srt.js: 映像コンテンツへの IoT 指向拡張プログラム埋め込みフレームワーク

栗原 一貴*† 橋本 美香*

概要. 本論文では代表的な時系列コンテンツの一つである映像コンテンツに対し，時刻情報の付与されたスクリプト言語を埋め込むことにより，映像と連動したプログラムを駆動しエンタテインメント価値を高める手法の再検討を行う．具体的には YouTube 等の映像コンテンツに対し公式にサポートされている字幕形式である srt ファイルの形で Javascript を埋め込み，映像のタイミングに合わせて外部 Web サービスや IoT 機器と連携することを容易に行えるフレームワーク srt.js の開発を行った．種々の具体例により本手法の有効性を示し，問題点と今後の展望について整理した．

1 はじめに

IoT (Internet of Things) という言葉で表わされるように，身の回りの様々なモノがインターネットに接続しインテリジェントに振る舞い我々の生活を豊かにする社会が到来しつつある．一般市民が趣味の工作として IoT 機器を制作するための要素技術も充実してきており，Intel Edison，Raspberry pi などの小型・安価な汎用コンピュータや，Sony MESH や webmo のように専門知識をほとんど必要とせずに気軽に IoT システム開発を行えるプロダクトも登場してきている．

そのような今日において，エンタテインメント用途のプログラミングは二極化している．一つは全てをことこまかにプログラミングした情報システムである．商用のゲームなどはこれに属する．多様なことが実現できる反面，システム開発のコストは大きい．一方，もう一つは IF-THEN ルールに単純化させたプログラミングである．IFTTT や myThings，および Sony MESH などは「もし〜がおきたら，〜する」という形のプログラミングを GUI 等で行うことによって，スマートフォンの機能や有名 Web サービスを連携したアプリの気軽な開発を可能にしている．

本論文は，人類がこれまで慣れ親しんでいた時系列コンテンツに対し現代的な機能拡張方法を検討するものである．ここで時系列コンテンツとは，提示される情報の順番が基本的には決まっていて，かつ比較的長い時間継続するコンテンツを指すこととする．音と動画像からなる映像コンテンツやプレゼンテーションなどがこれに該当する．これらの扱う情報は一方通行的にユーザに届けられ，またモダリティが限定されていたが，近年の各種 Web サービスや IoT 機器の普及により，時系列コンテンツの進行に連動してエンタテインメント価値を高められる可能性が広がった．たとえば我々がこれまでに発表した，プレゼンテーションの進行に合わせて予め準備した内容の Twitter の投稿を行う「ぴぴつい[1]」や，映像に応じて物理的な開放状態を調整する IoT 機器である「開放度調整ヘッドセット[2]」は時系列コンテンツを拡張するシステムの事例として挙げることができるだろう．

しかし可能性が拓かれた一方で，現状ではそのような複合的な時系列コンテンツを開発しようとした場合の汎用的な仕組みが不足している．それほどのインタラクション頻度は無いにもかかわらず，汎用のプログラミング環境を用いた完全にゼロからの開発を行うのは煩雑であるし，IFTTT などのような IF-THEN ルール型の簡易プログラミング環境は時系列コンテンツ拡張には対応していない．

そこで我々は，かつて Adobe Flash などで標準的に用いられていた，時系列コンテンツに連動した任意のプログラムをそのコンテンツの時系列上に埋め込み，鑑賞時にタイミングよくプログラムを実行するパラダイムを再検討し，現代的な形式での開発環境を再提案する．既存の時系列コンテンツ資源を活用し，低コストで多少の機能拡張をしたいと思った場合，ほとんどの制御は「時刻 X になったら処理 Y を実行する」という，互いに独立した WHEN-THEN 型ルールがコンテンツ中に比較的疎に散在する形になる．大部分の計算と通信がほぼリアルタイムもしくはノンブロックで行える現代の計算機環境においては，実際にプログラム実行にどれだけ時間がかかるかはあまり問題にならない場合が多いため，時刻情報とともに実行すべきプログラムを埋め込むのはごく自然な方法であり，開発にも適する方法である．またコンテンツ自体にプログラムが付随することに

* 津田塾大学，† Diverse 技術研究所

よって一元的な開発と管理，配布が容易に行える点も特徴である．

本稿では，時系列コンテンツへの埋め込みスクリプトによる機能拡張について，YouTube 等の映像コンテンツを対象とし，YouTube の字幕機能として Javascript を埋め込んでエンタテインメント性を拡張する手法にしぼり報告する．プレゼンテーションへのとりくみについては次報で報告する．

2 関連研究

時系列コンテンツをスクリプトにより拡張する研究は新規なものではない．Adobe Flash は時系列アニメーションコンテンツのオーサリングにおいて，指定された任意の時刻に実行されるプログラム（ActionScript）を埋め込む機能を実装しており，長年にわたり標準的に用いられてきた．その柔軟な拡張性の一方で，タイムラインと紐付ける形でプログラムを記述するパラダイムには以下の問題があったと分析される．

1. コンテンツがインタラクティブになればなるほどタイムライン通りに進行しないので，タイムラインから「実行の順と時刻を保証するもの」という意味合いが失われる．そのような環境でどのようなプログラム要素も便宜的にどこかの時刻に割り当てなければならないのは直感に反し，開発がやりにくい．
2. 全てのプログラムは，ある開始時刻と終了時刻を付与されて記述される．しかしそのプログラムがその時間内でちょうど実行終了する保証がないため，時間指定が便宜的になってしまう．
3. コンテンツのほとんどを唯一の ActionScript ブロックが占める場合，タイムラインがほぼ不要となる．

特に１は，コンテンツにインタラクティブ性をいかに付与するかが中心的課題だった当時において重要な課題だったと考えられる．Kurihara らによる AfterThought は，flash のもつ１および２の問題に対応し，プログラムを埋め込む拡張として，絶対時刻がなく，順序関係だけが意味を持つタイムラインを提案し，制約解消を用いた柔軟な UI を提案している[3]．また，３についてはのちにタイムラインを用いずに純粋に ActionScript のみでコンテンツ開発を行う環境である Adobe Flex（のちに Apache Flex）が登場して解決された．

本研究では，現代に即したプログラミングのパラダイムとして再度，絶対時刻つきタイムライン連動型のプログラム埋め込み方法の可能性について検討を行う．以前と比較して社会状況が好転した点として，以下が挙げられる．

- タブレット・スマートフォンなどの普及などにより，人々がエンタテインメントコンテンツに接する際の態度がより受動的なものになり，コンテンツに要求されるインタラクションの強度が必ずしも高くないものが認知されるようになってきた．
- 渡邊の「融けるデザイン」[4]にあるように，現代はコンテンツがユーザの時間を奪う時代であり，ユーザは奪われる時間に敏感にならざるを得ない．インタラクティブ性の少ない時系列コンテンツは，ユーザから奪う予定の時間を事前に提示しやすいため，インフォームドコンセントの観点から優れている．
- 計算機と通信の技術革新により，おおよその計算と通信がほぼリアルタイムもしくはノンブロックで行えるようになり，実際にプログラム実行にどれだけ時間がかかるかはあまり問題にならない場合が増えた．
- IFTTT や myThings などの登場により，既存のサービスに「少しだけ」独自要素を追加するプログラミングが受け入れられるようになってきた．時系列コンテンツへの「少しだけ」の独自要素の追加であれば，タイムライン上に疎にプログラムが埋め込まれる程度であるため，タイムラインというメタファーは破綻しない．
- 時系列コンテンツについては，YouTube やニコニコ動画などの動画共有サービスの普及により，独自要素を追加したいコンテンツは豊富に存在し，流通も容易になった．

現在，そのような取り組みとして既に実用化されている事例として，ホラー映画と観客のスマートフォンを連動させ，鑑賞中に実際に電話が着信通知する演出などを盛り込んだ「貞子 3D2 スマ 4D」[5]，主にポルノ映像と性家電を連動させる規格である+1D[6]などが挙げられる．我々はこのような映像連動コンテンツをエンドユーザ・プログラマが気軽に開発できるフレームワークを提供する．

後藤らの Songle API[7]や Kato らの TextAlive[8]は YouTube やニコニコ動画における音楽つき映像を拡張するための道具立てとして，音楽コンテンツの解析情報やブラウザ上でのスクリプト言語の開発・共有を可能にする機能を提供している．我々の提案した開放度調整ヘッドセットにおける音楽コンテンツとの連携動作は Songle API の活用により実現している．また Stepmania Gimmic[9]は音楽ゲームを Lua スクリプトにより拡張するものである．しかしこれらは音楽つき映像コンテンツを対象としており，それ以外の一般の映像コンテンツを扱うものではない．

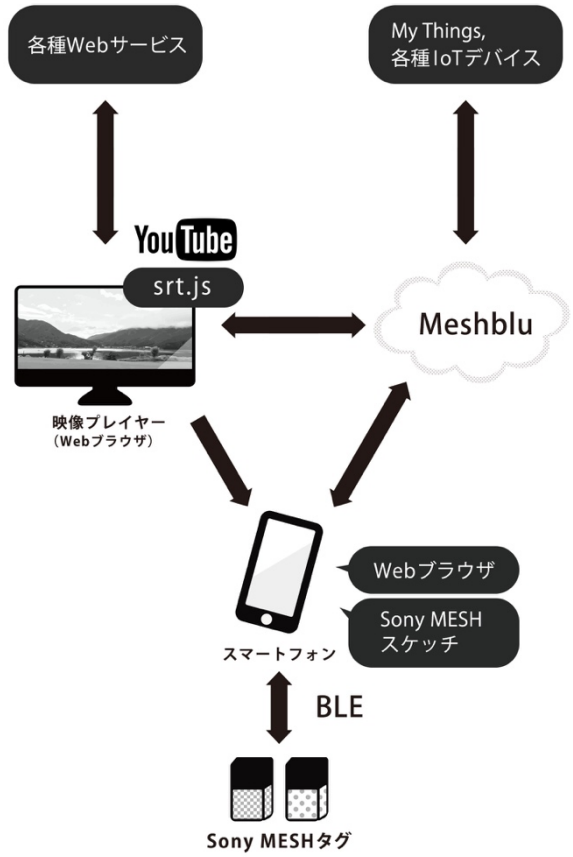

図 1　システム構成

中村ら[10]および松田ら[11]は YouTube の映像再生について，音量，再生速度の調整，効果音，視覚効果の追加等のカスタマイズをブラウザ上で行い，他者と共有するブラウザ拡張システムを提案した．本研究はこれらのように個人的な映像カスタマイズも支援できる点では共通性があるが，Javascript による任意のコード記述が可能であるため，プログラミングの知識は多少必要になるものの映像の再生コントロールの枠を超えたカスタマイズ，例えば IoT 機器との連動やインタラクティブ性の付与等が可能である．

ほかにも IoT 時代に即した新しいプログラミングパラダイムの提案として，以下の研究が挙げられる．加藤らによる f3.js は，IoT 関連システム開発で問題となる，ハードウェア設計とソフトウェア設計を統合的に開発するための環境である[12]．馬場らによる babascript[13]，および橋本らによる goldfish[14]は，人間を実行主体として扱うプログラミング，および実世界を対象としたプログラミングを行う基盤として提案されている．

なお，我々が提案する，既存時系列コンテンツへの「少しだけ」の独自要素の追加は，過去に著者らがゲーミフィケーションの周辺概念として提唱した Toolification of Games[15]と共通性を持つ．すなわちある目的を達成する際，専用の情報システムをゼロから開発するのではなく，一方で目的の達成を助ける外部エンタテインメント要素を小規模に追加する（いわゆるゲーミフィケーション）のでもなく，既存の優良エンタテインメントコンテンツの中に寄生させる形で追加するという点である．

3　提案システム

我々は映像連動コンテンツをエンドユーザ・プログラマが気軽に開発できるフレームワークを提案する．システム構成図を図 1 に示す． Web ブラウザ上の YouTube 等の映像プレイヤーで再生される映像に同期して，各種 Web サービスやスマートフォン，および各種 IoT 機器と通信し映像コンテンツを拡張する．拡張のためのプログラムは YouTube の映像コンテンツに対し公式にサポートされている字幕形式である srt ファイルの形で Javascript を埋め込み，映像のタイミングに合わせて実行することで実現する．以下に特記すべき要素について詳細に説明する．

3.1　srt.js 形式のスクリプト

srt.js 形式は，映像に対し付与される字幕情報のフォーマットとして一般的に普及している srt 形式によって，時刻に連動した Javascript のプログラム実行を指示するテキストフォーマットである．srt は図 2 のように，通し番号（1，5，9 行目），字幕開始時刻・終了時刻（2，6，8 行目），字幕テキスト（3，7，11 行目）を一つのブロックとして記載したもので，空行が区切り記号として用いられている．字幕テキストは改行して複数行にわたっても問題ない．したがって空行さえ用いなければ，字幕テキスト記載部にほぼ任意の Javascript の記述が可能である．字幕と同様に指定のタイミングで指定の Javascript コード群を実行することを指示するこのフォーマットを srt.js 形式と名付ける．図 3 は開始 1 秒後に "Once upon a time (以下略)"とアラート表示する処理などを含む srt.js 形式のファイルの例である．

srt.js 形式のファイルは YouTube 映像の所有者が公式の字幕情報として登録できるため，管理と配布が容易である．その他の srt.js の仕様を以下にまとめる．

- 再生中の映像と連動するため，html5 の動画オブジェクトや YouTube の iframe 埋め込み動画プレイヤーオブジェクトを player という名前で参照できる．
- 今どの字幕区間（すなわち，あるプログラム要素が実行される区間）にいるかを示すグローバル変数 index をもつ．どの区間にも入っていな

い場合は-1 が入る．

- あるイベントリスナが発火した際に，映像の再生箇所に対応した関数を実行するための仕組みである indexedFunction が用意されている．これにより，たとえばスマートフォンをシェイクするなどの同一のユーザ入力に対して，映像の再生時刻に応じた別の処理を行うことが可能である．
- 外部 Javascript を読み込むヘルパー関数である loadScript，および後述する IoT 機器連携機能が予め用意されている．

```
1
00:00:01,000 --> 00:00:04,000
Once upon a time, there was a boy.

2
00:00:05,000 --> 00:00:08,000
He lived in a small city.

3
00:00:10,000 --> 00:00:13,000
One day, he met a girl.

```

図 2　srt ファイルの例

```
1
00:00:01,000 --> 00:00:04,000
// srt.js example
alert("Once upon a time, there was a boy.");

2
00:00:05,000 --> 00:00:08,000
console.log("He lived in a small city.");

3
00:00:10,000 --> 00:00:13,000
var msg = "One day,";
msg = msg + " he met a girl.";

```

図 3　srt.js ファイルの例

3.2　映像と連動してプログラムを実行する映像プレイヤー

YouTube の iframe 埋め込み API を用いた静的なウェブサイトとして映像プレイヤーを実装した．これは *https://srtjs.azurewebsites.net/?v=xxx* にアクセスするだけで使用可能である．ただし xxx の部分には YouTube の映像の識別子である 11 桁の文字列が入る．本サイトはクライアントサイド Javascript のみで実装されている静的なサイトであるため，必要に応じてユーザが自身の管理するウェブサイトに組み込むことも容易である．

既に他者によって公開されている映像に対してユーザが作成した srt.js を連動させたい場合は，srt.js ファイルをネット上でアクセス可能な状態にしたうえで以下のように GET パラメータである surl にその URL を指定することで活用可能である．

https://srtjs.azurewebsites.net/
?v=xxx&surl=(srt.js の URL)

なお，本実装はスマートフォンを含む任意のブラウザで事前準備なく実行可能であり，プレイヤーを表示するウェブサイト全体のデザインができる点が利点である．これとは別に，chrome 拡張を用いて YouTube および Amazon Prime Video の公式サイト訪問時に起動する srt.js 連動映像プレイヤーを現在開発中である．

3.3　Meshblu

映像が再生されている web ブラウザと IoT 機器等を連携するための通信基盤として，IoT 通信プラットフォームである Meshblu を採用した．これは http(s) REST, web socket, MQTT などの様々なプロトコルで IoT 機器間の通信を仲介するサーバである．開発元である Octblu は誰でもすぐに利用できる Meshblu サーバを公開しており，それを活用することも可能である．また，IDCF クラウドを使用すると，Meshblu サーバを通じて IFTTT と類似のサービスである myThings と相互通信可能になるため，多様な機器，多様なサービスとの連携が容易に行える．srt.js では，http(s) REST API もしくは socket.io で Meshblu と接続する．

3.4　MESH 用の開発補助ライブラリ

気軽な IoT システム開発を実現可能な MESH には，Javascript による機能拡張をおこなえる SDK が準備されている．これを用いて，Meshblu と通信し subscribe（Meshblu への情報送信）および publish（Mesublu からの情報受信）が可能なコンポーネントを実装した．Meshblu を経由した myThing と MESH との連携手法は既に myThings 上で公開されているが，ポーリング周期が遅いため最大 15 分程度の通信の遅れがある．開発したライブラリは送受信ともにこの遅れを解決し，次項で述べる動画との連携機能を備えたものである．

3.5　鑑賞中の動画と IoT 機器を連動させる機構

鑑賞中の動画と IoT 機器との通信の確立のため，QR コードおよびスマートフォンを用いた方式を実装した．これは映像の再生時にその再生を一意に指定する ID と Meshblu への接続情報を含んだ QR コードを表示するもので，スマートフォンからカメラで読み取った URL にアクセスすると，動画との相互通信が確立する仕組みである（図 4）．srt.js 側で

は接続が確立した端末の承認リストを管理しており，接続確立後は承認された任意の端末間の通信が可能となる．ここでさらに，接続が確立したスマートフォン上で MESH のスケッチ（プログラム）が動いている場合は，スマートフォンの IP アドレス情報を照合することで鑑賞中の動画と MESH のスケッチとの通信も確立される．

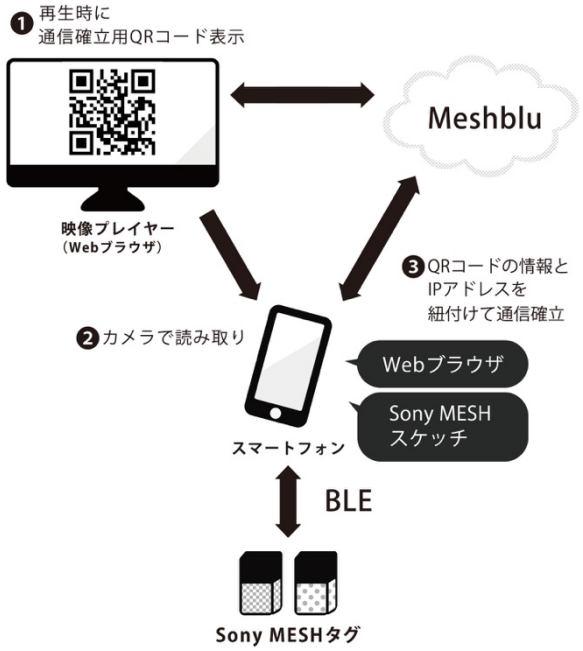

図 4　QR コードを用いたデバイス認証

4　ユースケースシナリオ

本節では提案システムの有効性を実証するために著者らが製作した映像連動コンテンツのプロトタイピング事例を列挙する．これらは一部を除きウェブサイトで公開されており，また論文と同時投稿された参考映像にも収録されているので参照されたい．
https://sites.google.com/site/qurihara/home/srtjs

4.1　ユーザの入力に基づく映像の再生制御

- MESH のボタンタグによる動画の再生と一時停止．
- 顔検出ライブラリを用いた，「見ている時だけ再生される映像」．
- GPS および地図 API を用いた，映像による道案内．
- ヘッドラインニュースにおいて，冒頭のニュース見出し部の各項目を閲覧時にスマートフォンをシェイクするとその項目の本編に移動する．
- ビブリオバトルの映像において，紹介されている書籍に興味を持った際にスマートフォンをシェイクするとその書籍を購入できる Amazon サイトへと誘導される．
- MESH の加速度タグが設置されたティーカップを頭に載せた状態で，それを落とさないようにラジオ体操を行うゲーム．

4.2　映像と同期した表現の拡張

- “Happy birthday to you”の歌における“Happy birth day dear XX”の XX の部分において，予め入力した名前を合成音声により読み上げる．
- QR コードにより通信を確立させた全てのスマートフォンに対し，新年のカウントダウン映像に同期して，“Happy new year!”と表示させる．
- 新年のカウントダウン映像に同期して，MESH の GPIO タグに接続された点火装置により花火を打ち上げる．
- 登場人物が念じる映像を再生すると，付近にあるアヒルの置き物が動く「超能力映像」を MESH の GPIO タグに振動モータを接続して実現する．
- 「スパイ大作戦」の「なお，この録音は自動的に消滅する」という台詞に同期して，ブラウザが強制終了する．
- 「いないいないばあ」を動物キャラクタが行う子供向け映像を webmo に載せたスマートフォン上で再生し，実際にキャラクタが振り向くようにする．
- 子供向けお笑い映像において，笑いがおこるタイミングで写真撮影する[16].

5　議論と展望

5.1　通信遅延

Meshblu を経由した IoT 機器との通信は開発が容易な反面，映像とのより厳密な同期が必要な場合は通信遅延が問題となる．その場合は webmo のように直接 srt.js 上から IoT 機器と通信できる機器を用いて遅延を減少させる以外に現状では対策はない．

5.2　脆弱性

提案システムは surl パラメータを用いて映像と連動した任意の Javascript を実行可能であるため，悪意のあるスクリプトに対し本質的に脆弱である．それに対し，以下の 2 つの対策を講じている．

- 映像の投稿者による公式の srt.js 以外が surl パラメータを用いて実行される場合には警告が表示される．
- eval()に代わるより安全度の高い実行環境である evel()が選択可能である．

今後より安全性を高めるには，ポータルサイトを

公開し，そこにアカウント登録しないと srt.js を実行できないようにし，さらにそれぞれの映像コンテンツに関し不特定多数ユーザからのコメントや評価を記録することで悪質なコンテンツを集合知的に排除する仕組みなどが有効であろう．

5.3 開発支援

現状では提案システムによるコンテンツ開発は，プレーンテキストによる Javascript の編集のみによって行われる．これは開発の自由度が高い反面，プログラム初心者には参入障壁が高い可能性がある．現在，ハッカソンでの API 提供等を通じて srt.js を活用してもらい，言語仕様やプログラミング体験についてのフィードバックを蓄積中である．今後中村ら[10], 松田ら[11]および Kato ら[8]のようにブラウザ上でグラフィカルな開発を行うことができ，内部で srt.js に変換されるミドルウェアを整備すれば，よりライトなユーザにも受け入れられるであろう．なお, 現状でも YouTube サイト上での字幕エディタや Aegisub などのオープンソース字幕エディタを用いれば，映像を見ながらタイムライン上にスクリプトを配置・編集することは可能である．

5.4 映像分析による srt.js 自動生成

映像を入力とし，映像認識技術や音声認識技術を用いることによってコンテンツを分析し，srt.js を自動生成するようなフィルタープログラムを開発することは興味深い研究対象である．例えば著者らが過去に開発した高速動画鑑賞システム CinemaGazer[17][18]の発話認識エンジンを用いて，映像中の発話区間のみ動画の再生速度を調整する「CinemaGazer 化スクリプト」などが自明に実現できる．さらに音声中の悲鳴やグロテスクな映像を認識できるなら，そのシーンのみを無音化したり非表示にしたりする「ホラー対策フィルタ」が実現可能であろう．IoT と連携する例では，映像中のシーンのムードを認識することで，映像鑑賞中のユーザの部屋の照明環境について Hue などを用いて自動調整することなども可能だろう．

謝辞

本研究は JSPS 科研費 JP15H02735, JP16H02867 の助成を受けたものです．

参考文献

[1] 栗原一貴. 放送化の時代のフレセンテーション支援システム. 情報処理, Vol.56, No.5, pp.465-471, 2015.

[2] 栗原一貴, 諸美咲. Openness-adjustable Headset : 開放度を調整可能なヘッドセットの開発, WISS'15 予稿集, pp.197-198, 2015.

[3] K. Kurihara, D. Vronay, and T. Igarashi. Flexible Timeline User Interface Using Constraints, *In Proc. of ACM SIGCHI'05*, pp.1581-1584, 2005.

[4] 渡邊恵太. 融けるデザイン. ビー・エヌ・エヌ新社, 2015.

[5] 貞子 3D2 スマ 4 D. http://www.sadako3d.jp/suma4d/.（2016/8/22 確認）

[6] +1D. http://www.plus1d.jp/.（2016/8/22 確認）

[7] 後藤真孝, 吉井和佳, 藤原弘将, Matthias Mauch, 中野倫靖: Songle: ユーザが誤り訂正により貢献可能な能動的音楽鑑賞サービス, インタラクション'12, pp.1-8, 2012.

[8] J. Kato et al. TextAlive: Integrated Design Environment for Kinetic Typography. *In Proc. of ACM SIGCHI'15*, pp.3403-3412, 2015.

[9] Stepmania Gimmic. https://hackmd.io/s/Hk6p0vDY.（2016/8/22 確認）

[10] 中村聡史, 石川直樹, 渡邊恵太. 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. WISS'13 予稿集, pp.19-24, 2013.

[11] 松田滉平, 中村聡史. 動画に対する音響的装飾の分析と視覚的装飾を可能とする手法の提案. エンタテインメントコンピューティングシンポジウム 2015 論文集, pp.458-464, 2015.

[12] 加藤淳, 後藤 真孝. IoT アプリケーションのソフトウェア・ハードウェアを単一コードベースで開発できる統合開発環境 f3.js. インタラクション'16 予稿集, pp.132-139, 2016.

[13] 馬場匠見, 橋本翔, 増井俊之. Babascript: 人とコンピュータの協調による処理実行環境. WISS'14 予稿集, pp.109-114, 2014.

[14] 橋本翔, 増井俊之. GoldFish: JavaScript と Android NFC による実世界 GUI フレームワーク. インタラクション 2012 論文集, pp.867-870, 2012.

[15] K. Kurihara. Toolification of Games: Achieving Non-game Purposes in the Redundant Spaces of Existing Games. *In Proc. of ACE'15*, pp.31:1-31:5, 2015.

[16] K. Tsukada and M. Oki. EyeCatcher: a digital camera for capturing a variety of natural looking facial expressions in daily snapshots. *In Proc. of Pervasive'10*, pp.112-129, 2010.

[17] K. Kurihara. CinemaGazer: a System for Watching Videos at Very High Speed. *In Proc. of AVI'12*, pp.108-115, 2012.

[18] K. Kurihara et al. Two-level fast-forwarding using speech detection for rapidly perusing video. *In Proc. of AH'14*, pp.19:1-19:2, 2014.

WISS2016 採録判定時のコメント

採録区分：ロング採録

判断理由：

「時刻 X になったら処理 Y を実行する」という、時系列メディアの再生時刻をトリガにしたイベントドリブンなアプリケーション開発を容易にする提案です。字幕用テキストフォーマットの拡張による JavaScript 記述方式、QR コードを用いたデバイス認証を含むサーバ・クライアント型実行環境に加え、ユースケースシナリオ等が評価され、ロング採録となりました。一方で、ターゲットユーザが不明確なことや、セキュリティ面での議論が表層的である点などに改善の余地があります。

レビューサマリ：

全体の構成：
時刻に関するイベントリスナを容易に書けるような Domain-Specific Language およびそれを実現するフレームワークの提案である。既存の字幕用テキストフォーマットを拡張してプログラミング言語を埋め込み可能とした。さらに、そのための実行環境を実装し、さまざまなユースケースシナリオを示した。

評価すべき点：
着眼点、ユースケースシナリオが面白く、説得的である。プロトタイプレベルとはいえ、QR コードを利用した認証や実デバイスを操作するためのシステム構成などよく練られている。

改良に向けたコメント等：

・ターゲットユーザがはっきり示されていない
誰が srt.js を使ったコンテンツを制作できるのかがはっきり書かれていない。関連研究の章で「エンドユーザ・プログラマが気軽に開発できるフレームワーク」と書かれているが、エンドユーザの定義があいまいである。
JavaScript のプログラミングさえできれば、srt.js 形式を学ぶのは難しくはなさそうだが、それだけでよいのか？ 4 章のユースケースシナリオは、JavaScript 初心者の映像作家でも作れるものなのか？ そうでないなら、どう設計してどう構築するのか？ 支援の方法はあるのか？ このような議論が深まっていれば、同様のシステムを構築する際の参考になるだろう。

・セキュリティに関する議論が表層的である
eval() が危険であるから evel() を使う、と書いてあるが、そもそもなぜ危険なのかを示すべきである。それによってはじめて、どう回避できるのかが議論できる。
なお、査読者からは、JavaScript のインタプリタをサンドボックスに入れる、JavaScript ではない簡易スクリプト言語を利用する、といった回避策が指摘されている。

本論文に対する各査読者の詳しいコメントは以下のページを参照のこと：
http://www.wiss.org/WISS2016Proceedings/oral/12.html
