

# ゲームプレイ動画の見どころアノテーションのための インタラクティブ視聴システムの試作

## 漫画的装飾のリアルタイムな付与

福井 拓真<sup>†</sup> 山西 良典<sup>††</sup> 松下 光範<sup>††</sup>

<sup>†</sup> 関西大学大学院総合情報学研究科 〒569-1095 大阪府高槻市霊仙寺町2丁目1-1  
E-mail: †{k137054,ryama,m\_mat}@kansai-u.ac.jp

**あらまし** 本稿では、ゲームプレイ動画の見どころアノテーションのためのインタラクティブ視聴システムを提案する。e-sportsの盛り上がりを受け、ユーザがゲームで遊んでいる様子を映したゲームプレイ動画から見どころシーンをまとめた「まとめ動画」が人気になっている。まとめ動画を作るためには見どころのアノテーションが必要となるが、ゲーム動画に対するアノテーションは複数の要因から非常にコストのかかる作業である。提案システムは漫画的装飾をリアルタイムに付与することを可能にすることで視聴の楽しさを損なうことなく、アノテーションデータを同時に獲得可能にする。

**キーワード** インタラクティブな動画視聴, リアルタイムアノテーション, 感性の外在化, e-sports

## Prototype of interactive gameplay viewing system for annotating highlights

### Comic-like decorations in real-time

Takuma FUKUI<sup>†</sup>, Ryosuke YAMANISHI<sup>††</sup>, and Mitsunori MATSUSHITA<sup>††</sup>

<sup>†</sup> Graduate School of Informatics, Kansai University Ryozenji 2-1-1, Takatsuki, Osaka, 569-1095 Japan  
E-mail: †{k137054,ryama,m\_mat}@kansai-u.ac.jp

**Abstract** This paper proposes a prototype of the interactive gameplay viewing system for annotating highlights. Due to more attention to e-sports, “compilation movies” consisting of highlights in gameplay in which someone plays are becoming popular. Annotation to each clip in gameplay is needed to create a compilation movie. However, the annotating gameplay is tough work for many reasons. The proposed system enables users to decorate comic-like effects to a movie while they view the movie in real-time; it provides both annotation data for analysis and enjoyment for annotators.

**Key words** Interactive watching movies, annotation in real time, externalizing affections, e-sports

## 1. はじめに

e-sportsの盛り上がりを受け、ゲームはプレイするものから見るものへも変化している。ゲームプレイの配信では、プレイヤーが遊んだ結果をそのまま動画としてライブ配信することが多い。ドラマやアニメのように演出や編集により制作者の意図がデザインされたものとは異なり、インタラクティブ性の高いメディアであるゲームを楽しむ様子は視聴者に作り込みのない感動や驚きを提供している。

ライブ配信されたゲームプレイを編集し、映像作品として楽しむ動きもある。例えば、ゲームプレイの動画から見どころと

なるシーンをまとめた「まとめ動画」は、動画コンテンツとして人気となっている。まとめ動画の制作では、視聴者が面白いと感じるシーンである「見どころの把握」が必要となる。まとめ動画の元となるゲーム動画は、1本の動画再生時間が数時間以上に及ぶことも珍しくない。まとめ動画の制作者は長時間動画を視聴し、視聴者がどのシーンを見どころだと感じるかを推測し見どころを把握する必要がある。動画編集のためには見どころはいつ起きたのかだけではなく、感嘆するようなシーンか、それとも笑えるようなシーンかといったどのようなタイプの見どころであるかも把握しなければならない。さらに、まとめ動画の視聴は、元となる動画の全体を大まかに把握することが目

的のダイジェスト動画とは異なり、見どころ（ハイライト）を短い時間で確認することが主要な目的となっている。動画時間を長くしすぎないためには、発見した見どころからどのシーンを使用するかという取舍選択を求められることになる。プレイヤーごとにカメラが存在するマルチカメラのゲームの場合、1つの見どころに対し、その場に居合わせたプレイヤーの数だけ視点（i.e., ゲーム画面）が存在する。そのため、クオリティの高いまとめ動画を制作するにはそれぞれのプレイヤーの視点を確認し、どのプレイヤーの視点を動画に採用するのかを考えなければならない。

### 1.1 ゲーム動画へのアプローチ

ゲーム動画は、視聴者とプレイヤーそれぞれの角度から研究されている。視聴者に向けては視聴の楽しさ向上、プレイヤーにとってはプレイのスキル向上がそれぞれ目的とされている。

梶浦ら [1] は、ゲームに音楽を付与することによって「楽しさ」にどのような影響を及ぼすのかプレイヤー、ゲーム動画視聴者それぞれの立場から検証を行った。また、Yun ら [2] はゲーム動画からキャラクターの動きや視点移動、攻撃シーンを認識し、それぞれの動きに対応したモーションチェアの動きを自動生成することで臨場感を演出し、ゲームプレイ観戦体験の向上を目指した。しかしこれらはゲーム動画視聴体験をより面白くするための研究であり、動画視聴に生じる課題を解決するための研究ではない。

ゲーム動画の視聴で生じる問題に取り組んだ研究としては、佐藤ら [3] や齋藤ら [4] の研究が挙げられる。ゲームプレイの振り返りはプレイした時間と同じだけの長さの動画を、どのような行動をとっていたか、どのような状況だったか考えながら集中して視聴する必要がある。これらの研究では、ゲームプレイの振り返りを支援するインターフェースを提案しており、プレイヤーが意味のある行動を取っていたかや、どちらのプレイヤーが有利だったのかを提示することでゲーム動画を視聴する負担を軽減している。しかし、これらはゲームプレイヤーが振り返りのために動画を視聴する際の問題に取り組んでおり、動画制作者がゲーム動画から見どころを把握し選択する困難さを解決することを意図したものではない。

### 1.2 本稿の貢献

見どころを把握する困難さを解決する手段として、ゲーム動画からの見どころ検出を考える。見どころを検出するためには、ゲーム動画の中で「いつ」「どのような」ことが起きると「見どころ」となるか、という3種類の要素を含んだアノテーションデータが必要となる。しかし、ゲーム動画に対してのアノテーションは、以下に挙げる複数の要因から人的・時間的コストが高い。

(1) ゲーム動画は、短時間でプレイヤーの行動や場の状況が変化するため、頻繁にどのようなことが起きたか記録する必要があり手間がかかる

(2) アップデートによってゲーム内に新たな要素の追加が行われると、それらの要素を反映させるために再度データを収集する必要がある

(3) 見どころについてのアノテーションデータが必要となるが、「面白い」という考えに個人の感性が関わってくるため、個人特有の感覚でないことを保証するには大人数でのアノテーションデータを収集しなければならない

(4) ゲームの登場から時間が経過するにつれ、動画に対する視聴者の見方が変化し、どのようなシーンを見どころと考えるかという基準が変わっていく可能性が考えられる

本稿では、ゲーム動画の見どころアノテーションのためのインタラクティブ視聴システムを提案する。多くの人に、継続的かつ簡単にゲーム動画に対してのアノテーションを行わせるシステムを構築することで上記の問題を解決できると考えた。提案システムでは、漫画で使用されるようなオノマトペや集中線といった装飾（以下、漫画的装飾）を、キー入力というワンアクションでリアルタイムに動画へ付与可能にすることで、いつ・どのような見どころが発生したかを記録できるようにする。オノマトペは、通常の言葉では表現しきれない感覚的な部分も含んでおり [5]、漫画的装飾を使用することにより視聴者は動画に対して感じた言語化しにくい感性も外在化することができる。視聴者は、動画に装飾をつけていくことでオリジナルのデコレーションがなされた自分だけの動画を制作し、自身がどのように装飾を施したか他ユーザに共有することができる。このシステムを利用することで少ないユーザの負担で、ゲーム動画に対しての実験環境下ではない客観的な見どころアノテーションデータを収集することができる。

## 2. 関連研究

映像から肝要となるシーンを探し出しまとめる映像要約の研究において、要約シーンを検出するための手法は数多く考案されている。最も一般的な手法として、動画に含まれる映像や音声を用いるものが挙げられる。祖父江ら [6] や福里ら [7] は、映像の特徴量を利用し、参考元となる漫画やレシピ画像と類似したシーンを検出することで要約を行った。日高ら [8] は、音声の韻律の変化から強調された音声を検出し、その音声が含まれる映像を抽出することで映像の要約を行った。また、望月ら [9] や長崎らの研究 [10] では、要約映像に使われやすいシーンの画像特徴量と音声認識を組み合わせることでイベントの推定や要約映像の生成を行なっている。このように映像や音声の利用は映像要約に有効な手法であるが、ゲーム動画に対して用いるにはゲームごとに特徴量が大きく変わることや、同一のゲームでもアップデートや個人の設定で映像の特徴量に変化が生まれることから必ずしも適しているとは言い難い。また、映像や音声のみではシーンに含まれる情緒的な意味を分析することは難しいという指摘もなされており [11]、面白さという感情的な要素が重要となる見どころシーンの検出には向いていないと考えられる。

映像要約に視聴者の視点を取り入れるための試みとしては、ユーザが動画に対して書き込んだコメントを利用する手法がある。小林ら [12] は、野球の試合に対する視聴者の実況コメントを収集することで要約に必要なシーンを検出した。齋藤ら [13]



図1 提案システムの実行画面。画面は「VALORANT」(注1)より。

は、紹介動画に必要なシーンに書き込まれるコメントの特徴を分析することで映像の要約を行なっている。これらの研究のように、コメントを用いることで視聴者がどのように感じたか、ということを要約に反映することができるが、ゲームタイトルごとに独自の用語が出現しやすいゲーム動画には適していないと考えられる。また、コメントは同じ内容でもユーザによって文章に違いがあることも多く、データ整形の手間が必要となってしまうが、略称や俗称が頻出するゲームプレイ動画ではより多くのコストがかかってしまう。

### 3. 提案システム

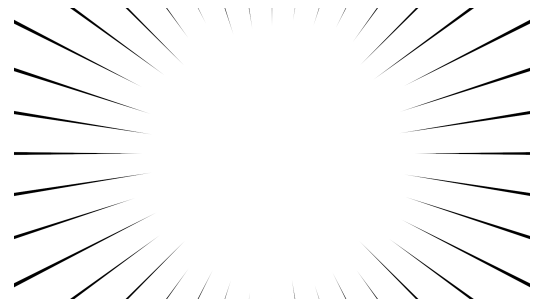
提案システムは、ゲーム動画を見ながらリアルタイムに漫画的装飾を付与可能にし、見どころのアノテーションを記録できる。提案システムの利用により、視聴時の「面白い」「迫力がある」といった感性を漫画的装飾によって強調可能となる。視聴者は、動画に抑揚をつけてデコレーション[14]するような感覚で楽しみながら、自分だけの動画を制作できる。本来単調となるアノテーション作業に対して、遊びの要素を織り込むことによって、積極的にシステムを利用してもらいながらゲーム動画に見どころを把握する手がかりを付与させるしかけとなっている。また、実験環境下以外で自主的にアノテーションを行わせることができるため、リアルな感性を表現したデータが得られる。

#### 3.1 システム構成

提案システムは、動画の再生機能と動画に対して漫画的装飾を施す機能を備えている。図1に、提案システムの実行画面を示す。以降、画面に表示される動画としてシューティングゲーム「VALORANT」(注1)のプレイ動画を使用する。動画の再生機能として画面下部にある Start で再生開始、Pause で一時停止、Stop で再生中断の操作が行える。動画下のシークバーで現在どこまで動画を再生しているか確認でき、シークバー上をクリックすることで再生時刻を変更することができる。動画を視聴しながらキーボードまたはゲームパッドを用いて対応するキーを入力することで、リアルタイムに画面上のランダムな位置に漫画的装飾が付与される。同時に複数の対応するキーを入力することで、複数種類の漫画的装飾を同時に付与することも



(a) オノマトペ



(b) 集中線

図2 提案システムで付与することのできる漫画的装飾。

可能である。付与可能な漫画的装飾は、Baek らの研究[15]を元に、漫画における出現頻度の高い上位10種類のオノマトペである「オ」「ザワ」「パチ」「ワ」「ドキ」「ン」「はっ」「?」「ゴ」「はあ」及び「集中線」(図2)とし、ユーザはこれらの中から使用する装飾を5種類選択する。5種類に絞らせることで覚えやすく[16]、操作が簡易になるだけでなく、自身が使用しなかった装飾を選択した他ユーザの動画を視聴する楽しみも生まれる。図1右下の Key ボタンを押すことで、入力設定画面へ移行し、使用する装飾の選択や選択した装飾を入力するキーを任意のキーやボタンに割り当てることができる。動画の再生が終了すると、いつ・どのような漫画的装飾が付与されたかというアノテーションデータが csv ファイルで出力される。

(注1) : <https://playvalorant.com/ja-jp/>



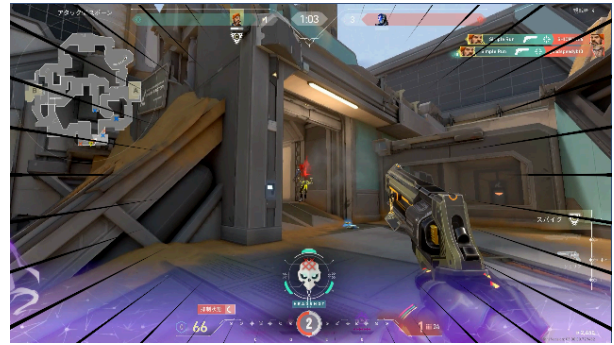
(a) 装飾の施されていないキルシーン



(b) 「オ」「ゴ」「ワ」の装飾が施されたキルシーン



(c) 「ザワ」「はっ」「？」の装飾が施されたキルシーン



(d) 集中線の装飾が施されたキルシーン

図 3 漫画的装飾の有無によるキルシーンの比較 (a と b, c, d) 及び

付与する漫画的装飾の違いによるキルシーンの比較 (b と c と d) .

これらは全て同一のキルシーンである。画面は「VALORANT」<sup>(注1)</sup>より。

### 3.2 提案システムの利点

リアルタイムに動画へ付与される既存コンテンツとして、ニコニコ動画<sup>(注2)</sup>に代表されるコメントが挙げられる。ユーザはコメントを利用することで自身の考えを言語化して表すことができるが、他ユーザに適切に伝わる文章を考え、入力するという手間が必要となる。動画の再生を一時的に停止するケースや、入力が間に合わずコメントを放棄するケースが発生する。提案システムを用いて漫画的装飾をつける際必要となる動作は、1つのキーを入力するだけであるため画面から目を離したり再生を一時停止することなく視聴時の感性を動画に素早く付与することができる。

ニコニコ動画のコメントは動画コンテンツそのものとのインタラクションよりも、動画コンテンツを介した他者とのインタラクションのためのしくみとも捉えられる。ニコニコ動画の最大の特徴は、動画上をコメントが流れていく点であり、コメントを通じて同じ動画を視聴したユーザとコミュニケーションが可能である。動画によっては、動画とは関係のないコメントの付与や、「弹幕」と呼ばれるような動画コンテンツの視聴を著しく阻害するようなコメントが投稿されることもあるが、このようなコメントの付与はニコニコ動画の楽しみ方としてユーザに受け入れられている。このことから、動画視聴を主眼においたインタラクションではないと考える。一方、提案システムで扱う漫画的装飾はメッセージ性が薄く、「どのようなシーンか」を端的・離散的に表現している。他ユーザが付与した装飾を視

聴する際も、漫画的装飾を付与したユーザの存在は意識されにくく動画が強調される。動画視聴時にユーザ間でのコミュニケーションは発生せず、付与される付加情報によって動画の視聴が妨害されることもない。提案システムは動画自体の面白さを損なうことなく享受することに主眼をおいている。提案システムの利用によって、視聴者が動画に対してどのように感じたのかを外在化し、自分自身の感性に向き合うことができる。

咄嗟の入力が可能な機能がある場合、ユーザはコメントよりも頻繁にその機能を利用することが知られている [17]。そのため、コメントを用いるよりも多くのアノテーションデータが収集されることが期待される。コメントの場合、「拍手」を意味する「8888」のようにネットスラングが頻出するため、ネットスラングを知らないユーザには意味が理解できない。一方で、漫画的装飾は漫画の読書経験があれば新規ユーザでも直感的に意味を理解可能である。コメントの場合、同一の意図であっても、異なる単語で記述されることがある。提案システムの漫画的装飾は不変のシンボルであるため、画一的なデータの取得が見込める。

漫画的装飾は、シーンの印象を変化させたり、強調したりする働きがある。例えば、シューティングゲームでは、敵を倒した際のエフェクトはどのシーンでもほぼ同一であり、全てのシーンに対して同じような印象を受けてしまいかねない。そこで、図3のように装飾を施すことで、特定のシーンの印象を他シーンより強調することができる。漫画的装飾の有無だけではなく、図3(b)、図3(c)、図3(d)の違いのように、漫画的装飾

(注2) : <https://www.nicovideo.jp/>

の種類を変化させることでシーンの印象が変わるため、視聴者自身が感じた特定の印象を強調可能である。他ユーザが装飾を施した動画を見ることで、動画の受け取り方の違いを楽しんだり、新たなゲームの観点に気づいたりするゲーム動画視聴の楽しみ方の拡張が期待される。動画への装飾の付与は、一視聴者ではなくコンテンツの提供者としての動画に参加している感覚も与えるであろう。

### 3.3 今後の展望

現在のシステムでは、漫画での出現頻度上位 10 種のオノマトペをライブラリに用意している。漫画は静的なコンテンツである一方で、ゲーム動画は動的なコンテンツである。そのため、オノマトペの漫画で用いられる場合と動画に重畳した場合の効果の違いを検証する必要がある。各オノマトペの使用回数やユーザテストから、動画に適した装飾を分析し、ゲームのプレイ動画へのアノテーションにおいて有用な漫画的装飾の集合を明らかにする。また、シーンの意味と付与された漫画的装飾の関係性を分析すると共に、それぞれの漫画的装飾が提示されるべき画面上の場所にも検討する。

ユーザが付与したアノテーションデータは、将来的に様々な用途での利用が期待できる。アノテーションデータから見どころとなるシーンを検出するだけでなく、その見どころの種類(「ゲームプレイが上手い」「面白い」など)の判別も可能になると期待される。さらに、マルチカメラのゲームでは、複数の動画から見どころシーンの比較を行うためのマーカーとしても参照可能であり、シーンの組み合わせ方のヒントとしての利用も期待できる。同一の装飾が付与されたシーンについて、ゲームのタイトルやジャンルを超えて参照することで複数タイトルのゲーム動画を対象とした動画の分析やまとめ動画の作成が考えられる。

## 4. おわりに

本稿では、ゲーム動画に見どころアノテーションを付与するための、漫画的装飾を用いたインタラクティブ動画視聴システムを提案した。提案システムは、視聴した動画への感性を漫画的装飾によって外在化させるとともに、動画の各シーンへの印象をアノテーションしたデータを取得可能にする。

今後は、ユーザテストを実施し、漫画的装飾の使われ方を分析する。また、アノテーションデータの参照による見どころの検出や見どころの分類についても取り組んでいく。

### 文 献

- [1] 梶浦久江, 中山伸一, “ブロック崩しゲームにおけるプレイヤーとゲームを見る人のフロー体験に与える音楽の影響,” デジタルゲーム学研究, vol.4, no.2, pp.13–22, 2010.
- [2] G. Yun, H. Lee, S. Han, and S. Choi, “Improving viewing experiences of first-person shooter gameplays with automatically-generated motion effects,” Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, no.320(14 pages), 2021.
- [3] 佐藤隼介, 梶並知記, “Fps ゲームにおける角待ちに注目した可視化インタフェース,” 日本デジタルゲーム学会年次大会予稿集, vol.12, pp.23–26, 2022.
- [4] 齋藤慎之介, 和三史弥, 梶並知記, “ガンダムアクションゲームを対象とした任意時間帯における優劣を可視化する振り返り支

- 援インタフェース,” 日本デジタルゲーム学会年次大会予稿集, vol.13, pp.29–34, 2023.
- [5] 山田香織, 工口陽平, 田浦俊春, “擬態語を用いた合成的な動きのデザイン方法の提案 (第 1 報 擬態語を見出しとした動きのデータベースの構築),” 日本機械学会論文集, vol.80, no.815, pp.DSM0210–DSM0210, 2014.
- [6] 祖父江亮, 山下隆義, 平川 翼, 藤吉弘亘, 中澤 満, C. Yeongnam, S. Björn, “レシピ画像と説明文を活用した類似シーン検索による調理動画の要約,” 精密工学会誌, vol.86, no.12, pp.1026–1033, 2020.
- [7] 福里 司, 平井辰典, 大矢隼士, 森島繁生, “アニメ作品におけるキーフレーム自動抽出に基づく映像要約手法の提案,” 画像電子学会誌, vol.42, no.4, pp.448–456, 2013.
- [8] 日高浩太, 竹内順二, 松浦宣彦, 茨木 久, 中島信弥, “音声の強調情報を利用したビデオコンテンツ短縮視聴方法の検討,” 画像電子学会誌, vol.34, no.5, pp.505–511, 2005.
- [9] 望月貴裕, 河合吉彦, 藤森真綱, 前澤桃子, 遠藤 伶, 浅見雄一郎, “ニュース要約映像作成支援システムの試作,” 映像情報メディア学会誌, vol.77, no.2, pp.262–271, 2023.
- [10] 長崎好輝, 林 昌希, 金子直史, 青木義満, “動画内の音と映像によるイベント推定タスクにおける時間方向クロスモーダルアテンションの導入,” 精密工学会誌, vol.88, no.3, pp.263–268, 2022.
- [11] L.-X. Tang, T. Mei, and X.-S. Hua, “Near-lossless video summarization,” Proceedings of the 17th ACM International Conference on Multimedia, p.351–360, 2009.
- [12] 小林尊志, 野田雅文, 出口大輔, 高橋友和, 井手一郎, 村瀬 洋, “Twitter の実況書き込みを利用したスポーツ映像の要約,” 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, vol.110, no.457, pp.165–169, 2011.
- [13] 齊藤義仰, 磯貝佳輝, 村山優子, “視聴者コメントを用いた動画検索支援のための紹介動画作成手法の提案,” 情報処理学会論文誌コンシューマ・デバイス&システム, vol.2, no.1, pp.74–81, 2012.
- [14] 鳥塚裕喜, 萩原将文, “効果素材を付加した対話型漫画生成支援システム,” 知能と情報, vol.32, no.6, pp.956–963, 2020.
- [15] J. Baek, Y. Matsui, and K. Aizawa, “Coo: Comic onomatopoeia dataset for recognizing arbitrary or truncated texts,” Proceedings of the 17th European Conference on Computer Vision, p.267–283, 2022.
- [16] C. Nelson, “The magical number 4 in short-term memory: A reconsideration of mental storage capacity,” Behavioral and Brain Sciences, vol.24, no.1, pp.87–114, 02 2001.
- [17] 香川健太郎, 伊藤淳子, 宗森 純, “動画共有システムに与える直感的絵文字コメント投稿機能と感情共有機能の効果,” 情報処理学会論文誌, vol.51, no.3, pp.770–783, 03 2010.