

キュレータの価値観に基づく 複数映像情報流の実時間編纂支援システム

藤澤 和輝^{1,a)} 平部 裕子^{1,b)} 荒川 豊^{1,c)} 安本 慶一^{1,d)}

概要：

2020年の東京オリンピックを機に、サッカーなどのスポーツ競技を対象とした新しいサービスの登場が期待されている。現在のスポーツのライブ中継は、スタジアムの複数箇所に設置されたカメラで撮影されている映像を、経験豊富なディレクターが動的に取舍・選択することで、放送する映像をリアルタイムに編纂し配信している。しかし、放送される映像は1種類のみであり、様々な嗜好を持つ視聴者の多様なニーズに応えられていない。熱心な視聴者やマニアックなファンが独自の価値観に基づき、スタジアムにおいて様々なアングル、ズームレベルで撮影された多数の映像ストリームから必要なものをリアルタイムに選択・編纂し配信できれば、同じスポーツ競技の映像から異なるコンテンツが複数編纂でき、様々な嗜好を持つ視聴者の多様なニーズに合わせた次世代ライブ中継サービスが実現可能になる。本稿では、キュレータと呼ばれるコンテンツの編纂者が、情報流と呼ばれる複数の映像ストリームやSNSストリームをリアルタイムに編纂する作業を支援するシステムを提案する。提案システムは、複数の映像情報流やSNS情報流をキュレータの価値観に合わせてスコア付けし、上位のもの数件のみを選択表示する機能と、コンテンツとして出力・配信する情報流（一つないし少数）を次々に切替えていく作業を支援する機能を提供する。デモでは、フットサルの試合を様々な角度、ズームレベルで撮影した映像情報流を題材に、キュレータが本支援システムを用いて複数情報流をリアルタイム編纂する様子を実演する。

1. はじめに

2020年の東京オリンピックを機に、サッカーなどのスポーツ競技を対象としたライブ中継（あるいは動画コンテンツの配信）や次世代のエンターテインメントサービスの登場が期待されている。現在、スポーツのライブ中継は、スタジアムの複数箇所にカメラとカメラマンが配置され、ディレクターがカメラマンと協調して映像を取捨・選択することで、放送する映像をリアルタイムに作成し配信している。しかし、放送される映像は通常1種類のみであり、様々な嗜好を持つ視聴者の多種多様なニーズに応えられていないと思われる。視聴者やファンが独自の価値観に基づき、スタジアムにおいてGoogle Glassなどを用いて様々なアングル、ズームレベルで撮影された映像ストリーム群から必要なものをリアルタイムに選択・編纂し配信できれば、様々な嗜好を持つ視聴者の多様なニーズに合わせた次

世代ライブ中継サービスが実現可能になると思われる。映像ストリームの選択や編纂作業は編纂者の価値観および視聴者の要求に合わせてリアルタイムに行う必要があり、その作業は非常に複雑なものになることが予想されるため、実時間での編纂作業を支援するシステムが必須となる。

本稿では、文献[1]で提案されている情報流キュレーション基盤に基づき、キュレータと呼ばれる編纂者が、情報流と呼ばれる多数の映像ストリームやSNSストリームからリアルタイムに必要なものを選択・融合し、新たなコンテンツとして編纂・配信する作業を支援するシステムを提案する。提案システムは、(1) 複数の映像情報流やSNS情報流をキュレータの価値観に合わせてスコア付けし、上位のもの数件のみを選択・表示する機能と、(2) それらの情報流からコンテンツとして出力・配信する情報流（一つないし少数）を次々に切替えていく作業を支援する機能を提供する。

(1)、(2)の機能を、情報流編纂インタフェース、アクションインタフェースとして設計・実装した。情報流編纂インタフェースでは、キュレータが、注目する選手やズームレベルの度合い、SNS情報流における指定したキーワードの出現数など、自身がどの項目をどの程度重要視す

¹ 奈良先端科学技術大学院大学 Nara Institute of Science and Technology

a) fujisawa.kazuki.ey4@is.naist.jp

b) hirabe.yuko.ho2@is.naist.jp

c) yutaka.arakawa@gmail.com

d) yasumoto@is.naist.jp

るかを価値関数として設定し、各映像情報流に予め付加されたメタ情報を用いて価値関数の値を計算する。また、価値関数の値が大きい幾つかの情報流を選抜・表示するとともにその変化の様子を可視化する。さらに、キュレータが、出力する情報流を動的に選択できるようになっている。アクションインターフェースではキュレータが編纂したコンテンツがリアルタイムに表示される。

これらのインターフェースを用いることにより、様々な価値観を持ったキュレータたちが、同じ映像・SNS 情報流から、複数の異なるリアルタイムコンテンツ（特定の選手やチームの動きに重点を置いたものや、観客席や監督・コーチの反応に重点を置いたもの、俯瞰視点に重点を置いたものなど）を編纂し配信することが可能になり、視聴者の多様なニーズに応えることが可能になると考えている。

本デモンストレーションでは、フットサルの試合を様々な角度、ズームレベルで撮影した映像情報流を題材に、キュレータが本支援システムを用いて複数情報流をリアルタイム編纂する様子を実演する。

以下、2 章で関連研究、3 章で提案システムの設計と実装について述べる。4 章では、デモの詳細について述べ、5 章で本稿のまとめを述べる。

2. 関連研究

好きなアングルやズームレベルを動的に切り替えるなど、視聴者のニーズを満たす動画配信を実現する研究がいくつか行われている。米澤ら [2] はニコニコ動画 [3] を利用した、視聴者協力型ライブ演出システムを提案している。このシステムは、ニコニコ動画を通じて演奏を配信する演奏者とその演奏を視聴するニコニコ動画の視聴者を対象としている。システムを利用する演奏者は演奏会場に複数の固定カメラを設定しており、視聴者はシステムに設定されているコマンドを用いて、視たい映像を撮影しているカメラに切り替えることができる。これにより、視聴者は演奏会場にあるそれぞれのカメラ画面を自分が視たいアングルやズームレベルに動的に切り替えながら演奏を視聴することができる。このシステムを用いて演奏者である配信者と視聴者が協力することで、映像の視覚的演出効果を高められることが示されている。

Louise ら [4] は、ダンスや音楽パフォーマンスといったライブ会場にいる複数の観客にタブレットを渡し、これらのタブレットを持つ観客がライブ会場のパノラマ映像をリアルタイムに自由にズームしたりパンする（左右に振る）ことができる仕組みを提案している。このシステムを通して得られる映像は、ライブ会場の外に居る視聴者に配信される。これにより、ライブ会場外の視聴者はタブレットを持つ観客の様々なアングルや趣向から生まれる映像を選択して視聴することができる。

上で述べたシステムは、視聴者は、編集され完成された 1

種類の映像を視聴するのではなく、さまざまな角度から複数のカメラで撮影された複数の映像の間を、視聴者が自身の意志で動的に切り替えて視聴できるという点で従来のライブ中継と一線を画しており、本稿の目的である、様々な嗜好を持つ視聴者の多様なニーズに合わせたライブ中継を実現すること、を部分的に満たしている。しかし、視聴者のそれぞれが自ら映像を切り替えなければならない点や、映像を好みに合わせて切替えた結果を他者に再配信したり、多数の映像情報流がある場合に、それらから重要なものだけを自動的に選別するといったキュレータ支援機能が提供されていないという点で、本稿の目的には不十分であると考えられる。

3. スポーツ競技を対象としたキュレータ支援システムの設計と実装

3.1 システムの概要

図 1 にシステムの構成図を示す。本システムはある特定のスポーツ競技について撮影・配信されている映像情報流と SNS につぶやかれた関連メッセージ（SNS 情報流と呼ぶ）をキュレーション（編纂）の対象としている。提案システムは編纂インターフェース、アクションインターフェースで構成される。キュレータはラベル付けされた映像情報流と SNS 情報流のキュレーションを行い、アクション（視聴者へのキュレーション結果のフィードバック）を行う。本システムが対象とするキュレータは、一般ユーザであり映像編纂のプロではないため、何の情報もなくキュレーションを行うことは難しい。そこで、キュレータの価値観に合わせて各情報流にスコア付けする価値関数をキュレータ自身が設定し、価値関数の算出結果をもとに数値の高い幾つかの情報流を編纂インターフェースに表示する。

SNS 情報流の内容およびその時間的変化の理解を促進するため、編纂インターフェースには、キュレータが設定したキーワードが SNS メッセージに含まれる（単位時間あたりの）頻度を可視化する機能も実現する。これによりキュレータが SNS 情報流から編纂結果のコンテンツに含めるメッセージをピックアップするのを支援する。

映像はリアルタイムに流れている。それと同時に価値関数も映像の内容に応じて実時間で計算・更新されていく。次節では、既存の映像キュレータ支援システムとの違い、価値関数の設定方法、計算方法について詳述する。

3.2 既存の映像キュレータ支援システムとの違い

既存の映像キュレータ支援システムとして、ビデオジョッキーソフトが挙げられる。VDMX5[5] 等のビデオジョッキーソフトでは、VJ と呼ばれる映像キュレータが、あらかじめ作っておいたり、リアルタイムで入手した複数の映像をリアルタイムに切替えて一つの映像に編纂し配信する。

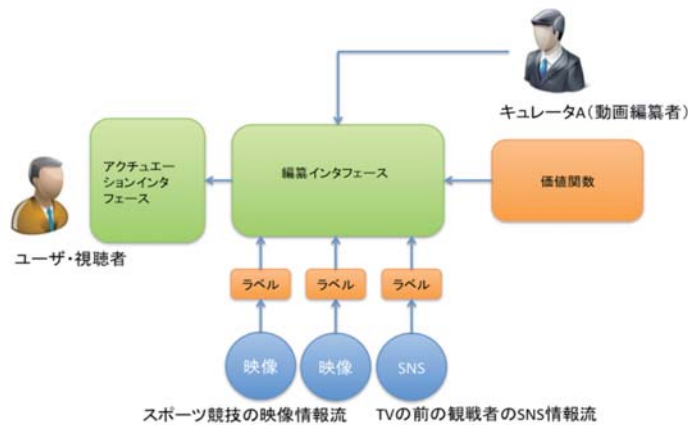


図1 キュレータ支援システムの構成

しかしビデオジョッキーソフトでは、映像を単純に切替える事や複数の映像を重ね合わせる事が主な機能であり、選択元の映像の数や動的に内容が変わっていくリアルタイム映像の数が多くなった場合、映像キュレータはそれら全ての映像の内容変化を追えなくなるという問題がある。一方、提案システムでは、映像を切替える事に加えて、映像情報流が多くなった時でも、価値関数により、切替える候補の映像情報流を、キュレータの価値観に合った情報流だけに絞り込む事や、配信する映像情報流における各時刻の映像フレームをどれか一つの切替元映像に限定することなく、1つのフレームの中にレイアウトを変えて複数表示することも可能である。また、提案システムでは、SNS情報流も絞り込み、映像情報流と同時表示をすることが可能である。

3.3 価値関数の設定法

価値関数は、キュレータによって設定され、各映像情報流の現在時刻における映像の重要度を反映する。ある映像情報流に対して価値関数を適用するには、その情報流の各時間区間に予めメタ情報が付与されている必要がある。フットサルの試合を撮影した情報流におけるメタ情報の項目の例を表1に示す。本稿では、各映像情報流の各時刻の映像に対し、表1のメタ情報の各項目に対する度合い(0-1の実数とする)が予め付与されているものとしている。映像情報流の各時刻の映像に対してリアルタイムにメタ情報を付与するためには、既存のオブジェクト認識手法が利用できるが、簡単のため、本稿では人出でメタ情報の付与を行うことにした。本稿では、キュレータは価値関数として、

表1 フットサル映像を例にした場合のメタ情報の項目

項目	説明
注目人物	映っている注目人物の大きさ
注目チーム	映っているチームのメンバの大きさ
俯瞰度	映っているコート面積の広さ
観客席	映っている観客席面積の広さ

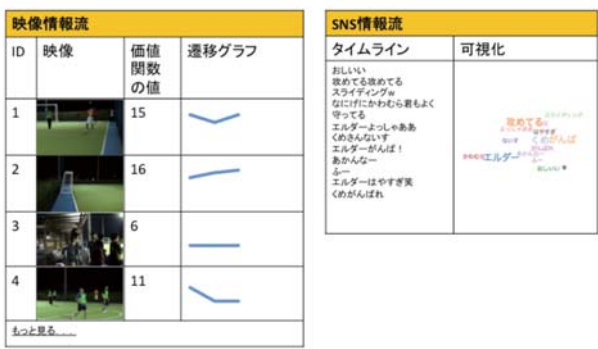


図2 編集インタフェースのイメージ

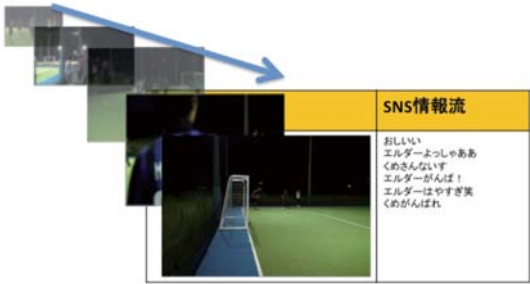


図3 アクチュエーションインタフェースのイメージ

表1で示した各項目に対する重みを設定するものとする。注目人物、注目チーム、俯瞰度、観客席の項目に対する重み(実数)をそれぞれ、 w_1, w_2, w_3, w_4 と表記する。ここで、 $w_1 + w_2 + w_3 + w_4 = 1$ とする。重みの数値が大きい程、その項目を重要視することを意味する。情報流 f の時刻 t の映像における表1の i 番目の項目の度合いを $f(i, t)$ と表記する。情報流 f の時刻 t における価値関数 $V(f, t)$ の値は、次のように算出される。

$$V(f, t) = \sum_{i=1}^4 w_i \cdot f(i, t)$$

関数の値が大きくなる程、キュレータにとって価値の高い映像であることを意味する。価値関数の値は、映像とともに編集インタフェースの画面上に表示される。

3.4 編集インタフェースとアクチュエーションインタフェース

図2に編集インタフェースのイメージ図を示す。図に示す通り、本インタフェースには、各映像情報流に対し、ID、映像、価値関数の値とその時間遷移グラフが、また、SNS情報流に対しては、ツイートのタイムラインと頻出キーワードの時間変化が可視化されたものが表示される。映像情報流に表示されている映像をキュレータが選択すると、図3に示すアクチュエーションインタフェースの画面がその映像に切り替わる。そしてキュレータがタイムラインのツイートをクリックすると、そのツイートがアクチュエーションインタフェースに反映される。キュレータは、

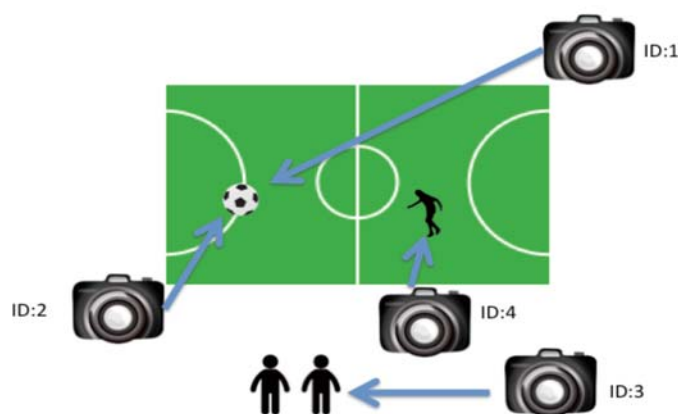


図 4 使用する映像情報流の撮影環境

編集インタフェースを見ながら、アクションインタフェースに出力する情報流を動的に選択していくことで、配信するコンテンツがリアルタイムに編集されていく。

4. デモ

フットサルの試合を様々な角度、ズームレベルで撮影した映像情報流を用いて、キュレータの価値観に基づいてリアルタイムに配信コンテンツを編集する様子をデモで示す。その際、タブレット PC とノート PC を用意し、タブレット PC で編集を行い、ノート PC にアクションを行う。使用する映像情報流の撮影環境を図 4 に示す。ID1 の映像情報流と ID2 の映像情報流はフットサルコートの対角線上の角でボール周辺の映像が撮影されている。ID3 の映像情報流は固定カメラによって観客の様子が撮影されている。ID4 の映像情報流は観客席で注目人物がずっと撮影されている。ズームレベルは ID1 と ID3 が比較的高めなのに対し、ID2 と ID4 は低めである。これらの取得した複数の映像情報流をサーバ上にアップロードして、ストリーミング配信でキュレータ支援システムに収集する。また、SNS 情報流の収集については Twitter API で行う。

5. おわりに

本稿では、キュレータと呼ばれるコンテンツの編集者が、情報流と呼ばれる複数の映像ストリームや SNS ストリームをリアルタイムに編集する作業を支援するシステムを提案した。また提案システムは複数の映像情報流や SNS 情報流をキュレータの価値観に合わせてスコア付けし、上位のもの数件のみを選択表示する編集インタフェースと、それによって次々に切替えられた情報流をコンテンツとして出力・配信するアクションインタフェースで構成される。今後はシステムに関するアンケートをとりユーザビリティを調べたり、映像情報流へのメタ情報付与の自動化について検討を進める予定である。

謝辞

本研究の一部は、JSPS 科研費 25280031 および 26220001 の助成のもと行った。ここに記して謝意を表す。

参考文献

- [1] 安本慶一, 山口弘純. 多数のデータストリームを実時間で融合・編集し活用するための次世代「情報流」技術, pp. 35–41. 情報処理 Vol.55 No.11 Nov.2014, 2014.
- [2] Takuro Yonezawa and Hideyuki Tokuda. Enhancing communication and dramatic impact of online live performance with cooperative audience control. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing, UbiComp '12*. ACM, 2012.
- [3] ニコニコ動画. <http://www.nicovideo.jp/>, Oct.2014.
- [4] Louise Barkhuus, Arvid Engström, and Goranka Zoric. Watching the footwork: Second screen interaction at a dance and music performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14*. ACM, 2014.
- [5] Vdmx5. <http://vidvox.net/>, Oct.2014.