

## スポーツハイライト映像作成のための Twitter 熱狂度に基づいたイベント検出

富田 大志<sup>†</sup> 道満 恵介<sup>††</sup> 井手 一郎<sup>†††</sup> 出口 大輔<sup>††††</sup> 村瀬 洋<sup>†††</sup>

<sup>†</sup> 名古屋大学 工学部 〒464-8603 愛知県名古屋市千種区不老町

<sup>††</sup> 中京大学 情報理工学部 〒470-0393 愛知県豊田市海津町床立 101

<sup>†††</sup> 名古屋大学 大学院情報科学研究科 〒464-8601 愛知県名古屋市千種区不老町

<sup>††††</sup> 名古屋大学 情報連携統括本部 〒464-8601 愛知県名古屋市千種区不老町

E-mail: <sup>†</sup>tomitat@murase.m.is.nagoya-u.ac.jp, <sup>††</sup>kdoman@sist.chukyo-u.ac.jp,

<sup>†††</sup>{ide,murase}@is.nagoya-u.ac.jp, <sup>††††</sup>ddeguchi@nagoya-u.jp

**あらまし** 本報告では、マイクロブログサービス Twitter への投稿を用いて算出した“Twitter 熱狂度”に基づいて、スポーツハイライト映像を作成するためのイベント検出手法を提案する。Twitter を用いた従来のスポーツ映像要約手法は、ツイートの言語やスポーツ種目に依存するほか、イベントではない場面を誤検出するという問題があった。一方、提案手法では、言語やスポーツ種目に依存しない文字列特徴を用いて“Twitter 熱狂度”を算出し、試合中のイベントを検出する。提案手法を実際のスポーツ試合に適用し、検出されたイベントと実際に放送されたハイライト映像に含まれるイベントを比較した結果、提案手法の有効性を確認した。

**キーワード** Twitter, スポーツ, ハイライト映像, イベント検出

## Event Detection based on Twitter Enthusiasm Degree for Making a Sports Highlight Video

Taishi TOMITA<sup>†</sup>, Keisuke DOMAN<sup>††</sup>, Ichiro IDE<sup>†††</sup>,

Daisuke DEGUCHI<sup>††††</sup>, and Hiroshi MURASE<sup>†††</sup>

<sup>†</sup> Faculty of Engineering, Nagoya University Japan

<sup>††</sup> School of Information Science & Technology, Chukyo University Japan

<sup>†††</sup> Graduate School of Information Science, Nagoya University Japan

<sup>††††</sup> Information and Communications Headquarters, Nagoya University Japan

E-mail: <sup>†</sup>tomitat@murase.m.is.nagoya-u.ac.jp, <sup>††</sup>kdoman@sist.chukyo-u.ac.jp,

<sup>†††</sup>{ide,murase}@is.nagoya-u.ac.jp, <sup>††††</sup>ddeguchi@nagoya-u.jp

**Abstract** In this report, to create a sports highlight video, we propose a method for detecting events based on “Twitter enthusiasm degree” calculated from posts to Twitter. The conventional sports video summarization methods that make use of Twitter, not only depend on both languages and sports, but also often detect non-event scenes. Meanwhile, the proposed method detects sports events using a “Twitter enthusiasm degree” calculated from several kinds of string features independent from languages and sports. We applied the proposed method to actual sports games, and compared the detected events with the events in actually broadcasted highlight videos. As a result, we confirmed the effectiveness of the proposed method.

**Key words** Twitter, sports, highlight video, event detection

### 1. はじめに

近年、放送映像の処理技術に対する需要が高まっており、特

に自動要約技術が注目を集めている。なかでもスポーツ映像は多くの人が視聴するが、種目によっては試合が長時間に及ぶことがあるため、要約技術への期待が高い。スポーツ映像を要約

表 1 ロンドンオリンピック 女子サッカー決勝（日本対米国）戦において投稿されたツイート

時刻	タイムゾーン	ツイート内容
19:55:07	UTC-6	GOOOOOOOOOOOOOOOOOAA AAAAAAAAAALLLLLLLLLL!!!
19:55:08	UTC-8	Go #TeamUSA!
19:55:10	UTC+2	US. leads Japan 1-0 at halftime in gold medal women's soccer match.
19:55:12	UTC+9	うああああああああああ 切り替えの遅さが目立つなあ
19:55:15	UTC-5	GOAL #TEAMUSA !!!!!!!!!!! Up 2:0 in the 55th Minutes!!!

する際には、視聴者の視点を考慮することで、より臨場感や興奮を伝えることができると考えられる。以降、このように臨場感や興奮を視聴者に伝えられる要約映像のことをハイライト映像と呼ぶ。

ハイライト映像の作成に関して、試合会場の観客の歓声の大きさをを用いる手法 [1] や、番組実況チャットの書き込みを用いる手法 [2]、番組中の音声や動き特徴を用いる手法 [3] が提案されている。しかしこれらの手法は、会場にいる観客や実況チャットに書き込むユーザからの情報を用いるため、幅広く視聴者視点の情報を得られているとはいえない。ハイライト映像の作成には、広いユーザ層、多くのユーザが利用するサービスを情報源とする必要があるため本研究では情報源として、マイクロブログサービス Twitter<sup>(注1)</sup> に注目した。

Twitter は、2006 年に米国でサービスを開始して以来、ユーザを増やし続けている。2012 年 3 月時点のアクティブユーザ数は全世界で 1 億 4,000 万人を超えており [4]、そのうち 70% 以上は米国以外から利用している [5]。また、最近では即時性から、Twitter とスポーツの相性の良さが認知されつつある。2012 年に開催されたロンドンオリンピックにおいては、大会期間中に投稿されたオリンピックに関するツイート（Twitter への投稿）数は 1 億件以上になった [6]。オリンピックのような世界規模の大会でなくとも、アメリカンフットボールのプロリーグである NFL やサッカーの欧州選手権である UEFA EURO など、ある程度知名度が高い大会であれば、オリンピックと同じようにその大会に関するツイートが多く投稿される [7]。実際にオリンピック期間中に投稿されたツイートの例を表 1 に示す。この例は、サッカーの試合中にゴールが決まった瞬間のツイートである。得点に対する歓喜や落胆といった、ゴールが決まったことに対するツイートが様々な地域から様々な国の言語で投稿されている。このことから、その時に視聴しているスポーツの状況を Twitter へ投稿するユーザは多数存在するといえる。

このような Twitter とスポーツの関係に着目し、Twitter へ投稿されたツイートをを用いてスポーツ映像を要約する手法 [8] [9] [10] が提案されている。これらの手法は、ツイートに含まれる語句を解析することでイベントを判別するため高いイベントの認識率を得ることに成功している。しかし、語句を基にイベントを判別する際は、スポーツ種目ごとにイベントに関

する語彙を記述した辞書が必要なことに加えて、分析対象のツイートの言語に制約がかかるという問題がある。これらの手法は、いずれもツイート数の増減を基にしてイベントの発生を検出している。これに関して、Shamma ら [11] は、米国の大統領選挙中の党首討論番組における議論の流れを Twitter 上でのユーザ間の議論を基に解析し、映像に対するアノテーションを試みた。その結果、ツイート数が討論の終了時に増大する傾向が確認され、放送された番組中の盛り上がりのみでツイート数が増加するのではないことを示した。このことから、スポーツ試合の放送映像においてもイベント以外の場面でツイート数が増大する可能性があり、ツイート数を基にした検出手法では、誤検出が起きてしまうと考えられる。そこで我々は、イベント以外でのツイート数の増大を考慮し、ツイート数に加えツイート中の文字列特徴を用いた得点イベントの検出手法を提案した [12]。この手法では、得点のみを検出対象としていたが、視聴者が重要だと感じるイベントは得点イベント以外にも存在する。

そこで本報告では、ツイート中の興奮を示す特徴や、他のユーザのツイートを引用したツイート（リツイート）の数など、ツイート自身に含まれる特徴を考慮して、スポーツ試合のハイライト映像を作成するためにイベントを検出する手法を提案する。以下に、本手法のアプローチと適用範囲を述べる。

#### a) アプローチ

本手法では、Twitter ツイートの文字列特徴を基に算出した Twitter 上におけるユーザの盛り上がりのことを「Twitter 熱狂度」と呼び、この値に基づいてハイライト映像に含むべきイベントを検出する。ここでは、ハイライト映像に含むべきイベントのなかでも視聴者の興奮が高まるようなイベントを検出する。視聴者の興奮が高まるようなイベントとは、得点イベントや得点に結びついたイベント、得点になりかけてならなかったイベントなどのことである。これらのイベントは、他のイベントに比べて、試合の勝敗により密接に関係するイベントであるため、試合の流れを理解するうえで重要である。なお、算出した熱狂度に対してしきい値処理を行うことで、得点イベントのような、視聴者をより興奮させるイベントを分離することもできる。

#### b) 適用範囲

本研究では、スポーツ種目の中でも特にサッカーと野球の試合を対象とした。この 2 種目のスポーツは、得点に至る過程や試合の構造などの性質が異なるため、これら 2 種目に関して提案手法を適用し、有効性を確認することができれば、スポーツ種目に依存しない手法であることが示せると考えた。また、日本国内においては、サッカーと野球は他のスポーツに比べ人気大幅に高い [13] ため、テレビ中継される機会も多い。世界的にも、国際連合加盟国よりも国際サッカー連盟加盟国の方が多い [14] ことが示すように、サッカーは人気が高いスポーツである。従って、この 2 種目に関しては各国の言語で書かれたツイートが多く集まるため、それらを対象にすることで、言語に依存しない手法であることを示せると考えた。

以降、2. で Twitter について述べ、3. で提案手法を詳述す

(注1) : Twitter, <https://twitter.com/>.

表 2 リツイートとハッシュタグの例 (ユーザ名は XXX に置換)

日時	ツイート内容
2012/9/6 11:59:12	RT @XXX: 駒野のクロス最高!! #ntv #daihyo #ツイート観戦 #jfa
2012/9/6 12:07:24	RT @XXX: 清武の今日のパフォーマンスは素晴らしいね。 #jfa #daihyo

る。その後、4. で提案手法の評価実験の結果と考察を述べ、最後に 5. でまとめる。

## 2. Twitter について

Twitter は、投稿文字数が 140 文字までに制限されたマイクロブログサービスである。Twitter への投稿は「ツイート」と呼ばれ、近況やその時感じたことなどを書き込むが、他者との交流のために用意された機能や慣習もある。その 1 つである「リツイート」は、他者のツイートを引用して自分のツイートを投稿する機能である。その例を表 1(c) に示す。このツイートは、XXX という名のユーザのツイートを引用して他のユーザが投稿したものである。リツイートを行う際には、引用するツイートの最初に「RT」を付して投稿する。なお、特定のユーザを話題にしたツイートや、他のユーザに宛てたツイートを投稿する場合は、そのユーザ名の先頭に「@」を付す。

他にも、ある特定の話題に関するツイートであることを明示するために、ツイートに「ハッシュタグ」を付して投稿することがある。表 1(c) に示したツイートの末尾に、「#」に続けて記されているのがハッシュタグである。このようにツイートすることで、特定の話題に関するツイートのみを検索できるようになり、同じ事柄に興味を持つユーザを見つけたり、その話題に関する他のユーザの反応を調べるのに役立つ。

本研究では、Twitter 社が提供している Twitter Streaming API<sup>(注2)</sup>を用いてスポーツ試合中のツイートを収集する。この API は、検索語を指定することでその検索語が含まれるツイートのみを収集することができるため、本研究では各試合に関連するハッシュタグを検索語として設定することで、その試合に関するツイートを収集した。

## 3. 提案手法

提案手法は、Twitter から収集するツイートとスポーツ試合の映像を入力とし、ツイート中の文字列特徴を用いて Twitter 熱狂度を計算することでハイライト映像に含めるイベント区間を出力する。本手法では、以下の 3 つの特徴を用いる。

- (1) 興奮を示す特徴 1 :  $E$
- (2) 興奮を示す特徴 2 :  $L$
- (3) イベントの区切りを示す特徴 :  $R$

これらの値は窓幅  $w$  ( $=30$ ) 秒ごとに計算し、最大値を 1、最小値を 0 に正規化する。なお、これらの特徴を用いる理由については後述する。

上記の 3 つの特徴を用い、 $w$  秒ごとに Twitter 熱狂度を次式

で算出する。

$$S = E \times L \times (1 - R) \quad (1)$$

ここで、Twitter 熱狂度は最大値を 1、最小値を 0 として正規化する。このように、 $E$  と  $L$  の積に  $R$  を 1 から減じた値を乗じることで、興奮を示す特徴量  $E$ ,  $L$  が大きく、イベントの区切りを示す特徴量  $R$  が小さいほど Twitter 熱狂度  $S$  が大きくなり、ハイライトに含めるイベントとして検出されやすくなる。

以降、実際のサッカーの試合中に投稿されたツイートから得られた結果を基に、各特徴と熱狂度について詳述する。なお、ここで説明に用いる試合は、2012 年 6 月 23 日に行われた UEFA EURO2012 のスペイン対フランス戦である。この試合では、19 分と後半のアディショナルタイム (AT) 1 分の時に得点が入った。後半の AT1 分の得点はペナルティキック (PK) によるものであった。

### 3.1 興奮を示す特徴

得点に関するイベントが発生した瞬間において、試合を観戦しているユーザのなかで興奮や落胆などの感情が発現し、それがツイートの文面に表れる傾向がある。従って、ユーザが興奮状態にあることを検出できれば、得点関連のイベントを検出できるのではないかと考えた。一方で、ツイートは口語体で投稿されることが多いため、語気を荒げる記号として、感嘆符 (“!”) が用いられることが多い。語気を荒げるのは興奮していることを示し、感嘆符の出現数が多いほど、より興奮していると考えた。そこで我々は、感嘆符の出現数を興奮を表す特徴量  $E$  として用いる。図 1(a) に試合中のツイートに含まれる感嘆符の文字数を  $w$  秒ごとに数え、最大値を 1、最小値を 0 に正規化した結果を示す。19 分と後半の AT1 分の得点時に感嘆符が多く現れている。また、試合終了時にも多く現れている。

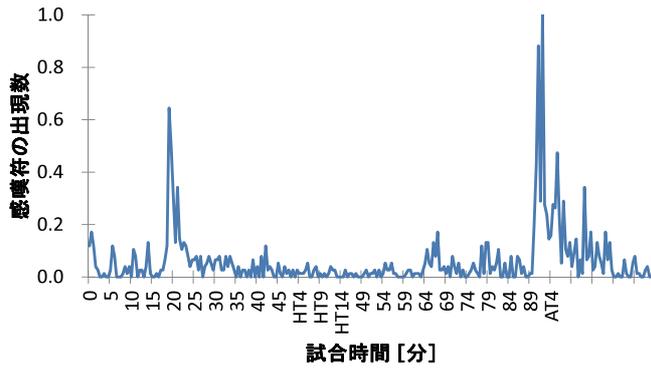
また、“入ったああああああ”、“GOOOOOOAAAAALLLLL!!!”のように同じ文字を繰り返して興奮を表現することも多い (以下、繰り返し文字表現と呼ぶ)。そこで、このような繰り返し文字表現を含むツイート数も興奮を表す特徴量  $L$  として用いる。なお、高橋ら [15] は、同じ文字が 5 回以上出現する繰り返し文字表現を含むツイートをを用いて感情的なツイートの抽出に成功している。そのため、本研究においても 5 回以上の繰り返し文字表現を含むツイートを興奮を表すものと考えた。図 1(b) に繰り返し文字表現を含むツイート数を  $w$  秒ごとに数え、最大値を 1、最小値を 0 に正規化した値を示す。この結果においても、得点時に大きな値が現れている。

以上に基づいて、本研究では興奮を表す特徴として感嘆符の出現数と繰り返し文字を含むツイート数の 2 つの特徴を用いる。

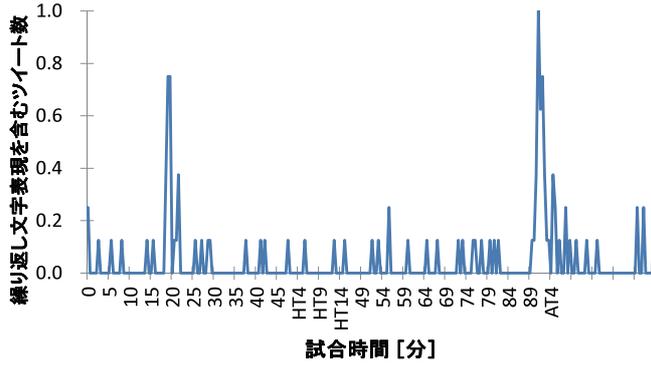
### 3.2 試合の区切りを示す特徴

既存研究 [11] において、イベントの区切りではツイート数が増大することが確認されている。このことから、スポーツの試合の区切りやイベントの後においてもツイート数が増大し、イベント検出に影響を及ぼすと考えられる。このため、試合やイベントの区切りを示す特徴が必要である。ここでは、ツイートの数に着目した。2. で述べたように、リツイートは他者のツイートを引用してツイートを投稿する機能である。リツイート

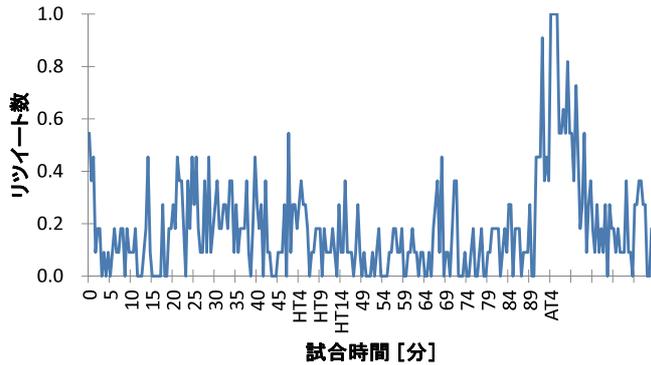
(注2) : Twitter Streaming API, <https://dev.twitter.com/docs/streaming-apis/>.



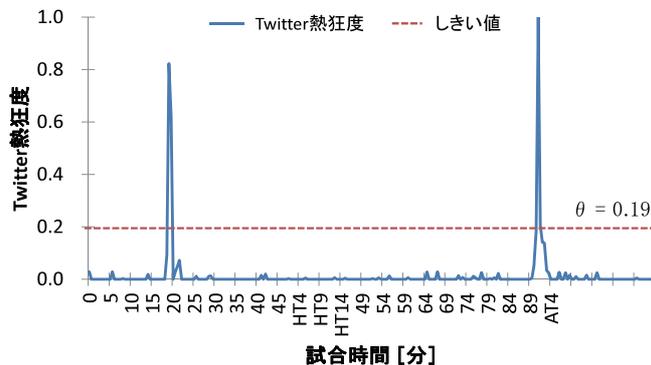
(a) 感嘆符の出現数



(b) 繰り返し文字表現を含むツイート数



(c) リツイート数



(d) Twitter 熱狂度としきい値

図 1 提案手法で用いる特徴量と Twitter 熱狂度の算出例

を行う際には、気に入ったツイートを発見して、先頭に“RT”と付し、場合によっては自身のコメントも付け加えて投稿するため、通常のツイートを投稿する際よりも手間がかかる。このような作業を、まさに得点が入りそうな瞬間に行うとは考えにくく、他者のツイートを引用した議論は、イベントの終了時や

その後に行うと思われる。従って、試合やイベントの区切りとその他の場面を区別する特徴として、リツイート数を用いる。図 1(c) に試合中のリツイート数を  $w$  秒ごとに数え、最大値を 1、最小値を 0 に正規化した結果を示す。得点直後と試合終了時にリツイート数が特に増加している。この観察に基づいて、試合の区切りを示す特徴  $R$  としてリツイート数を用いる。

### 3.3 Twitter 熱狂度としきい値

イベントの検出に用いるしきい値は次式で算出する。

$$\theta = \mu_S + \alpha \times \sigma_S \quad (2)$$

$\mu_S$  は Twitter 熱狂度の 1 試合中の平均、 $\sigma_S$  は標準偏差である。提案手法では、Twitter 熱狂度がこのしきい値を超過した時点をハイライトに含めるイベントとして検出する。このしきい値の値を小さくすると、特に熱狂度が高くなる得点イベントの他に、得点に結びついたイベントや得点が入りかけたものに入らなかったイベントなど、得点イベントほどではないものの、熱狂度が高いイベントを検出できるようになる。

Twitter 熱狂度とそのしきい値を算出した結果を図 1(d) に示す。式 (2) において  $\alpha = 2$  として、この試合の Twitter 熱狂度に対するしきい値を算出した結果、 $\theta = 0.19$  だった。Twitter 熱狂度がしきい値を超えたのは、19 分と後半の AT1 分であり、実際に得点があった時点と一致する。このように、正しくイベントを検出できたため、しきい値の設定は適切だったと考える。

## 4. 実験・考察

提案手法の有効性を評価するための実験について述べる。

### 4.1 方法

評価実験に際して、表 3 に示すスポーツ試合のツイートを用いた。ツイートは、大会ごとに表 4 に示すハッシュタグのいずれかを含むものを収集した。また、しきい値を決める式 (2) において、 $\alpha = 2$  とした場合と  $\alpha = 1$  とした場合を比較することで、検出されるイベントの性質や有効性を確認した。

評価方法として、提案手法によって検出されたイベントと、放送されたテレビ番組のハイライト映像が含む試合中のイベントを比較した。また、参考として我々が以前提案した得点イベント検出手法 [12] で検出された結果も示す。得点イベント検出手法は、以下の 3 つの特徴を用いる。

- $T$ : ツイート数
- $E$ : 感嘆符 (“!”) の出現数
- $R$ : リツイート数

これらの特徴量を最小値 0、最大値 1 になるように正規化し、式 (3) に従って得点イベントらしさ  $G$  を算出する。

$$G = T \times E \times (1 - R) \quad (3)$$

得点イベントらしさ  $G$  を最小値が 0、最大値が 1 になるように正規化し、式 (4) に従って算出したしきい値を超えた箇所を得点イベントとして検出する。

$$\theta_G = \mu_G + 2 \times \sigma_G \quad (4)$$

ここで、 $\mu_G$  は得点イベントらしさの 1 試合中の平均、 $\sigma_S$  は得

表 3 実験データ

試合	種目	試合開催日	対戦カード	時間	ツイート数
s1	サッカー	2012/07/01	UEFA EURO2012 決勝 (スペイン対イタリア)	165 分間	410,363
s2	サッカー	2012/12/16	FIFA CWC 決勝 (チェルシー対コリンチャンス)	120 分間	22,075
b1	野球	2012/10/30	プロ野球 日本シリーズ 第 3 戦 (日本ハム対巨人)	270 分間	28,979
b2	野球	2012/10/31	プロ野球 日本シリーズ 第 4 戦 (日本ハム対巨人)	270 分間	35,359

表 4 試合中のツイートを収集する際に検索語に設定したハッシュタグ

種目	大会名	ハッシュタグ
サッカー	UEFA EURO2012	#EURO2012, #Euro2012, #euro2012
サッカー	FIFA CWC	#CWC
野球	プロ野球 日本シリーズ	#lovefighters, #kyojin

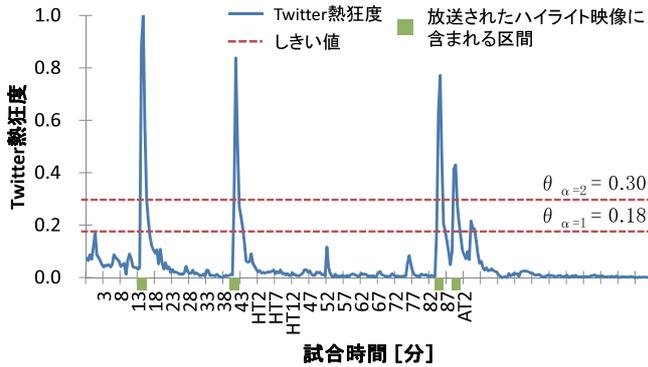


図 2 試合 s1 に対する提案手法の適用結果

点イベントらしさの 1 試合中の分散を示す。

これにより、提案手法によって検出されるイベントの性質や有効性を調査した。

#### 4.2 結果

ツイートから算出した Twitter 熱狂度としきい値の例を図 2 に示す。しきい値を横方向の点線を表示し、テレビ番組で放送されたその試合のハイライト映像に含まれるイベント区間を横軸上の矩形で示した。なお HT はハーフタイム、AT はアディショナルタイムを意味する。実験結果として、Twitter 熱狂度がしきい値を超えた時点もしくは放送されたハイライト映像に含まれる場面に発生したイベントの詳細を表 5~表 8 に示す。ハイライト映像に含まれるイベントの内容は放送局によって異なるため、ここでは放送局ごとにハイライト映像に含まれていたイベントを示す。実験の結果、提案手法は視聴者の興奮が高まるようなイベントを検出できた。そのため、視聴者視点に基づくスポーツ試合中のイベント検出に有効であるといえる。

#### 4.3 考察

実験結果に基づき、提案手法が検出できたイベント、できなかったイベント、参考手法 [12] との比較、提案手法の言語とスポーツ種目に対する非依存性の 4 つの観点から考察する。

##### a) 提案手法が検出できたイベント

提案手法により算出した Twitter 熱狂度と、 $\alpha = 2$  とした場合のしきい値を用いてイベント検出を行った結果、サッカーの 2 試合と野球の 1 試合に対して全ての得点イベントを検出できた。また、残る野球の 1 試合に対しても、10 件の得点イベント

表 5 試合 s1 において検出されたもしくは実際に放送されたイベントの種類ごとの数 (試合中の得点イベントは全 4 件)

イベント	a 局のハイライト映像内のイベント数	参考手法 [12]	提案手法	
			$\alpha = 2$	$\alpha = 1$
得点	4	4	4	4
区切り	0	1	0	2
その他	0	0	0	0

表 6 試合 s2 において検出されたもしくは実際に放送されたイベントの種類ごとの数 (試合中の得点イベントは全 1 件)

イベント	a 局のハイライト映像内のイベント数	参考手法 [12]	提案手法	
			$\alpha = 2$	$\alpha = 1$
得点	1	1	1	1
区切り	0	1	1	1
その他	2	2	0	3

表 7 試合 b1 において検出されたもしくは実際に放送されたイベントの種類ごとの数 (試合中の得点イベントは全 10 件)

イベント	ハイライト映像内のイベント数			参考手法 [12]	提案手法	
	a 局	b 局	c 局		$\alpha = 2$	$\alpha = 1$
得点	4	7	8	8	7	10
区切り	1	1	1	1	1	1
その他	0	6	7	5	1	6

表 8 試合 b2 において検出されたもしくは実際に放送されたイベントの種類ごとの数 (試合中の得点イベントは全 1 件)

イベント	ハイライト映像内のイベント数			参考手法 [12]	提案手法	
	a 局	b 局	c 局		$\alpha = 2$	$\alpha = 1$
得点	1	1	1	1	1	1
区切り	1	1	1	1	1	1
その他	10	2	5	9	6	12

のうち 7 件を検出でき、しきい値を  $\alpha = 1$  まで下げると全ての得点イベントを検出できた。

従って提案手法は、スポーツ試合中のイベントのなかでも重要な得点イベントを高い精度で検出できた。放送されたハイライト映像との比較の結果、本手法は、複数の放送局がハイライト映像を放送した試合において、すべての放送局がハイライト映像に含めたイベントを検出した。よってハイライト映像に含

むべきイベントを正しく検出できたといえる。また、得点以外のイベントではヒットやファインプレーを検出した。これは、ヒットによって塁が埋まる過程における視聴者の興奮の高まりや、ファインプレーによってピンチを凌いだことに対する視聴者の喜びが Twitter 熱狂度に表れたためであると考えられる。これらを検出することは、視聴者視点によるイベント検出において重要なため、本手法は視聴者の視点を考慮したハイライト映像作成のためのイベント検出手法として有効である。

しかし、日本シリーズ第3戦（試合 b1）において、本手法は誤って試合終了を検出した。これは、勝利に対する喜びや敗北に対する落胆によって、Twitter 熱狂度が大きくなったためであると考えられる。試合中のイベントではないイベントを検出してしまったため、本手法の目的からすると誤検出であるが、この結果は試合の結果が決まる瞬間が視聴者にとって興奮を高める瞬間であることを示しており、ハイライト映像に含めるべきイベントである可能性がある。今後、試合終了時をハイライト映像に含めるべきか検討する必要がある。

#### b) 提案手法により検出できなかったイベント

放送されたハイライト映像に含まれており、提案手法によって検出できなかったイベントもあった。例えば、日本シリーズの第3戦（試合 b1）における5回表の得点イベントも、 $\alpha = 2$ としたしきい値を用いて場合、検出できなかった。このイベントは、無死1、3塁からダブルプレーの間に得点したもので、得点イベントではあるものの、盛り上がりは小さかった。このような得点イベントも、本手法では検出が難しい。試合の進行を把握するためのイベント検出という点では未検出になるが、視聴者の視点に基づいたイベント検出という点では、本手法による結果は妥当と考える。

#### c) 参考手法 [12] と提案手法の比較

また、提案手法と参考のために示した得点検出手法の検出結果を比較したところ、得点検出手法では提案手法よりも多く得点イベントを検出できていた。しかし、得点検出手法は得点と区切り以外のイベントも多く検出していた。一方で提案手法は、 $\alpha = 2$ のしきい値を用いたときはその他のイベントの検出数は少なく、 $\alpha = 1$ とするとその他のイベントを検出するようになっていた。しきい値により得点とそれ以外のイベントを分離できる点で、提案手法は参考のために示した得点検出手法よりも優れているといえる。

#### d) 提案手法の言語とスポーツ種目に対する非依存性

提案手法の言語非依存性に関して、本実験で使用したツイートは英語、日本語、スペイン語、イタリア語など複数の言語が万遍なく混在していた。提案手法により、これらのツイートをを用いてイベント検出ができたことから、提案手法が言語に依存しないことを確認した。また、サッカーと野球という異なる種類のスポーツに対しても同じ手法でイベントを検出できており、スポーツ種目に依存しにくいことも確認した。

## 5. まとめ

本報告では、Twitter に投稿されたツイートの文字列特徴から算出した“Twitter 熱狂度”に基づき、スポーツ試合のハイ

ライト映像作成のためのイベント検出手法を提案した。提案手法では、感嘆符の出現数と繰り返し文字表現を含むツイート数を興奮を示す特徴として、リツイート数を試合の区切りを示す特徴として利用した。評価実験の結果、提案手法がハイライト映像作成のためのイベント検出に有効であることと、提案手法の言語非依存性、スポーツ種目非依存性を確認した。今後は、新たな特徴の導入やイベントとして検出する対象の見直しなどを行うことで、視聴者視点をより考慮したイベント検出手法の実現を目指す。

## 文 献

- [1] Y.-L. Chang, W. Zeng, I. Kamel, and R. Alonso, “Integrated image and speech analysis for content-based video indexing”, Proc. IEEE Int. Conf. on Multimedia Computing and Systems 1996, pp.306–313, June 1996.
- [2] 宮森 恒, 中村 聡史, 田中 克己, “番組実況チャットに基づく視聴者視点を利用した放送番組のビュー生成”, 第 16 回データ工学ワークショップ (DEWS2005), 4B-i9, Feb. 2005.
- [3] I. Otsuka, R. Radhakrishnan, M. Siracusa, A. Divakaran, and H. Mishima, “An enhanced video summarization system using audio features for a personal video recorder”, IEEE Trans. Consumer Electronics, vol. 52, issue 1, pp.168–172, Feb. 2006.
- [4] Twitter Blog, “Twitter turns six”, <http://blog.twitter.com/2012/03/twitter-turns-six.html/> (2013/2/7 参照) .
- [5] SemioCast, “Twitter reaches half a billion accounts More than 140 millions in the U.S”, [http://semioCast.com/publications/2012\\_07\\_30\\_Twitter\\_reaches\\_half\\_a\\_billion\\_accounts\\_140m\\_in\\_the\\_US/](http://semioCast.com/publications/2012_07_30_Twitter_reaches_half_a_billion_accounts_140m_in_the_US/) (2013/2/7 参照) .
- [6] Twitter Blog, “ロンドンのハイライト”, [http://blog.jp.twitter.com/2012/08/blog-post\\_12.html/](http://blog.jp.twitter.com/2012/08/blog-post_12.html/) (2013/2/7 参照) .
- [7] Twitter, “Twitter で 2011 年を振り返る”, <http://yearinreview.twitter.com/ja/tps.html/> (2013/2/7 参照) .
- [8] 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”, 信学技報, MVE2010-162, Mar. 2011.
- [9] J. Lanagan and A.F. Smeaton, “Using Twitter to detect and tag important events in live sports”, Proc. 5th AAAI Int. Conf. on Weblogs and Social Media, pp.542–545, July 2011.
- [10] S. Zhao, L. Zhong, J. Wickramasuriya, and V. Vasudevan, “Human as real-time sensors of social and physical events: A case study of Twitter and sports games”, Tech. Rep. TR0620-2011, Rice Univ. and Motorola Labs, June 2011.
- [11] D.A. Shamma, L. Kennedy, and E.F. Churchill, “Tweet the debates: Understanding community annotation of uncollected sources”, Proc. 1st ACM SIGMM Workshop on Social Media, pp.3–10, Oct. 2009.
- [12] 富田 大志, 道満 恵介, 井手 一郎, 出口 大輔, 村瀬 洋, “Twitter を用いたスポーツ試合中のイベント検出に関する検討”, 信学会 HCG シンポジウム 2012 論文集, pp.492–498, Dec. 2012.
- [13] 三菱 UFJ リサーチ&コンサルティング・(株) マクロミル, “2012 年スポーツマーケティング調査プレスリリース”, [http://www.murc.jp/publicity/press\\_release/pr-121011.pdf](http://www.murc.jp/publicity/press_release/pr-121011.pdf) (2013/2/7 参照) .
- [14] Fédération Internationale de Football Association, “FIFA’s Member Associations”, <http://www.fifa.com/aboutffifa/organisation/associations.html/> (2013/2/7 参照) .
- [15] 高橋 雄太, 片岡 義雅, 浅井 洋樹, 山本 祐輔, 秋岡 明香, 山名 早人, “繰り返し表現を含んだ感情的なツイートの抽出”, 第 4 回データ工学と情報マネジメントに関するフォーラム (DEIM2012) 論文集, C3-6, Mar. 2012.