

ぱらぱらアニメによる体験データの要約・編集支援システム

System Of Summarization And Edit Experiences By Using Animated Cartoon

小関 悠 角 康之 西田 豊明 間瀬 健二*

Summary. We proposed the system which summarize, edit and share experiences in this paper. First, this system detects meaning scenes of large video data from wearable and ubiquitous cameras. Then, video data is summarized and arranged by converting to animated cartoons which are easier to check through in a short time than video format. Users also can edit a layout of these cartoons with a own intention. And everyone can view a output of this system, because it don't depend any special hardware or software. So, it allows users to share experiences easily by using internet, especially e-mail and blog.

1 はじめに

昨今、ビデオカメラや各種のセンサー機器が小型化し、また安価になったことから、これらを組み合わせることで、体験データを容易に取得することの出来る体験記録システムが複数提案されている [1][2]。ここでいう体験データとは映像・写真・音声・センサー情報・ドキュメントなど、複数のメディアから構成されたデータのことであり、体験者本人が後で観賞することにより、従来のビデオデータなどより詳細に振り返ったり、新しい気付きをより得ることが出来るようなものを指す。また体験データには体験者以外の第三者が観賞することで、体験を共有し、新しい観点を見出す、といった用途も考えられる。

こうした体験記録システムでは、膨大な量のデータを容易に得ることが出来る。一方でこうしたデータを鑑賞しようとした場合、現状では依然として、ビデオを見たり本を読んだりというようには容易に行えない。原因としては大きく二つが挙げられる。

一つには、得られるデータがユーザーが直接扱うにはあまりに大量であるためである。ホームビデオカメラ一台で運動会を撮影していたような場合と異なり、多くの体験記録システムでは複数のビデオカメラを用いた常時撮影により大量の映像データを取得するため、それらの中から目的の場面を感覚的に探したり、目視でデータ全体を確認したりすることは非常に困難になっている。このため、データからあまり意味のない部分を事前に排除する、一覧性の高い状態でデータを提示する、ユーザーの望む観点で簡単にデータを並べ変える、といった要約・編集技術が新しく求められる。

もう一つには、これまでに提案されてきた体験記

録システムでは、得られたデータを鑑賞するために特殊なディスプレイや専用のデータベースなど、特別なハードウェアやソフトウェアを必要とすることが多いという現状がある。結果、興味深い体験データを得ることが出来たとしても自由に鑑賞出来るのはごく少数に留まり、重要な用途の一つであるはずの体験共有が非常に困難となっている。

このような背景から、ユーザーが簡単に利用可能で、特別なハードウェアやソフトウェアを必要とせず多くの人が観賞出来る、体験データの要約・編集・共有システムが求められている。

そこで我々は、自動的に映像データの要約を行い、ユーザーに分かりやすい形で表示し、かつユーザーが好みの観点で体験を振り返ることが出来るというシステムを提案、実装する。得られた体験データはインターネット上で簡単に配布出来る上、観賞に特別なソフトウェアやハードウェアを用いないため、標準的なインターネット環境では誰でも簡単に人の体験データを共有することが出来る。なお、体験データ取得の基となる体験記録システムには体験キャプチャシステム [3] を用いている。体験キャプチャシステムはウェアラブルカメラと設置型のカメラを組み合わせることで同一シーンを複数視点で撮影する一方、同様にウェアラブルセンサーと設置型のセンサーを用いることで「話す」や「見る」といったインタラクションとその対象を自動的に認識する、という特徴がある。

以下、本稿では本システムを用いた体験データの生成から共有までを順に説明する。本システムの概要は図1の通り。まず、体験キャプチャシステムから得られたデータを、シーンという概念を用いながら要約する手法について述べる(2章)。次にぱらぱらアニメの技法を用いて、ユーザーに分かりやすい形で体験データを提示する手法について説明する(3章)。以上によりユーザーは一覧性の高い体験データを容易に得ることが出来る。さらに得られた体験

© 2005 日本ソフトウェア科学会 ISS 研究会。

* Yu Koseki, Yasuyuki Sumi, Toyooki Nishida (京都大学情報学研究科 / ATR メディア情報科学研究所) and Kenji Mase (名古屋大学情報連携基盤センター / ATR メディア情報科学研究所)

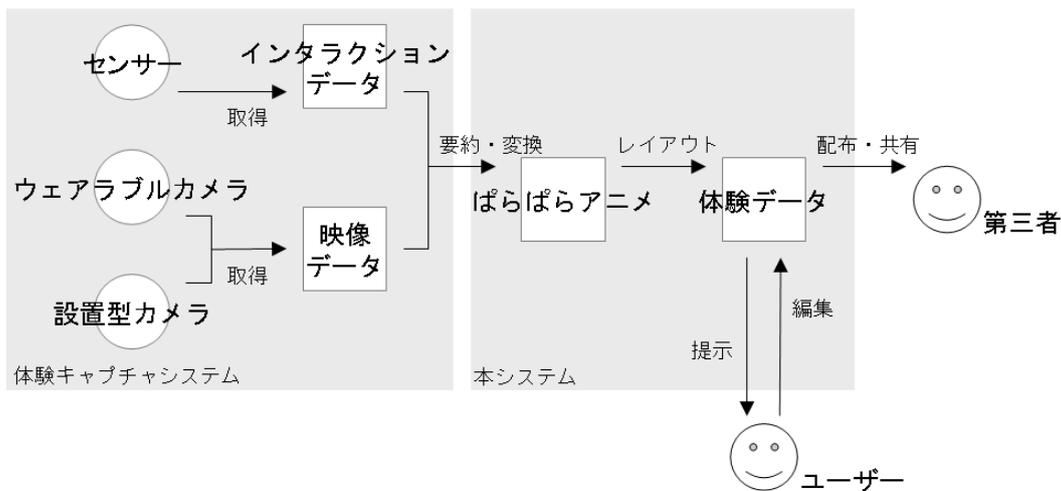


図 1. 体験キャプチャシステムと本システムの概要

データはユーザーが好きな形に編集出来ること、これにより体験が様々な観点で見直せることについてまとめる(4章)。最後に、以上のような過程で編集された体験データが第三者へ容易に配布可能であることと、その利点について簡単に述べる(5章)。

2 データの自動要約

本システムでは、体験とは複数のシーンの列なりであると仮定している。例えばある一人の博物館来場者の体験を記録した時、その体験は「絵画Aを見ながらXさんと話す」「絵画Bの前でYさんの話を聞く」「絵画Cと絵画Dの前をうろろする」...というような意味のある複数のシーンに分割される、といった具合である。本システムはウェアラブルカメラや設置型カメラを用いる体験キャプチャシステム特性上、展示会や説明会などの閉じた空間でのイベントでの体験記録に適正を持っている。

本稿でいうシーンとは、体験を時間的に、その意味に応じて適切に分断したものと定義出来る。このシーンという概念を用いると、体験データを適切に要約するということは、適切にシーン分けし、かつそれぞれのシーンの重要度を適切に定めることである、と言い換えられる。

以前、我々は体験キャプチャシステムから得たインタラクション・データに対して、前後のインタラクションとの関係性を基に群平均法を用いたクラスタリングを行うことで、映像データが意味のあるシーンの列なりへ自動的に分解されるというアルゴリズムを提案した[5]。本システムでもこの手法を踏襲し、大量の映像データやセンサーデータから、自動的に意味のある体験データを得る。例えば図2では、5つのインタラクションが「Aと話す合間にBと話す」シーン1と「Cと話す」シーン2に分割されている。



図 2. シーン分けの例

こうして得られた個々のシーンの重要度は、クラスタリングにおけるクラスター(=シーン)の大きさと近似的に見なすことが可能だろう。例えば図2の場合では、シーン1の方が複数のインタラクションを含んでいるので、シーン2より重要であると考えられる。ただし、含まれるインタラクションの種類に応じて、シーンの重要度は若干異なるとも考えられる。例えば「誰かに見られていた」シーンより「誰かを見ていた」シーンの方が、当人にとっては重要と推測出来る。また「誰かを見ていた」だけのシーンより「誰かと話していた」シーンの方が、一般的にはより重要なイベントが存在すると予測可能である。

以上のような考えにより、本システムでは個々のインタラクションについて表3のように点数付けを行っている。シーンの重要度は、そこに含まれるインタラクションを表に従って点数化した、その合計とする。図2の場合、シーン1は2点のインタラクション4つから構成されているので合計8点である。同様にシーン2は「Cと話す」だけから構成されているので2点になる。

なお、体験キャプチャシステムでは、センサー情報から自動的に解釈されるインタラクションとして「誰か(何か)を見る」「誰かに見られる(どこかにいる)」「誰かと見つめ合う」や「誰かに話す」「誰かに話される」「誰かと話し合う」といった種類が存在する。よって自動要約によって得られるシーン

インタラクションの種類	点数
見る	1点
見られる	0点
見つめ合う	2点
話す	2点
話される	2点
話し合う	4点

図 3. インタラクションの点数化

はその合成,例えば「Aさんと話したあと、Bさんに話しかけられる」や「展示物Cの近くでDさんと話し合う」といった種類になる。

3 体験データの提示

映像データは原則的に、撮影時間と同じだけの鑑賞時間が必要となる。もちろんノイズや不要な場面をカットしたり、早送りやメタタグを用いた頭出しのような機能を用いることで、鑑賞時間のある程度削減することは可能である。しかし何にせよ一貫性が低く、検索に不向きであるため、はじめに述べた通り膨大な量のデータを扱う場合は、映像という形式は扱いづらい。

こうした問題点を解決するため、本システムでは映像データをばらばらアニメに変換して用いる。これにより必要な鑑賞時間を大幅に減らすことを可能にしている。結果、例えば一日の出来事を大雑把に把握したり、あやふやな記憶の中から目的のシーンを選び出すというような、膨大な映像データが対象である場合は非常に難しい作業を簡単にこなすことが出来る。

ばらばらアニメへの変換は以下の順で行う。まず図4のように、それぞれのシーンが含む個々のインタラクションについて、センサーの発火時間に合わせた静止画を映像データから切り出す。体験キャプチャシステムでは通常、一つのインタラクションは複数のセンサー発火から意味付けられているので、図中「」のような、なるべく個々のインタラクションの中心に存在するセンサーの発火に合わせて静止画を切り出す。切り出しをセンサーの発火時間に合わせるのは、その瞬間にセンサーを持ったインタラクション対象がカメラに捉えられている可能性が高いからである。

図5左はこのようにして切り出された静止画の例である。センサー発火に合わせて切り出したため、インタラクションの対象である人物が大きく画像の中心に表示されていることが分かる。

次に、より明示的にインタラクションの対象を示すため、センサーの位置情報を用いてハイライト処理を行う。図5右は図5左にハイライト処理を行った結果である。インタラクションの対象でない左側

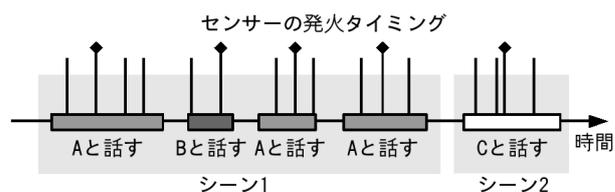


図 4. センサー発火と静止画の切り出し



図 5. 切り出された画像とハイライト処理

の人物がモノクロになって潰れていることが見てとれる。これにより、複数の人物が画像に存在しても、誰がインタラクションの対象であるかを簡単に見てとることが出来る。

続いて、こうした画像をシーンごとにゆっくり順に切り替える、ばらばらアニメへとまとめる。つまり、一つのばらばらアニメは一つのシーンの画像を順番に表示していく。これにより映像データの意味的情報を失わず、またある程度の臨場感を保ったまま、時間的には大幅に圧縮して閲覧することが可能となる。ばらばらアニメは先に演算したシーンの重要度に従って、大・中・小の三段階のサイズをとる。

最後に、このようにして得られたシーンごとのばらばらアニメを、左から右、上から下へと時間に沿って並べて表示する。図6は本システムによって最初に表示される体験データの例である。複数のシーンがそれぞれの重要度に応じたサイズで、漫画のように整然と並べられていることが分かる。データの要約から表示に至るここまでの過程は、全て自動的に、短時間の間に行われるため、体験記録システムに参加したユーザーは図6のような体験データをすぐに手にすることが出来る。

4 体験データの編集

このようにして得られた体験データは、しかしあくまで体験キャプチャシステムのデータから自動的に導かれたものであり、ユーザーの意図は全く反映されていない。そのため体験データによっては、例えば自分にとって重要と思われたシーンが小さく表示されていたり、あるいは一続きであると考えていたシーンが複数に分割されて表示される場合がある。また通常、シーンは時間順に並べられているが、人によってはこれはインタラクションの対象ごとに並べた方が何か分かることがあると思うかもしれないし、あるいは全く異なるレイアウトをとることが出



図 6. 生成された体験データ

来れば意外な発見があるかもしれない。

よってこれまでに述べた手法で得た、シーン分けとそれぞれのシーンの重要度、配置は全て一つの叩き台として捉え、最終的には体験記録システムのユーザー本人が、それぞれの意図でもって体験データのレイアウトを自由にデザイン出来るようにすべきである。

本システムはこうしたニーズに応えるため、得られた体験データのレイアウトを簡単に編集することが出来る。例えば、それぞれのシーンはドラッグにより好きな場所に移動可能である。移動先に別のシーンが存在した場合、二つのシーンは自動的に結合される。逆にシーンの中から適当なインタラクションを選び出し、別のシーンとして切り出すことも可能である。加えてそれぞれのシーンの重要度、すなわちサイズも変更出来る。こうした操作は、全て通常のマウス操作だけで簡単に行える。

具体的には、図7がシステムで表示されるシーンの例である。ハイライト処理された静止画にカメラの種類や時刻が添えられ、さらに会話のあったインタラクションであれば、その長さに応じた大きさの吹き出しが話者の近くに出現する。シーンはマウスで選択されると図8のような縁が赤くなり、縦に三つのアイコンが新しく表示される。これは上からそれぞれ「シーンの早送り/巻き戻し」「シーンの移動」「シーンから現在表示の画像を切り出し」(図9)の機能を示しており、クリックやドラッグでそれ

ぞれの操作が可能である。

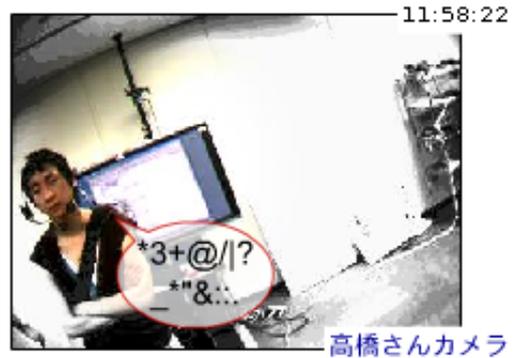


図 7. 表示されたシーン



図 8. 選択されたシーン

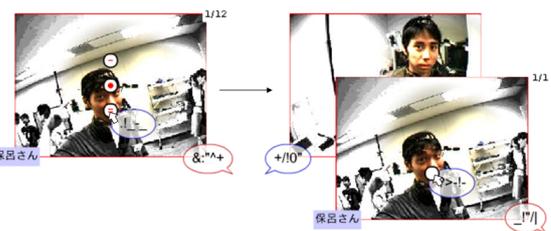


図 9. ドラッグによるシーンの切り出し

より高度な操作としては、シーンを時間やインタラクションの対象に応じてソート、あるいは分類し直す機能がある。図10は全てのシーンを時間(左右)とインタラクションの対象(上-下)で分類し直した時のレイアウトである。序盤(図10左端部)では様々な対象とインタラクションを交わしていたが、終盤(図10右端部)では限られた対象とだけインタラクションを交わしていたこと、また中盤まで特定の対象とずっとインタラクションを交わしていたこと(図10下端部)などが見てとれる。



図 10. 並べ変えられたレイアウト

こうした機能により、体験記録システムのユーザーは、容易に自分の意向を反映させた体験データのレイアウトを作成することが出来る。例えば特定の人物との会話シーンだけを集めたり、例えばある時刻のインタラクションだけを集中的に表示させたり、関係のある人物を近付けて並べたりというレイアウトが考えられる。より具体的には、展示会のポスターに興味を持った順に目立つよう配置させたり、学会でどの時間に忙しく動いていたかなどが分かるように並べたり、ということが可能になる。

なお、それぞれのレイアウトは好きな時に複数保存出来る。保存されたレイアウトはまた、好きな時に呼び出して再編集することが可能である。つまり一つの体験データに対して、様々な観点で作成したレイアウトを保存しておくことが出来る。体験データは通常どれくらい前に体験したのかによって、受ける印象が異なっていく。本システムではその差異を楽しむため、かつての印象で作成したレイアウトや、今の印象で作成したレイアウトを同時に保存し、閲覧することが可能になっている。

5 体験データの共有

以上のようにして得られた体験データとそのレイアウトは、ひとまとめのファイルとして生成される。また本システムの体験データ観賞・編集部分は Macromedia Flash で実装されている。このため、通常的环境下では体験データを得るだけですぐに観賞出来る。つまり、体験データのファイルをまとめてインターネットにアップロードしたり、あるいは電子メールに送付することで、第三者も簡単にその体験を楽しむことが出来るということである。また体験データを得た第三者は、体験者の用意したレイアウトを楽しめる一方で、体験者同様に全く新しいレイアウトを作成し、新しい観点を持ち込むことも出来る。

今日、インターネット上では種々の個人的なデータが公開され、共有する枠組みが出来上がりつつある。例えば、各種の blog システムは Trackback や Permalink のような実装により、日記という一種の体験データを容易に公開し、ネット上で共有することが可能になった。bloc (<http://www.bloc.jp/>) のようにイベント情報を共有するシステムや、Flickr (<http://flickr.com/>) のように写真を共有するシス

テム, Google Maps (<http://map.google.com/>) のように地域情報と共有するシステム, さらには orkut (<http://orkut.com/>) や mixi (<http://mixi.jp/>) のように個人情報そのものを共有するシステムまでもが生まれている。こうした状況を鑑みると, 体験記録システムがより普及すれば, 体験データも同様に共有の道を歩むと考えられる。例えば美術館での体験データをそのまま個人の blog に掲載したり, 学会発表者の体験データをまとめて学会のサイトに載せたり, というような使い方がそう遠くない時点で実現するかもしれない。

6 おわりに

本稿では, 体験記録システムによって得られた大量のデータに対し, シーンという概念を用いながら映像データをばらばらアニメに変換することで, 従来に無いような分かりやすい体験データ観賞システムが実装出来ることを説明した。また体験データ観賞システムは同時に編集システムとして, ユーザーがそれぞれの観点で自由にレイアウト可能であることと, これにより可能な展開についてまとめた。最後に特別なソフトウェアやハードウェアなしに体験データが共有可能ということについて述べた。

今後は様々な場面で体験記録を行うことで, 体験データの要約・編集・共有がどのように発揮されるかについてより深く調査したい。

7 謝辞

本研究は情報通信研究機構の委託研究「超高速知能ネットワーク社会に向けた新しいインタラクション・メディアの研究開発」により実施した。

参考文献

- [1] 相澤 清晴: ライフログの取得と処理 -ウェアラブル, コピキタス, 車-, 人工知能学会全国大会, 2005年6月
- [2] 河村 竜幸, 福原 知宏, 武田 英明, 河野 恭之, 木戸 出 正継: 実世界対象物とのインタラクションによる体験映像の整理, 日本認知科学会第20回大会, pp.240-241, 2003年
- [3] 角 康之, 伊藤 禎宣, 松口 哲也, Sidney Fels, 間瀬 健二: 協調的なインタラクションの記録と解釈, 情報処理学会論文誌, Vol.44, No.11, pp.2628-2637, 2003年11月
- [4] 坂本 竜基, 角 康之, 中尾 恵子, 間瀬 健二, 國藤 進: コミックダイアリ: 漫画表現を利用した経験や興味の伝達支援, 情報処理学会論文誌, Vol.43, No.12, pp.3582-3595, 2002年12月
- [5] 小関 悠, 角 康之, 西田 豊明, 間瀬 健二: シーン推定と漫画技法を用いた体験要約システム, 人工知能学会全国大会, 2005年6月