利用する事で，これまで手動で煩雑な作業が必要だった，音楽の区間・小節・拍に同期した映像切り替え表現やエフェクトが簡単に実現できる．また Web システムの特性を生かし，素材映像をユーザが持ち寄る事で，素材集めと管理の手間が減った．また作った物を共有する手間も少なくなった．加えて，音楽の構造情報を利用することで，調和度を保ったまま，任意の音楽と映像の組み合わせを試す事ができる，自由度が高いシステムになった．

今後の課題は，まず映像の表示ルールをユーザが編集できるインタフェースの実装である．現在は 3 つの映像素材からなる映像セットを，ハードコーディングされた表示ルールによって，音楽の構造特徴に連動させて表示している．音楽特徴とエフェクトのマッピングを，ユーザが Web 上で自由に編集しプログラムできるインタフェースがあれば，もっと自由度が高まり，かなりクリエイティブな表現が実現可能と考えられる．

また別の点として，提案システムで制作したミュージックビデオの構成を，AfterEffects など既存の動画編集ソフト向けにエクスポートする機能が望まれる．ミュージックビデオのクリエイターに話を伺ったところ，音楽構造の把握と映像への利用は，今のところほぼ手作業であり，非常に大変とのことであった．そのため支援効果の大きい分野だと考えられる．

提案システムは Web サービスなので，比較的容易にユーザに試用してもらう事ができる．そのために早い段階での一般公開が望ましい．その過程で，ユーザがどのような音楽と映像の組み合わせを好むか，嗜好のデータが収集できるので，データに基づいた学習システムや，サジェストシステムも今後の発展が予想される．

参考文献

[1] 後藤 真孝, 吉井 和佳, 藤原 弘将, M. Matthias, 中野 倫靖. Songle:音楽音響信号理解技術とユーザによる誤り訂正に基づく能動的音楽鑑賞サービス. 情報処理学会論文誌, 54(4):1363–1372, 2013.

[2] 岩宮 眞一郎. 音楽と映像のマルチモーダル・コミュニケーション 改訂版. 九州大学出版会, 2011.

実世界オブジェクトへの機能拡張によるインタフェースの提案

松嶌 信貴　　赤池 英夫　　角田 博保*

概要. 近年登場したメガネ型デバイスにおける新たなインタフェースを提案する．従来のメガネ型デバイスにおけるインタフェースにはそれぞれ問題点が挙げられる．音声やジェスチャを使った手法にはコマンドの種類が制限される問題がある．複雑な操作を可能とするポインティング手法については空中や自分の身体に表示・操作する手法があるが，前者では触覚フィードバックが得られず，後者では身体の一部を専有される問題がある．いずれの手法もユーザの周辺の環境を活用してはいない．そこで本研究では周辺にある実世界のオブジェクトを活用し，各オブジェクトの特徴やアフォーダンスに応じた機能を拡張するインタフェースを提案する．カメラを用いてユーザ視界からオブジェクトを認識するとともに，ユーザインタフェースを重畳表示し，直接触ることによる操作を実現する．物体を動かす，傾ける，回転させる，表面へのタッチ等さまざまな操作を割り当てることができ，広く応用が考えられる．未来ビジョンとしては，現実世界と仮想世界をシームレスに利用する方法として，実世界オブジェクトの機能を拡張する手法を示した．

1　はじめに

本研究では，ユーザの周辺にある実世界オブジェクトを活用することによる，メガネ型デバイスにおける新たなインタラクションを提案する．

ウェアラブルなコンピュータであるメガネ型デバイスが登場し，多様なアプリケーションが開発されている．メガネ型デバイスへの代表的な操作手法には，空中での手・指による操作があるが，触覚フィードバックを得られない．また視線による操作では Midas Touch Problem が知られている．いずれの手法についても，ユーザの周辺の環境を活用してはいない．

そこで本研究では，ユーザの周辺にある実世界のオブジェクトを活用し，直接触って操作できる機能拡張インタフェースを提案する．



図 1. タイマー: カップ麺の拡張

2　提案手法

ユーザの周辺にある実世界オブジェクトにその特徴やアフォーダンスに応じたインタフェースを拡張し，直接触っての操作を実現する．カップ麺のような円筒型オブジェクトを活用する例を考えると，垂直軸で回転させることにより，タイマーの時間を設定する機能が拡張（図 1）できる．物体への操作は他にも，動かす，傾ける，表面へのタッチ等の操作を割り当てることができ，広く応用が考えられる．

3　実装システム

ユーザ視界に情報を表示するためにヘッドマウントディスプレイを使用する．周辺にある実世界オブジェクトや手・指を認識するために，ヘッドマウントディスプレイに併せて頭部に RGB・深度カメラを装着する．

現在の試作では，オブジェクトに貼付した AR マーカから ARToolKit[1] を用いて実世界オブジェクトを認識し，深度ビットマップから OpenNI[2] の手・指認識ライブラリを用いて手・指を認識している．RGB・深度カメラには PrimeSense 社の CARMINE 1.09[3] を使用している．

認識したオブジェクトには，その特徴に応じたユーザインタフェースをヘッドマウントディスプレイにより重畳表示する．オブジェクトの形状に沿って表示することにより，実世界のオブジェクトに機能が拡張されたように感じられる表示を目指す．ヘッド

Copyright is held by the author(s).

* 電気通信大学大学院 情報理工学研究科 情報・通信工学専攻

[1] http://www.hitl.washington.edu/artoolkit/
[2] http://www.openni.org/
[3] http://www.primesense.com/

マウントディスプレイには，Sony 社のヘッドマウントディスプレイ HMZ-T2[4]を使用している．また，画像処理や提示するユーザインタフェースの生成のためにパーソナルなコンピュータを使用している．

システムを軽量化ワイヤレス化するため，現在スマートフォンによる実装にとりかかっている．スマートフォン画面はレンズを通して見ることにより，ヘッドマウントディスプレイのような表示器として使用できる．スマートフォン内蔵のカメラ画像から AR ライブラリ vuforia[5]を用いてオブジェクトを認識している．

4 関連研究

周辺の環境を活用したインタフェースには，C. Harrison らの OmniTouch[1] がある．ウェアラブルな深度カメラと小型プロジェクタにより，ユーザ前方の平面に対してインタフェースを投影し，操作できる．しかし，平面へのタッチ操作のみを対象としており，物体に対するタンジブルな操作を実現する本研究とは異なる．

周辺にあるオブジェクトに機能を拡張し，タンジブルな操作を実現した研究として，J. Huber らの LightBeam[2] がある．深度カメラとプロジェクタにより，紙やマグカップにインタフェースを提示し，直接触れることによるタンジブルな操作を提案している．しかし，手に把持またはテーブル上に配置した状態での使用方法の調査が主であり，ウェアラブルなメガネ型デバイスを使用する本研究とは異なる．

周辺オブジェクトを活用した現実拡張インタフェース [3] は著者らの試作を発表したものである．

5 おわりに

本稿では，メガネ型デバイスにおけるインタラクション手法として，周辺にある実世界オブジェクトに機能を拡張する手法を提案した．当日は実装したアプリケーションのデモを行う．

参考文献

[1] C. Harrison, H. Benko, and A.D. Wilson. OmniTouch: Wearable Multitouch Interaction Everywhere. *UIST 2011*, pp.441-450, 2011.

[2] J. Huber, J. Steimle, C. Liao, Q. Liu, and M.Muhlhauser. LightBeam: Interacting with Augmented Real-World Objects in Pico Projections. *MUM 2012*, Article No. 16, 2012.

[3] 松嶌, 赤池, 角田. 周辺オブジェクトを活用した現実拡張インタフェースの提案. ヒューマンインタフェースシンポジウム 2013, pp.545-548, 2013.

[4] http://www.sony.jp/hmd/products/HMZ-T2/
[5] https://developer.vuforia.com/

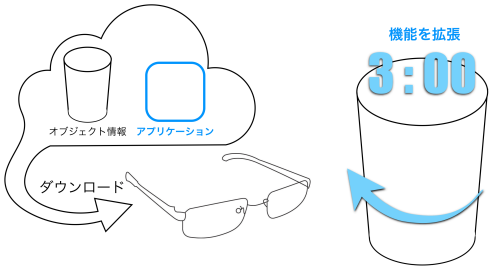
未来ビジョン

本研究が目指すのは現実世界と仮想世界を区別なく便利に利用できる世界である．本研究では，メガネ型デバイスを用いて実世界オブジェクトに情報を付加するだけではなく，オブジェクトの機能を拡張する．

カップ麺のタイマーは一例に過ぎず，日常的に使っているさまざまなオブジェクトに機能を追加することが考えられる．実世界オブジェクト向け拡張機能をアプリケーションとして，メガネ型デバイスからダウンロードして使える仕組みをつくれば，製品を製造している企業が機能拡張アプリケーションを提供することもできるし，第三者が開発し公開することも考えられる．

実世界オブジェクトを認識し，対応するアプリケーションを探すのが困難となりそうであるが，例えば製品バーコードの情報を使用することでオブジェクトを特定でき，アプリケーションとともに特徴量情報等のオブジェクト情報をダウンロードすることで解決できる．ユーザ自身が周囲のオブジェクトに好みの機能を設定してもよいだろう．

実世界オブジェクトを認識し，アプリケーションのダウンロードが自動で行われるようになると，現実世界と仮想世界とをシームレスに利用できるようになるだろう．



手本データを自動抽出する運動指導システム

黒田 修平　　入江 英嗣　　吉永 努 *

概要. 人が新しく運動を学び始める場合，一般的にはその運動の熟練者から指南を受ける．自分自身では気づきにくい運動の誤りを，熟練者に正してもらうためである．しかし，熟練者を要することは，機会やコストの面で習得の敷居を増加させてしまう．熟練者に代わって，コンピュータがユーザの運動の誤りを指南できれば，多くのユーザの運動習得を支援することができる．そこで我々は，人間の関節座標を取得してユーザの運動の流れを検出し，運動の正しさを自動的に判定して指導するシステムを提案している．このシステムは動きをいくつかのポーズの連続としてとらえ，一つ一つのポーズについて，ユーザが正しい体勢をとっているか判断する．この判断について，我々の提案している相対チェックポイントを用いることにより，体格によらずに細やかな運動判定を行うことができる．しかし相対チェックポイントは手動で作成せねばならず，妥当なチェックポイント作成には運動とシステム双方への精通を必要とする．また，新しい運動に対するチェックポイント記述に多大な時間がかかってしまう．

そこで本論文では，練習したい運動の関節座標軌跡データから，チェックポイントを自動生成する手法を提案する．フレーム間類似性に着目した要約技術を利用して特徴的なキーフレームを抽出し，そのキーフレーム中の関節座標データから相対チェックポイントを生成する．これにより，様々な運動の指導がたった一回の教師のお手本によりできるようになる．また教師データの配信により，多くの人にとって様々な運動の練習が身近なものとなる．

1　はじめに

人が新しく運動，体操，ダンスなどを習得しようとする場合，独力での練習を難しくする理由に，自身が上手くできているかの判断が難しいことが挙げられる．このため，インストラクターなど，その運動に精通した者の協力を得て，正しいフォームを指導してもらうことが効果的とされている．しかし，他者に協力を求める際には，外部要因の制約により自由な時間に気軽に学ぶことができず，結果として運動を学ぶ機会が妨げられる．

我々は独力での運動の上達を促すべく，ユーザを自動で指導するシステムを開発している．参考となる技術としては Chaves らの運動認識技術 [1] と同様に，対象となる運動をコマ割りし，特徴的なポーズの連続として表現する．特徴的なポーズを，一定の時間内に同じ順番でとることができれば運動に成功したと判定する．我々は，関節間の相対的な位置関係を相対チェックポイントとすることで，様々な体格のユーザの運動を識別できるようにし，ユーザを指導するシステムを提案した [2]．もしユーザが正しい運動を行えなければ，該当箇所を反復練習させ，運動の上達を促す．

本論文では，提案システムを様々な運動に対応できるように以下の課題に取り組む．まず，相対チェックポイントの妥当性である．ユーザが効果的な習得を行うために効果的なチェックポイントの抽出は熟練者のノウハウに依存している．次にチェックポイント定義の時間的コストである．チェックポイントはそれぞれ手動で定義しており，新しい運動への対応に時間がかかってしまっている．

これらの課題に対し，特徴フレーム抽出技術を用いることにより，上記二つの点を改善し，ユーザが望む運動の練習を効果的に行うシステムを実現する．

2　提案手法

運動指導システム [2] をより正確に，かつ拡張性を高めるため，我々はチェックポイント自動生成手法を提案する．まずユーザは，お手本となる運動の関節座標の軌跡データを得る．これはデータとして配布されたものを取得するほか，フレーム毎の関節座標を取得できるデバイスの前で，運動を習得済みの人間に実際に運動を行ってもらって得る．次にそのデータから，類似度の低いいくつかのフレーム（キーフレーム）を抽出し，そこに記録された関節座標の相対的な位置関係をチェックポイントとして生成する（図 1）．他のフレームと類似しないフレームは，運動の中でも特徴的な重要ポーズのフレームと考えられ，妥当なチェックポイント抽出の自動化が可能となる．

Copyright is held by the author(s).

* Shuhei Kuroda, Hidetsugu Irie, Tsutomu Yoshinaga, 電気通信大学大学院情報システム学研究科情報ネットワークシステム学専攻

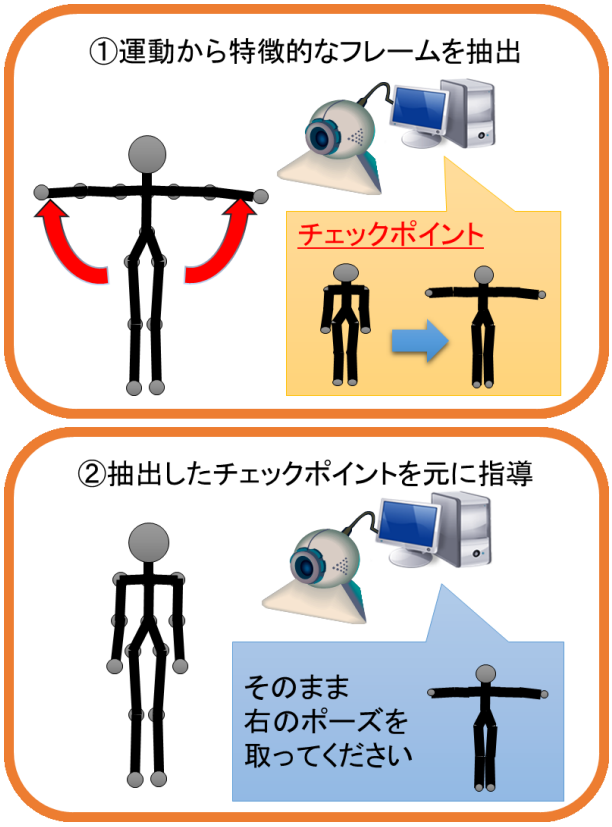


図 1. 提案手法概要

3 アルゴリズム

キーフレームの抽出にあたり，Huang らが提案した，3D ビデオのフレーム間の類似関係を元にキーフレームの抽出を行い，自動的に内容を要約するシステム [3] と同様のアルゴリズムを用いる．まず，何枚のキーフレームを抽出するのかのレートを定義する．また，3D ビデオのモデルを構成するボクセルをひとつひとつを，モデルの重心を中心とした球座標系 (r, ϕ, θ) に変換する．次に，対象となる N 枚のフレーム中にある，任意のフレーム i,j 間の類似度を，ボクセル間のユークリッド距離の合計にて算出する．この類似度とは，フレーム j のモデルを 360 ° 回して，フレーム i のモデルとユークリッド距離的にもっとも近づいたときの差異の和である．次いでこの類似度を元に，フレーム i から j までの間全てのフレームと，フレーム i からフレーム i-j までの全てのフレームとの間に生まれる歪みコストを求める．この歪みコストを元に，3D ビデオの要約に最適となるキーフレームを算出する．

Huang らと異なり，我々の手法では関節同士の距離に着目するため，フレーム i,j 間の類似度を算出するのにモデルを 360 ° 回転させたりせず，それぞれの関節座標同士のユークリッド距離を用いる．また求められたキーフレームに記録されている関節座標間の相対的な位置関係を相対チェックポイントとして生成し，[2] に組み込む．

4 システム構成

本システムでは，ユーザの関節座標を取得するために Kinect を用いる．Kinect はユーザの全身を RGB カメラと赤外線センサーで捉え，20 個の関節座標を取得することが可能である．Kinect を計算機に接続しシステムを起動すると，ユーザへの指導を表示するウィンドウ，RGB カメラで撮影したユーザの関節上に赤いマーカを描画するウィンドウ，そして同様に RGB カメラの撮影の様子の上にボタンが表示されたウィンドウがポップアップされる．このボタンから運動指導か，手本データ録画を選んで，Kinect に映るよう運動する．

5 未来ビジョン

チェックポイント自動生成手法を用いた運動指導システムが確立されれば，人々は熟練者の支援に頼ることなく，様々な運動を上達させることができるようになる．これにより，人々にとって運動はより身近なものとなり，スポーツやダンスなどが盛んに行われるようになる．これを通じ，我々は人々の競技スキルの向上のみならず，人々の健康度の底上げを図っていく．

参考文献

[1] Thiago Chaves, Lucas Figueiredo, Alana Da Gama, Cristiano de Araujo, Veronica Teichrieb, ”Human Body Motion and Gestures Recognition Based on Checkpoints ”, Virtual and Augmented Reality 2012 14th Symposium, 28-31 May 2012

[2] 黒田修平，放地宏佳，吉見真聡，吉永努，入江英嗣，”相対座標を用いた運動指導システム”，DICOMO2013 シンポジウム (2013).

[3] Peng Huang, Adrian Hilton, Jonathan Starck,”Automatic 3D Video Summarization: Key Frame Extraction from Self-Similarity”, 3DPVT’08 - the Fourth International Symposium on 3D Data Processing, Visualization and Transmission(2008).

EnchantMoon を利用した漢字・部首認識を活用した描画補助ツールの提案

若林 花見 * 井川 裕貴 武藤 寛 瀬川 典久 杉野 栄二 澤本 潤†

概要. 近年タブレット端末が普及してきているとともに、タブレット端末用のアプリケーションを用いて絵を書く人が増えてきた。従来のペイントアプリケーションでは、画面上のキャンバスの上部や側部にペンの種類や色、太さを変更するためのツールバーが表示されている。しかし、これらの機能性が操作を複雑にし、自由な発想を妨げていた。そこで本研究では、シンプルなインターフェースを持つ enchantMOON を用い、enchantMoon に部首認識機能を追加し、部首と手書き描画機能をつなげることで、描画補助を行う仕組みを提案する。

1 はじめに

本格的なイラスト描画を行うデジタルデバイスとしてペンタブレットや液晶タブレットが挙げられるが、近年タブレット端末が普及してきているとともに、タブレット端末用のアプリケーションを用いて絵を書く人が増えてきた。タブレット端末で絵が描ければ、改めて機材を買い揃える必要がなく、趣味や遊び目的で絵を描く人にも容易に環境が整う。

従来のペイントアプリケーションでは、インターフェースとしてウィンドウの中央に描画用のキャンバスがあり、上部や側部にツールバーというレイアウト構成がほとんどである。それらのツールを使って描いたものに対してエフェクトをかけるが、その操作は直感的なものではなく、使い方を覚える必要がある。

これでは単純に絵を書くことが目的であるのにも関わらず、手間が増えてしまった。そして、この手間がユーザにストレスを感じさせていた。デッサンやアイディアスケッチの場合このようなストレスを感じることは発想の妨げになってしまう。

そこで、我々の研究グループでは、油絵などで絵を描くときに利用するパレット (図 1) を参考にシステムを実装する。パレットは、絵を書く人がパレット上に絵の具を置いてパレット上で色を混ぜて、欲しい色を作成しその色を塗っていく。

本研究で目指すのは、手書きタブレットに適したパレットみたいな機能を実装することである。本研究では、部首認識を行う部品を画面の自由なところに設置し、利用者が漢字を書くことで、機能を追加できるシステムを提案する。

本研究で用いる enchantMOON[1] とは、ユビキタスエンターテイメント社が開発したタブレット端末である。ほとんどの操作をデジタイザーペンによる手書きと指のタッチ（図 2）のみで行うシンプルなインターフェースである。筆跡ストロークのデータ構造は、X,Y の座標と筆圧の 3 つで構成されていて JSON 形式で保存される。ストロークはオブジェクトとして扱うことができ、さらに、JavaScript で記述したプログラムを機能として追加することができる。本研究の実装にあたってこのインターフェースとストロークに対する動作への機能追加の容易性から、enchantMOON が最適であると考えた。



図 1. パレット



図 2. enchantMOON における指での指示例

本稿では、EnchantMoon を用い、部首認識を使いその部首の組み合わせで、機能を追加する仕組みについて提案を行う。

2 関連研究

手書き認識の研究 [2, 3] は、様々行われており、実用化もされている。そのため、本研究ではそれら

Copyright is held by the author(s).

* カンガルースタジオ
† 岩手県立大学ソフトウェア情報学部

の知見を利用して実装を行っている。

また、手書きをユーザインタフェースとして利用している一つの例として、[4] があげられる。利用者が、手書きでジェスチャーを書くことで、機能を呼び出すことを実現している。

3 プロトタイプシステム概要

利用者は、本プロトタイプシステムを実装したプログラムを起動し、画面上に表示する（図 3）。



図 3. プロトタイプシステムの起動

図 3 で示すように、シンプルな四角が画面上に表示される。利用者は、この四角を自由な場所に移動し、配置することが可能である。

本システムは、次にあげる動作を行（図 4）。

1. 利用者が、四角の中に、漢字を記述する。その記述した物は、部品として画面上に置かれる。
2. 利用者は、必要なときに、漢字の表面を指で触れる。
3. 漢字の認識を部首単位で行い、機能が呼び出される。

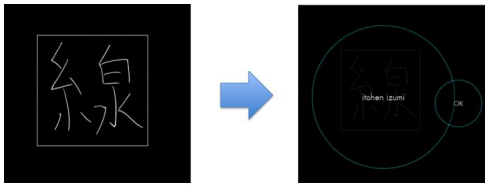


図 4. 部首認識

本システムの認識は、DP マッチングを用い実装している。DP マッチングを用いたのは、実装が容易であるのと、すべての漢字を認識させるわけではなく、限られた部首を認識するためだけに利用するので、必要十分に動作するからである。

図 5 は、システムの動作例である。利用者が、四角の中に「消」の文字を記述しておきその文字を触れることで、今まで書いた手書きストロークが消去する。この文字は、部品として扱われるので、利用者はこの「消」を触れるだけで、ストロークの消去が可能になる。

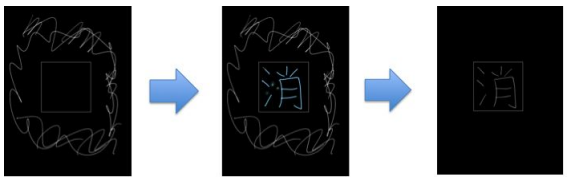


図 5. システム動作例

そして、この部首の組み合わせで、さまざまな機能の割り当てを行う。例えば、「へん」が機能で「つくり」がパラメータを表すなどである。この漢字は、当然の事ながら、創作漢字も含まれる。今後は、認識される部首と、機能の関係を決める文法を定めて実験を行う。

4 おわりに

本研究では、enchantMoon を用い、部首認識を行う部品を、画面上に設置し、そこに漢字を記述することで、描画補助機能を自由に配置し、起動できることを示した。今後は、さまざまな機能を実装し、部首と機能の関係の文法を決定し、さまざまな実験を行う。

参考文献

[1] enchantMOON ; The Hypertext Authoring Tablet http://enchantmoon.com/

[2] 伊藤 等, 中川 正樹:Hidden Markov Model に基づくオンライン手書き文字認識 , 電子情報通信学会技術研究報告. MVE, マルチメディア・仮想環境基礎 97(207), 95-100, 1997-07-25

[3] 福島 貴弘, 中川 正樹:確率モデルに基づくオンライン枠なし手書き文字列認識電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解 98(489), 25-30, 1998-12-17

[4] 坂東 宏和, 福島 貴弘, 加藤 直樹, 中川 正樹:枠なし手書き文字列認識における誤認識訂正インタフェース情報処理学会論文誌 43(6), 1996-2005, 2002-06-15

講義中の板書を支援する人物半透明化カメラの実装

山根 恵和　　岡部 正幸　　梅村 恭司 *

概要. 通常，講義などの場において，黒板やホワイトボードを用いて説明を行うという形式がとられている．ここでの問題は，講義者が黒板やホワイトボードに文字や図を書いているとき，講義者と文字が重なる際に，受講者からは，黒板の文字が見れなくなることである．そこで本システムでは，受講者が，講義者および講義者により遮られた文字を同時に見ることができる「半透明人間カメラ」を提案する．文字や図が書かれた黒板と半透明化を施した講義者の映像を別スクリーンへ出力することで，受講者はそのスクリーンを見るだけで講義者により遮断されて見ることが不可能であった文字や，講義者の説明箇所を正確に把握することが可能となる．

1　はじめに

講義や会議などの場では，黒板やホワイトボードを使用した説明が主流である．最近では電子黒板などの電子ボードも注目を集めているが，いずれも講義者が文字や図を書く，または書いたものを説明するときに，講義者の手や身体により文字や図が遮蔽されてしまい受講者から見えなくなる問題がある．電子ボードを扱う場合には，書き込みや保存などの機能の切り替えをマルチタッチで行うことで，腕や体の移動を最小限にして文字や図の遮蔽を防止するソフトウェアがある [1]．しかし，この方法では，講義者の様子は見ることができるが，電子ボードの前に講義者がいることで遮蔽された文字を見ることができない点や，電子ボード以外では適用できない点が挙げられる．このような問題を解決するために，本システムでは受講者が講義者によって遮られた文字と，講義者の動きを同時に見ることができる「半透明人間カメラ」を提案をする．このシステムにより，深度情報により講義者を判別して半透明化を行い，スクリーンに出力することで，受講者は，スクリーンを通して講義者の陰にある文字および講義者の様子を同時に把握することが可能となる (図 1)．

2　既存システムの問題

Kinect を用いて人物を認識し，その部分を背景画像によりマスクすることで人間の透明化を行うシステムが存在する [2]．しかし，このシステムでは，講義者を透明化するときに講義者の輪郭が残ってしまう問題や，背景画像の更新がないので透明化された領域では黒板の文字が更新されないという問題がある．また，講義者の容姿を完全に消去することは，それを見ている受講者にとって不自然であり，注意が散漫してしまう恐れがある．そこで本システムでは，講義者のいる領域を拡大した後，講義者の消去および色付けを行うと共に，講義者以外の領域を常時更新するというアプローチを図った．

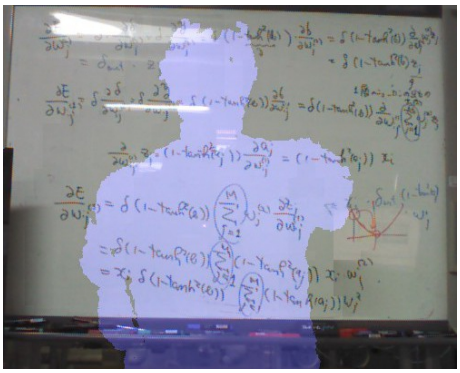


図 1. 半透明人間での講義

3　システムの概要

図 2 に本システムの構成を示す。本システムは，カメラにより RGB・深度情報を取得し，取得した情報を基に PC により講義者を半透明化を施した後，スクリーンに出力する処理をリアルタイムで行う．システム起動時，カメラには黒板が画面中央に映るように，カメラを配置する．なお，システム起動時に，背景領域と判断するための距離を決定するため，講義者は画面に映らないようにする．PC では，取得された深度データにより，講義者をマスクする領域と，黒板や文字などの背景領域を判別し，講義者の半透明化および黒板に書かれた情報の更新を行う．PC により講義者が半透明化された映像は，スクリーンや外部モニタなどに出力される．受講者は，出力された映像を見るだけで瞬時に，黒板の文字と講義者の様子を捉えることが可能となる．

Copyright is held by the author(s).

* Yamane Yoshikazu and Umemura Kyoji, 豊橋技術科学大学 情報・知能工学系, Okabe Masayuki, 豊橋技術科学大学 情報メディア基盤センター

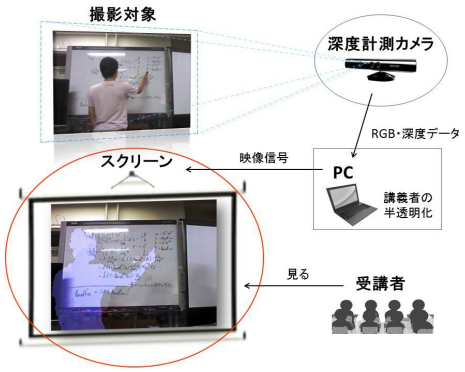


図 2. システム構成

4 システムの実装

本システムは，入力装置として Microsoft 社の Kinect を用いて実装を行った．kinect により取得した深度データを基に，マスク領域と背景領域を割り当てるが，その方法として，深度データの閾値を用いて領域を判定した．深度データの閾値の決定は，画像中央付近の深度データを優先的にサンプリングし，それらの平均値により決定しているので，黒板がカメラ画面の中心にある場合，カメラと黒板の間の距離が閾値となる．ある画素における深度は，閾値以上であれば背景領域，閾値未満であればマスク領域として割り当てられる．なお，背景領域は黒板やホワイトボードなどに相当し，背景画像として随時更新する領域である．また，マスク領域はカメラと黒板の間に存在するオブジェクト (講義者など) に相当し，そのオブジェクトの背景領域によりマスクを行い，オブジェクトを色付けすることで半透明化を行う領域である．最新の画像に対して，それぞれの領域に応じた処理を行うことで，講義者を半透明化することが可能となる．

4.1 深度画像の分割

Kinect からの各種情報は 30fps で取得しているが，各フレームに対してすべての画素を探索し，1pixel ごとの処理を行うことは，メモリ使用量が多くなることや，計算時間が増加することで処理が重くなるため現実的ではない．そのため，pixel ごとの深度データをいくつかのブロックにまとめることで，扱う画素を減少させた．なお，1 ブロックは 10 × 10 ピクセルとして深度画像の分割を行い，各ブロック内でサンプリングデータを選ぶ．そして，そのブロックが背景領域またはマスク領域のどちらに属するのかを多数決を行うことで決定し，各ブロックの値を背景領域またはマスク領域に割り当てる．これにより，領域の割り当てを高速化することができるので，講義者の半透明化および背景の更新をリアルタイムで処理することが可能となる．

4.2 マスク領域の拡大

領域を割り当てた際に，背景更新領域とマスク領域の境界面である講義者の輪郭部分が，領域の誤分類により半透明処理されずに表示されてしまう．例えば，図 3 左中央の丸で囲んだ部分のように，腕や手にマスク領域の割り当てがない箇所が存在するので，この部分が消去されずに残ってしまう．そこで，マスク領域の拡大を行うことで問題を解決した (図 3)．本システムでは，余分にマスク領域の割り当てを行っても，半透明化の際にオブジェクトのみを色付けするため，システムの動作に影響はしない．深度データにより割り当てられたマスク領域に対して，それぞれ 8 方向に 1 ブロックだけシフトしたマスク領域を作成し，それらの論理和を取ることでマスク領域の拡大の実装を行った．マスク領域を拡大することで，講義者を消去する領域が広がるため，輪郭を残すことなく講義者を消去することができる．マスク領域により消去した後，オブジェクトに色付けを行うことで半透明化することができる．



図 3. マスク領域の拡大．(左):元画像，(右):オブジェクトの色付け

5 まとめ

本研究では，講義などの場において，受講者が，講義者によって遮られた文字と，講義者の様子を同時に見ることができるシステムの開発を行った．既存のシステムにおける，透明化した際に講義者の輪郭が表示されてしまう問題や，講義者の様子が把握不可能になる問題，および背景が更新されない問題に対して，マスク領域の拡大と色付け，背景領域の常時更新を行うことにより解決した．本システムによって講義者の半透明化を行うことにより，実際の講義などの場で，講義者により遮断されて見ることが不可能であった文字，および講義者の様子を見ることができる画期的な講義が可能となる．

参考文献

[1] 永野直, 栗原一貴, 渡辺裕太, 藤村裕一, 皆月昭則, 林秀彦. マルチタッチ電子黒板による教材提示性能の改善, 鳴門教育大学情報教育ジャーナル No.7, pp. 23–27, 2010.

[2] 中村薫, 田中和希, 宮城英人. 秀和システム. Kinect for Windows SDK プログラミング C# 編, pp. 153–162, 2012.

浮遊ノートシステムにおける道案内機能の実装

小野澤 清人　　入江 英嗣　　吉永 努 *

概要. ロボットが活躍する場は，産業のような専門的な環境が多く，操作するためには特別な知識や技術が必要であった．しかし，近年になり自動お掃除ロボットのような誰でも簡単に操作することができるロボットが登場し，人間の生活をサポートし始めている．我々は，ロボットがユーザの様々な手助けを行う「浮遊ノート」システムの提案・実装を進めており，このシステムにお遣い・道案内機能を持たせるためのナビゲーションシステムを構築した．このシステムでは，スマートフォンと連動して地図タップという簡単な操作でロボットの目的地を設定し，自動設計された経路をロボットが自律飛行する．しかし，実装したアルゴリズムでは，狭い道など GPS が取得しづらい地点での到達失敗が多い．また風の影響により進む方向が変化し，目的地に到達しないことがある．そこで本稿では，飛行ロボットの移動推定や加速度センサを用いたフィードバックシステムを実装し，提案ナビゲーションシステムのロバスト性を高める．加速度センサ及び過去の GPS の値から現在地を推定し，実際の値と大きな差がある場合には，GPS の測定ミスと判断し，推定値よりナビゲーションを行う．また，操縦指示と異なる機体の動きを加速度センサにより検知し，移動に補正をかける．これらのシステムを構築・導入することで，ナビゲーションシステムの精度向上を行い，道案内機能を実現する．

1　はじめに

ロボットは原子力施設の保守や宇宙開発，産業など専門的な環境での利用が主流であり，これらを操作する人間にも特別な知識や技術が必要とされてきた．しかし近年になり，自動掃除ロボットのように誰でもできる簡単な操作で，人間の生活をサポートするロボットが登場している．我々は，ロボットが人間のパートナーとして存在し，ユーザを様々な場面で手助けする将来を目指し，「浮遊ノート」システムを提案している [1]．このシステムは，携帯情報端末を飛行ロボットが把持してユーザの望む位置に滞空することにより，両手が使えない場面や目が離せない場面でも容易に情報端末を利用できることを目指したお供ロボットである．現在我々は「浮遊ノート」システムの更なる機能拡充を目指して，お遣い・道案内機能を開発している．既に行った実装では，手持ちのスマートフォンと連動し，ユーザが目的地を地図上でタップすることにより，経路を自動生成し，GPS 及びコンパスセンサを利用してロボットを目的地へ到達させる [2].

スマートフォンでの目的地設定は，どのユーザに対しても簡単な操作を実現したが，ロボットの目的地到達率は 50% であった．主な原因は建物沿いの狭い通路など GPS の精度が悪い地点での不正確な位置データであった．また，風により機体が流されそうになった場合は GPS による制御が追いつかずに想定外の場所へ飛行してしまった.

自動ナビゲーション可能な浮遊ノートによるお使い機能実現とそのアプリケーションを検討するためには，これらの課題を解決しなければならない．そこで本稿では風に流されたときの補正及び不正確な GPS データ取得の対策を実装する．機体が風に流され本来の指示とは異なった方向に進むことを加速度センサにより検知し，補正を加えナビゲーション能力の向上を図る．また，GPS データの誤取得を防ぐために GPS の値の移動推定を行う．取得した GPS の値と推定値に差が見られる場合は，GPS の精度が悪いと判断し，推定値を用いてナビゲーションを行う．これらの提案より，ナビゲーションシステムの目的地到達率の向上を行い，道案内機能の実現を目指す．目指している道案内機能は，携帯情報端末で現在地や目的地を確認しなくても，ロボットの後を追えば目的地に到着することのできるシステムである.

2　提案手法

2.1　加速度による移動方向フィードバック

提案する動作を図 1 に示す．加速度センサを利用することで，飛行ロボットの速度を算出できる．ナビゲーションにより前進や旋回の指示を出すときの加速度の値と，飛行ロボットが指示を受けた後の加速度の値を比較し，受けた指示とは異なる方向に進んだことが検知できた場合，風に流されたと判断し，誤って進んだ方向とは逆向きの移動指示を出すことで補正する.

Copyright is held by the author(s).

* Kiyoto Onozawa, Hidetsugu Irie, Tsutomu Yoshinaga, 電気通信大学 大学院情報システム学研究科 情報ネットワークシステム学専攻

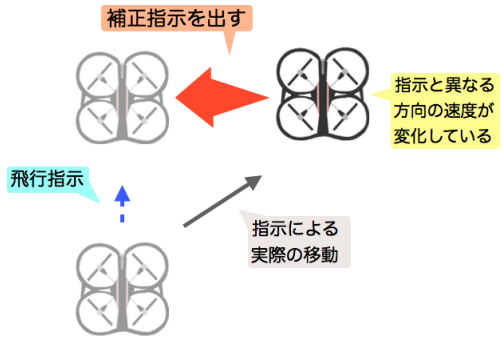


図 1. 加速度センサによる補正

2.2 GPS による移動推定

ナビゲーションアルゴリズムの基本動作では，その時の GPS 及びコンパスデータより目的地までの距離及び角度を計算し，前進または旋回の指示を出して目的地到達を目指す．このナビゲーションの実行中に GPS の測定に外乱があることを想定し，それまで取得した飛行ロボットの速度データ及び過去の GPS データより現在値を推定する．図 2.2 に推定を行いながら，移動する例を示す．

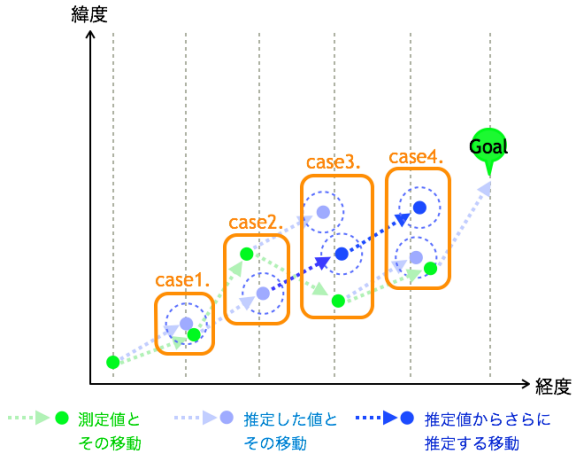


図 2. GPS の実測値と推定値の推移

case1 のように，測定値と推定値が近似する状態の時，測定値を利用してナビゲーションを継続する．

case2 のように測定値と推定値の差が大きい状態の時，GPS の測定ミスと判断し，推定値をロボットの現在地と想定したナビゲーションを行う．次の地点の推定は今回の推定値を元としたものと，測定値を元としたものの二通りを用意する．

case3 は，全ての推定値と測定値が異なる状態を示す．この場合は，GPS の測定ミスと判断し，2 つの推定値のうち測定値に最も近い推定値を現在として想定し，ナビゲーションを行う．次の地点の推定は測定値及び採択された推定値を元に計算する．

case4 では，いずれかの推定値と測定値が近似する状態を示す．この時，近似した推定が正しいと判断し，測定値を利用してナビゲーションを行う．次の地点の推定には測定値を利用する．

3 実装と評価方法

ナビゲーションシステムでの飛行ロボットへの指示は，旋回と前進のみ行われる．加速度センサにより指示以外の移動を検知した場合，その方向とは逆向きの方向へ移動する指示を出す．

飛行ロボットに命令を出す際，どの程度の速度で移動するかを決定することができる．そのため，移動方向フィードバックを行う際，移動前後の速度比較により，どの程度異なる方向に進んだかを計算し，進んだ距離に応じた速さで移動する命令を出す．

GPS データの現在地推定において，GPS の値が更新される時間の間隔及びこの時間内の移動速度から緯度，経度の推定値を算出する．この結果を利用し，提案手法で示した移動推定を行う．

これらの制御アルゴリズムを浮遊ノートに追加し，ナビゲーションシステムを利用した実験を行う．直線経路，曲がる地点を含む経路，GPS を取得しづらい地点を目指す経路それぞれの経路をナビゲーションさせ，目的地到達率がどのように改善されるか評価する．さらに，道案内アプリケーションの実験を行い，ユーザが飛行ロボットの跡を追うことで指定した目的地にたどり着けるかどうかを検証・評価する．

4 未来ビジョン

移動方向フィードバックシステムにより，安定して自律飛行する浮遊ノートシステムが実現すれば，お供ロボットによる様々な生活支援アプリケーションを安定して実装することができる．このことによりユーザの荷物を任意の場所に届けることや，ユーザが車内の渋滞などで身動きが取れない時に，代わりに任意の地点の様子を見に行かせるなど，人々を様々な形でサポートすることができるようになる．我々は，このようなナビゲーションを行うアプリケーションを開発し，「浮遊ノート」システムを発展させ，ロボットが人間のパートナーとして存在する未来を目指す．

参考文献

[1] 芝星帆:“飛行型ロボットと携帯情報端末を利用した「浮遊ノート」システムの提案”, 情報システム学研究科 情報ネットワークシステム学専攻 修士論文 (2012).

[2] 小野澤清人:“タッチ指示によるお供ロボットナビゲーション”,DICOM2013 シンポジウム (2013).

[3] Tianmiao Wang:“Autonomous Control And Trajectory Tracking Of Quadrotor Helicopter”, IEEE International Conference(2012).

日本ソフトウェア科学会　研究会資料シリーズ ISSN 1341-870X No. 73
2013 年 12 月 4 日発行

第 21 回　インタラクティブシステムと
ソフトウェアに関するワークショップ（WISS2013）論文集