

# YoutubeLive のチャット解析による盛り上がりの検知及び VR による演出

2010390 須山瑛樹 成見研究室

## 1 はじめに

近年 esports 分野の市場規模は拡大傾向にあり、2024 年 7 月には、賞金総額約 6,000 万ドルの"ESPORTS WORLD CUP"も開催され、注目を集めた。このような規模の大きな大会を現地に赴き観戦することは、esports の楽しみ方の一つである。しかし、世界大会や規模の大きな各地域大会は現状海外での開催が主であり、大会の現地観戦には金銭的・時間的なコストの問題も存在する。

誰もが手軽に観戦するだけであれば、多くの esports 大会は Youtube 等の配信サイトで配信しているためこれらを視聴すればよい。しかし、現地では観客の歓声や拍手、舞台演出で会場全体が盛り上がりを見せていたとしても配信映像には映らないため盛り上がりが見えづらく、臨場感のある観戦体験ができるとはいえない。

## 2 研究目的

本研究では、VR を用いて安価で手軽に臨場感のある esports の試合観戦体験を可能にするシステムの構築を目指す。リアルタイムでの試合観戦を可能にするための土台作りとして、試合のアーカイブの映像を用いた実装を行う。Youtube のライブ配信のチャット機能から試合の盛り上がりを検出し、VR 空間の疑似会場において適切なタイミングで演出を行うことで、現地の会場にいるような体験を可能にする。

## 3 既存研究

### 3.1 チャット量を比較した盛り上がり箇所の判定

高橋は配信アーカイブを 1 分ごとのチャット数を比較し、チャットの多い上位 3 つの区画のクリップを作成した [1]。本研究ではより細かい区間でのチャット数の比較に加え、盛り上がっているときのチャットに含まれるキーワードを解析し、それらのキーワードの出現頻度でも盛り上がりを判定する。

### 3.2 テキストチャット解析によるクリップ生成

上村は既存のハイライト動画から元動画の盛り上がったシーンを把握し、そのシーンで入力されたチャットを解析することで機械学習モデルを作成した [2]。機械学習モデルを用いてクリップを新たに生成し、既存のハイライト動画と比較していた。本研究ではクリップの作成ではなく、視聴に演出を加えることを目的とする。

### 3.3 テキストチャット解析による感情分析

Thakre らは Stanford Sentiment Treebank v2 に基づいて、入力されたチャットを形態素解析し、それぞれの形態素について 5 ポイントの感情値 (非常に肯定的~非常に否定的) を割り当てチャットの感情を読み取った [3]。また、絵文字は実際の文章に比べて皮肉などの暗喩的な意味合いが含まれることが少なく、投稿者のストレートな感情を表現していることが多いと考えられるため、感情を読み取るうえでのヒントとし、これらの情報から視聴者のリアルタイムの感情を分析し、配信者にフィードバックした。本研究では、感情の肯定的、否定的だけでなく、4 つの感情に分け、その大きさを検知することで、より細かい感情の分析を行った。

## 4 提案システム

### 4.1 システム概要

システム概要を図 1, 図 2 に示す。YouTube のライブ配信のテキストチャットデータを解析して観戦者の心情を把握するのが前半部分である。後半部分では、VRChat にアップロードした疑似会場に観戦者の感情に合った演出を施して臨場感のある試合体験を可能にする。

配信のアーカイブからテキストチャットデータを取得し、チャット数の比較から盛り上がったタイミングを検出する。盛り上がったタイミングのチャットを形態素解析し、視聴者の心情の種類とそのタイミングでよく入力されているチャットの内容を SVM による機械学習で紐づけて、類似のチャットが入力された場合はその心情に合った演出のタイミングと判断し時刻を記録する。また、GG(Good Game の略語) というスラングが含まれたチャット (以下、「GG チャット」と称する) の頻度から試合の終了時刻を、絵文字から適切な拍手のタイミングも記録する。記録されたデータをもとに、VRChat にアップロードした疑似会場に演出を加え、MetaQuest2 に投影する。

### 4.2 対象とした esports

テキストチャットを解析するうえで、チャットの量が多いほうがより会場の雰囲気と正確に把握することができるため、本研究の対象の esports としてプレイヤー数が多く人気の高いシューティングゲーム「VALORANT」を選定する。また、同様の理由から、大会配信の視聴者数が多い YouTube チャンネルである

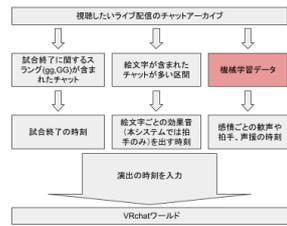
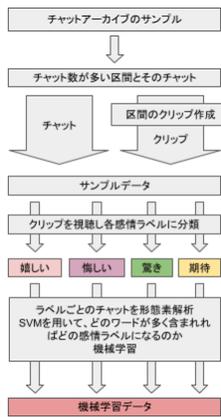


図 1 システム概要 (チャット取得～機械学習データ作成)

図 2 システム概要 (~VR-Chat への出力)

「VALORNT // JAPAN」を対象のチャンネルとする。

### 4.3 試合終了時間の予測のためのパラメータ決定

GG チャットから試合終了時刻を予想するため、パラメータを決定する実験を行った。GG チャット同士の入力された間隔が何秒以内であれば連続して入力されたと判断するかを表す間隔  $X(s)$  と、何回連続で入力されれば試合終了と判断するかを表す連続回数  $Y(回)$  のパラメータを、それぞれ  $(X, Y) = (1, 5), (1, 10), (1, 15), (3, 10), (3, 15)$  と変化させて実験を行った。また、これらの結果から  $(5, 5)$  についても追加で実験を行った。41 試合の実際の試合終了時刻と各パラメータによって予測された試合終了時刻との誤差の比較は表 1 のとおりである。未検知の試合数が最も少なく、誤差 5 秒未満の試合数が最も多いため、本システムにおいてパラメータは  $(5, 5)$  を採用した。

## 5 評価実験

### 5.1 検出した盛り上がりシーンの評価

本システムにより「嬉しい」、「悔しい」、「驚き」、「期待」の 4 種類のラベル付けがなされた 11 本のクリップを、9 名の被験者にはラベルを隠して視聴してもらい、各クリップのラベルの予想とその感情の大きさの度合いを質問した。感情の大きさはクリップ区間の 10 秒ごとのチャット量 (以下、「チャットスコア」と称する) との関係性を調査

表 1 試合終了予測実験の結果

パラメータ(X,Y)	誤差			未検知
	5s未満	5s以上 10s未満	10s以上	
(1,5)	22	11	2	6
(1,10)	19	6	1	15
(1,15)	19	3	0	19
(3,10)	24	4	2	11
(5,10)	26	5	3	7
(5,5)	28	6	4	3

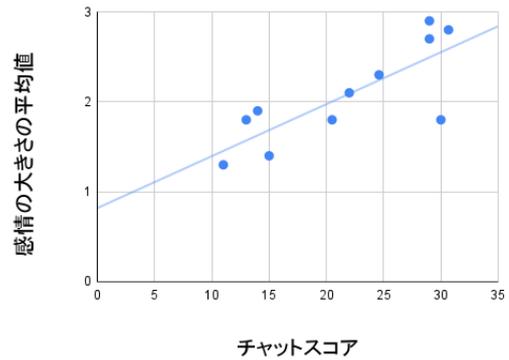


図 3 チャットスコアと感情の大きさの関係

した。「期待」に関しては一致率が 100 %、「嬉しい」「驚き」に関しても約 60 % で一致していたが、「悔しい」に関しては正確に検出できなかった。また、チャットスコアと感情の大きさには図 3 のとおり正の相関関係があり、相関係数は 0.79 であった。

### 5.2 システムの使用感の調査

実際に VRChat の疑似会場にてアーカイブを 1 試合視聴してもらいその感想を聞いた。肯定的な意見として、「手軽に巨大なスクリーンで視聴できる」ことや、「試合終了と歓声のタイミングの一致、雑踏や歓声により大勢の存在を感じられる」というものがあつたが、否定的な意見として、「好プレイの時の歓声がずれていた」、「動画の音声がかき消されていた」、「会場に配置されたマネキンが動かず違和感があつた」などが挙げられた。

## 6 まとめ

esports の配信アーカイブのチャットを分析し、臨場感のある試合観戦体験をするためのシステムを構築した。「期待」「嬉しい」といったプラスの感情の検出は被験者とおおむね一致したが、「悔しい」などのマイナスな感情の検出ではうまくいかなかった。リアルタイム視聴への対応方法の検討や、マイナスな感情の検出などが今後の課題点としてあげられる。

## 参考文献

- [1] 高橋慶至, 青森大学 ソフトウェア情報学部 ソフトウェア情報学科 卒業研究, YouTubeLive のチャットデータを用いた盛り上がり箇所の判定, 2022 年
- [2] 上村岳人, 名古屋工業大学大学院 工学研究科 工学専攻 情報工学系プログラム 修士論文, インターネット配信動画におけるテキストチャット解析によるハイライト動画生成, 2023
- [3] Yash Thakre, AhivamSudan, Sujal Agrawal, Aveg Ganorkar, Mining Comments and Sentiments in YouTube Live Chat Data, Shubhangi Tripude, 2024 4th International Conference on Intelligent Technologies(CONIT)