

研究論文

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

齊藤 義仰^{1,a)} 磯貝 佳輝¹ 村山 優子¹

受付日 2011年5月20日, 採録日 2011年12月2日

概要: 動画投稿サイトには膨大な動画が投稿されており, 自分が興味のある動画を探すのは難しく, 予想していた動画内容と実際の動画内容が異なることがある. 動画共有サービス利用者が見たい内容の動画を適切に選択できるようにするためには, 動画の内容がどのようなものか見せるための短時間の紹介動画を参考にして, 動画を見るかどうかを判断してもらう方法が考えられる. 本稿では, 動画のシーンに対して視聴者がコメントを時系列データとして投稿できる動画共有サービス上において, 動画選択支援のための紹介動画作成手法を提案する. 提案手法では, 単位時間あたりのコメント数に着目し, コメントの多い箇所が必要なシーンであるという仮説のもと紹介動画の作成を行う. 作成アルゴリズムを検討するため, 紹介動画に必要なシーンについて調査を行った. その結果, 単純にコメント数のみでは紹介動画は作れないことが分かった. そこで, コメントの多い箇所のうち不要なシーンについて, コメントの内容に着目した除外方法について検討を行った. 検討結果から得られたいくつかの手法について, 単純にコメント数のみを用いた手法と比較し, 評価した結果について報告する.

キーワード: インターネット放送, インタラクティブ TV, 動画要約

A Method of Digest Video Production to Support Video Search Using Audience's Comments

YOSHIA SAITO^{1,a)} YOSHIKI ISOGAI¹ YUKO MURAYAMA¹

Received: May 20, 2011, Accepted: December 2, 2011

Abstract: Since these video sharing services have several million videos, it is difficult to find interesting videos for users. To find an objective video, the users typically search several videos for keywords and then select a video from thumbnail images and titles of the searched videos. However, these thumbnail images and titles are not always consistent with the users' expectations. We believe to provide short-time videos which show contents of these videos are able to support finding an objective video for the users. In this paper, we make the short-time video for introduction using the number of the comments per second with the assumption that a scene which has many comments is suitable for a scene of the short-time video. To make the short-time video, we conducted a preliminary experiment for scenes of the short-time video. From the experiment, we found unnecessary scenes in scenes which had many comments. Therefore, we studied to eliminate unnecessary scenes from the much commented scenes. The algorithm gained from the study was compared with an algorithm which uses only number of comments.

Keywords: Internet broadcasting, interactive TV, video summarization

1. はじめに

近年, インターネットの普及による広帯域化が進み, ネットワークを介した動画配信や動画共有サービスが日常的に使われるようになった. その代表例として, YouTube [1]

¹ 岩手県立大学
Iwate Prefectural University, Iwate 020-0193, Japan
^{a)} y-saito@iwate-pu.ac.jp

やニコニコ動画 [2] があげられる。これらの動画投稿サイトには膨大な動画が投稿されているが、膨大な動画の中から短時間で、自分が興味のある動画を探すのは難しい。一般的に動画を探す場合、特定のキーワードで動画検索を行い、興味のある動画の絞り込みを行う。しかし、動画を選ぶ際の情報はタイトル、サムネイル画像や短い紹介文のみであり、動画の内容がどのようなものかは予想するしかない。よって、実際に見てみたときに予想していた内容と実際の動画の内容が異なることがある。

動画共有サービス利用者が見たい内容の動画を適切に選択できるようにするためには、動画の内容がどのようなものか見せるための短時間動画（ここでは紹介動画と呼ぶ）を参考にして、本編動画を見るかどうかを判断してもらう方法が考えられる。紹介動画を作成する方法としては、動画投稿者などが手動で作成する方法と、アルゴリズムなどにより自動で作成する方法がある。人の手により手動で作成する場合は、作成者の手間と時間を要する。また、紹介動画を作成するための知識が必要となる。アルゴリズムなどにより自動で作成する場合は、人の手間や時間を必要としないため、容易に紹介動画を作成することができる。動画を要約する手法としては、映像・音声の特徴を利用する手法が一般的である。また、メタデータを利用することでさらに正確な要約を行う手法も提案されている。しかし、メタデータを利用する場合は、あらかじめメタデータが入力されていることが前提となるため、入力負担が問題となる。

一方、ニコニコ動画では、動画に対して視聴者がコメントを行うことができる。動画に対してコメントを行った場合、コメント投稿と同時にコメントを行った時点の動画の再生時間が記録される。視聴者が動画を再生し、自分や他の視聴者がコメントを行った再生時間に到達すると、映像上にコメントが表示される擬似同期型のコミュニケーションサービスを提供している。投稿されたコメントは動画の各シーンに関するメタデータと見なすことができる。よって、投稿されたコメント情報を時系列ごとに分析することで各シーンの性質を推測することができれば、メタデータ入力の負担なしに適切な紹介動画を作成できると考えられる。

本稿では、単位時間あたりのコメント数に着目し、コメントの多いシーンが紹介動画に必要なシーンであるという仮説のもと、紹介動画を自動で作成する手法を提案する。まず、紹介動画に必要なシーンを実験により調査し、紹介動画のシーン構成を明らかにする。次に、紹介動画に必要なシーンに含まれるコメントの特徴を分析し、コメントの特徴から必要なシーンを抽出するための手法を検討する。さらに、各手法が紹介動画に必要なシーンをどの程度抽出できるかを評価し、視聴者コメントを用いた紹介動画作成手法の実現可能性を示す。

本稿の構成を以下に示す。2章では、動画要約に関する関連研究について述べる。3章では、視聴者コメントを用いた

紹介動画作成手法のモデルを示す。4章では、紹介動画に必要なシーンと視聴者コメントの関係調査を行う。調査により、紹介動画を構成するシーンを明らかにする。5章では、コメントの特徴を分析することで、コメントの特徴から紹介動画に必要なシーンを抽出するためのアルゴリズムをいくつか提案する。6章では、再現率・適合率を用いて提案アルゴリズムの評価を行う。7章では提案アルゴリズムの実用化に向けた課題について考察し、8章で本稿をまとめる。

2. 関連研究

本章では、動画要約に関する研究として、映像・音声を用いた手法とメタデータを用いた手法をいくつかあげ、本研究との違いについて述べる。

2.1 映像・音声を用いた動画要約

要約動画作成の際には映像の動きや、音声といった動画に含まれる情報を用いるのが一般的である。

三浦らは、料理映像の自動要約手法について提案している [3], [4]。料理映像は食材を加工し、完成させるまで一連の流れから構成されている。このうち、「包丁で食材を切る」や「フライパンで炒める」といった手や器具の動きについて映像には一定のパターンが生まれる。このパターンを利用し、映像の要約を作成するというものである。

Michael らは、野球中継における自動要約手法について提案している [5]。野球中継のようなスポーツは、選手を応援する観客の歓声上がる。また、中継の進展に合わせてアナウンサーが実況を行う。これらの音声情報を用いた映像要約手法について提案している。

2.2 メタデータを用いた動画要約

映像要約手法としてメタデータを用いる方法が存在する。その一例としてメタデータを用いた手法と、利用するメタデータの内容について述べる。

Takahashi らは、メタデータを用いた動画要約について提案している [6]。ここでいうメタデータとは、MPEG-7 という動画に付随される番組の内容を示す。MPEG-7 のファイルには映像や音声情報以外に各シーンの内容について記述されており、このシーンの内容に関する記述を用いて動画の要約を作成する。

宮森らは、番組実況チャットを用いたテレビ番組の自動動画要約について提案している [7]。テレビ番組でのスポーツ観戦について、テレビを見ながら個々の感想などを書き込む。この際、応援しているチームが優勢になると、喜びのコメントや顔文字が現れるという特徴を利用し、番組のメタデータを作成する。このメタデータの応用例として、ストーリーボードや動画要約をあげているが、要約手法については述べられていない。

また、青木らはニコニコ動画に投稿されている動画に関

して、単位時間あたりのコメント数と映像要約の関係性について調査を行っている [8]。当該研究では、コメント数が多いところは面白いシーンに該当するが、単純にコメント数が多いところをとっただけでは映像要約への応用は難しいと指摘している。

2.3 本研究との比較

動画要約の研究では、映像の変化からシーンを検出するという映像に着目した手法や、音声の変化からシーンを検出するという音声に着目した手法が一般的である。しかし、映像・音声だけではシーンの情緒的な意味を分析することが難しいといった問題も指摘されている [9]。

メタデータを用いた映像要約について、動画に行われるコメントを動画とは異なる付随したメタデータにとらえ、これを利用するという点では似ていると考えられる。しかし、これらのメタデータ作成において作成者が付与するものに対して、本研究で用いるメタデータとは視聴者が付与するコメントであり、メタデータ作成に対する手間が存在しない。

また、一般的な動画要約では、動画の内容を短時間で理解するというを目的としている。そのため、要約された動画を見た後、要約される前の動画を見ることを目的としていない。一方で、本研究で提案する紹介動画の目的は、動画の内容がどのようなものであるかを見せ、視聴者が行う動画視聴の判断を支援するものである。既存の要約動画と紹介動画の違いとしては、要約動画が話の起承転結を理解するための重要シーンにより構成されているのに対し、紹介動画は元動画の視聴判断を目的としているため、一般的な要約動画とはシーン構成が異なると考えられる。たとえば、推理ドラマのように、トリックや結末を知ってしまうと元動画を見る必要性が著しく低下する動画などについては、シーン構成を工夫する必要がある。そのため、紹介動画作成のためのアルゴリズムが必要となってくる。

3. 視聴者コメントを用いた紹介動画作成手法

本章では、視聴者コメントを用いて紹介動画作成を行うシステムのモデルを提案する。また、提案システム実現に向けた課題について述べる。

3.1 提案システム

本研究で作成する紹介動画とは、元となる動画を視聴者に短時間で紹介し、視聴者の判断で見る、または見ないという意味決定を行うために利用するものである。提案システムでは、紹介動画を作るためニコニコ動画のコメントと同様の、視聴者のコメントとコメントを行った時点の動画の再生時間を利用し、紹介動画を自動で作成する。提案システムのモデルを図 1 に示す。動画を視聴するユーザは、提案システムから動画をダウンロードし、動画を視聴しながら時系列ごとにコメントを行う。投稿されたコメントとコメントを行った時点の動画の再生時間は、提案システムに蓄積される。提案システムは、これらのコメントと再生時間から、紹介動画に必要なシーンを分析し、短時間の紹介動画を作成する。作成した紹介動画を、面白い動画を探す視聴者や、目的の動画を探す視聴者に提供し、動画選択を行う際の判断基準に利用可能とする。視聴者から自発的に集まるコメントを動画のメタデータとして用いるため、動画投稿者やサービス提供者の負荷を増加させることはない。

3.2 提案システム実現に向けた課題

提案システムを実現するにあたり、本研究では、次に述べる 2 つの課題について取り組む。

1 つ目の課題は、紹介動画はどのようなシーンで構成されるのかを明らかにすることである。本研究で作成する紹介動画は、紹介動画を見た後に、オリジナルの動画を見ることを想定している。このように、一般的な要約動画とは性質が異なるため、紹介動画を構成するシーンも異なると考えられる。そのため、まず紹介動画に必要なシーンとは、動画中のどのようなシーンなのかを調査する必要がある。

2 つ目の課題は、紹介動画に必要なシーンと視聴者コメントの間に、どのような関係があるのかを明らかにすることである。視聴者コメントは、面白いシーンや印象に残ったシーンに対して投稿されるものである。そのため、視聴者コメントを時系列で集計し、単位時間あたりのコメント数が多いシーンを見つけることができる。これらのシーンは動画にとって重要なシーンであるため、紹介動画に必要なシーンとして利用できると考えられる。したがって、本研究では、コメント数の多いシーンが紹介動画に必要な

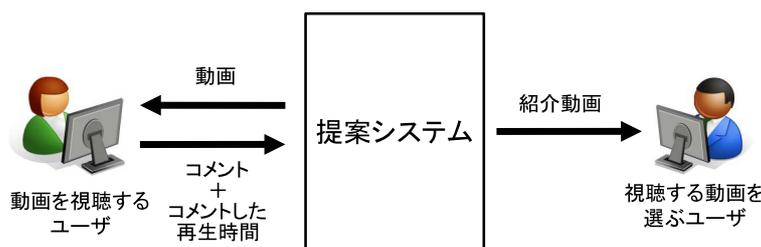


図 1 提案システムのモデル

Fig. 1 System model.

シーンであるという仮説を立て、その検証を行う。また、コメント数だけで解決できないシーンが存在すると考えられるため、それらの例外的なシーンについての対応を検討する。

4. 紹介動画に必要なシーンと視聴者コメントの関係調査

紹介動画を作成するにあたり、必要なシーンについて予備調査を行った。調査に用いる動画は、ニコニコ動画に投稿されている動画を用いた。その中から選んだジャンルは「ゲーム」と「料理」である。ゲームカテゴリはニコニコ動画に投稿されている動画のジャンルのうち最も多いためである。また、料理カテゴリは手順という意味で起承転結を持っているためである。各ジャンルから動画を2種類で計4種類の動画 (Game A, Game B, Cooking A, Cooking B) を再生数の多い順に選んだ。Game Aは鬼ごっこのように追ってくる敵から逃げまわるゲーム動画、Game Bはアクションゲームのプレイに合わせて機械音声で実況するゲーム動画である。Cooking Aは鶏肉を使って数種類の料理を作っていくという料理動画、Cooking Bはおつまみ料理を数種類作っていくという料理動画である。動画共有サービスを利用したことのある学生10名に、これらの4種類の動画を視聴してもらい、各動画について30秒前後(25~35秒)の紹介動画に必要なシーンを列挙してもらった。さらに、列挙したシーンには重要度を「A:絶対必要~E:どちらでもよい」として5段階でつけてもらった。

視聴者コメントについては、使用した4種類の動画からコメントを動画ごとに10,000件取得した。取得した各コメントデータは、日付、タグ、コメント番号、ユーザID、コメントが投稿された再生時間、コメントの内容により構成されている。

4.1 紹介動画に必要なシーン

被験者からの回答から得られた紹介動画に必要なシーンについて、各シーンの重要度を集計するため、重要度A~Eにそれぞれ5~1ポイントの数値を割り当て、必要なシーンとして選択された秒区間ごとにポイントを割り当て、その合計値を計算した。各動画における紹介動画に必要なシーンとその重要度を図2に示す。

Game Aの重要度の最大値は、500秒付近で19ポイントとなっている。また、840秒から1,030秒までは被験者が誰も選択しなかったことが分かった。選択傾向の特徴として、前半から中盤にかけて大きく選択割合が多くなっている箇所が多いことが分かった。

Game Bの重要度の最大値は、340秒付近で20ポイントとなっている。選択傾向としてはGame Aの重要度と同様に中盤から後半にかけての450秒から550秒は選択されない傾向にあることが分かる。また、選択の傾向としては重要度の高い箇所は150秒前後や280秒前後、330秒前後といった中盤に偏る傾向があることが分かった。

Cooking Aの重要度の最大値は、開始直後の20秒で18ポイントとなっている。また、選択の傾向としては、中盤である460秒から600秒までは被験者が誰も選択しなかったことが分かった。選択が多くなる傾向として、前半に集中していることが分かった。前半部分以外では、最後に近い箇所でも局所的ではあるが11ポイントとして重要度が高いことが分かった。

Cooking Bの重要度の最大値は、250秒前後で12ポイントとなっている。選択傾向として、重要度が大きく伸びた直後の270秒から440秒までは被験者が誰も選択していないことが分かった。また、重要度の合計値が他の動画の重要度に比べて伸びないのは、選択箇所にばらつきが生じ、重複する箇所が比較的少なかったためと考えられる。

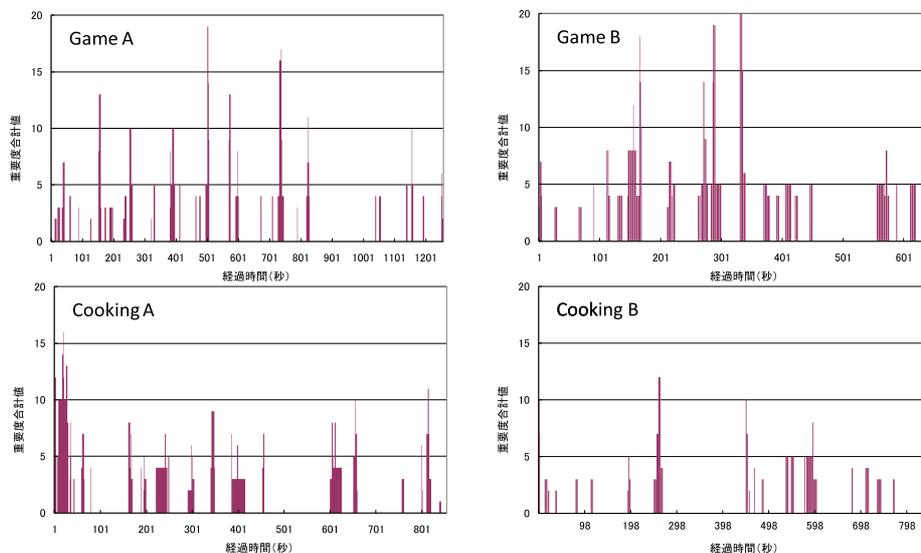


図2 各動画における紹介動画に必要なシーンとその重要度

Fig. 2 Necessary scenes for the introduction video and their degree of importance.

4.2 紹介動画に必要なシーンとコメント数との関係

各動画から取得した 10,000 件のコメントについて、1 秒あたりのコメント数の集計を行った。コメント数が多いシーンが紹介動画のシーンとして重要度が高いかを検証するため、適合率 (precision) と再現率 (recall) から F 値 (F-measure) を求めて検証を行った。F 値は、動画要約の精度を求めるためにしばしば利用される [10], [11], [12]。適合率と再現率、F 値はそれぞれ式 (1)~(3) で求めることができる。

$$\text{precision} = R/C \quad (1)$$

$$\text{recall} = R/N \quad (2)$$

$$F\text{-Measure} = 2 \cdot \text{precision} \cdot \text{recall} / (\text{precision} + \text{recall}) \quad (3)$$

ここで、 C = 全体のうち正解集合、 N = 抽出結果、 R = 抽出結果のうちの正解集合である。F 値は 0 から 1 の値をとり、1 に近いほど正解集合に近いことを示し、0 に近いほど正解集合とは異なることを示す。

コメントの多いシーン (N) と正解シーン (C) として F 値を求めた。コメントの多いシーンは、秒ごとのコメント数をコメントの多い順に並べたとき、上位 30 件に該当する箇所とした。また、正解シーンは、重要度の合計値が高い順に並べ替えたときにポイントの多かった箇所 30 秒分とした。各動画に対して F 値を求めた結果、Game A は 0.03、Game B は 0.07、Cooking A は 0.5、Cooking B は 0.2 となり、総じて F 値は低い値であった。そのため、コメントの多いシーンは正解シーンと単純には一致しないことが分かった。

F 値が低い原因として、コメントは多いが紹介動画には不必要なシーンと、コメントは少ないが紹介動画には必要なシーンの存在が考えられる。これらのシーンは単純なコメント数のみを用いた場合、除外・抽出することができない。そこで今回は、コメントは多いが紹介動画には不必要なシーンをいかに除外するかに注目し、F 値向上を試みる。

5. 視聴者コメントを用いた紹介動画作成アルゴリズムの検討

調査により、コメントの多いシーンは正解シーンと単純には一致しないことが分かった。本章ではコメントは多いが紹介動画には不必要なシーンを除外するアルゴリズムを検討し、F 値向上を試みる。

5.1 コメントが行われた時間をずらす手法

まずすべての手法を行う前処理として、視聴者が動画を見てコメントを入力するまでの時間を考慮する必要がある。視聴者は印象に残る動画のシーンを見て、そのシーンに対してコメントを入力する。しかし、コメントを入力している間も動画は再生されつづけるため、コメントの対象

となった動画のシーンとコメントが入力されたタイミングにズレが生じる。コメントが動画シーンに対するメタデータとして不正確なものになってしまうため、ズレを補正することが必要になる。

5.2 タグを含むコメントを除外する手法

コメント変化について特徴を調べたとき、特定のシーンで 1 秒間に瞬間的なコメントが行われるという特徴があることが分かった。加えて、このシーンは被験者の選択がないということが分かった。このシーンは動画に対するリアクションとは異なり、今回用いた動画の投稿サービスであるニコニコ動画の特有のコメントの特徴にあたる弾幕 (同じ時間に複数の同一コメントを行うという特徴) である。この弾幕コメントの特徴について分析を行った結果、コメントに対して色づけや文字の大きさが変化されていた。

この動画共有サービスであるニコニコ動画では、動画上にコメントを流すことで他の視聴者にコメントを見せることができる。この中で装飾機能として、コメントに色付けや文字の大きさを調節 (タグ) することができる。この特徴を利用し、特殊な装飾が行われているコメントを除外することでこのような不要なシーンを除外することが可能であると考えられる。

5.3 秒ごとの重複したユニーク ID を除外する手法

瞬間的にコメントが増える弾幕の特徴について、コメントを行った視聴者の ID について着目した。すると、特徴的に現れている弾幕を行っている視聴者の ID は同一であることが分かった。これは、弾幕コメントは一部の視聴者によるコメントであるということである。そのため、1 秒間に同一 ID の視聴者が行っているコメントを無効化し、コメントの数を数えることで弾幕状にコメントが行われているシーンを除外することができるのではないかと考えられる。

5.4 w を含まないコメントを除外する手法

時系列ごとに視聴者がコメントを行う可能な動画共有サービスであるニコニコ動画の特徴として、動画に対するリアクションとして面白いという意味をこめてコメントに「w」を打ち込む傾向がある。この「w」は笑い (warai) の頭文字をとったものであり、面白いシーンが現れたときにこのようなコメントを入力する傾向がある。一方、入力されたコメントについて反論のコメントが入力される場合や、心ない視聴者による煽りコメントや、それに対する中傷コメントが見られる場合がある。このように、映像とは直接関係のないところで、コメントが増えるという特徴が稀に見られる。「w」を含むコメントを抽出することで、動画のシーンに対するリアクションとしてのコメントのみを抽出することができ、動画とは無関係なコメントを除外することが可能なのではないかと考えられる。

6. 評価

提案した「コメントが行われた時間をずらす手法 (秒ずらし)」、 「タグを含むコメントを除外する手法 (タグ)」、 「秒ごとの重複したユニーク ID を除外する手法 (ユニーク ID)」、 「w を含まないコメントを除外する手法 (w 抽出)」 の 4 種類を組み合わせる評価を行った。 組合せについては以下のとおりである。

- 「秒ずらし」のみ
- 「秒ずらし」 & 「タグ」 & 「ユニーク ID」
- 「秒ずらし」 & 「w 抽出」
- 「秒ずらし」 & 「ユニーク ID」 & 「w 抽出」
- 「秒ずらし」 & 「タグ」 & 「ユニーク ID」 & 「w 抽出」

秒ずらしについては 0~9 秒の間で行い、0 秒とすれば秒をずらさない場合と同じであるため、すべてに秒ずらしを組み合わせたと同じであるため、すべてに秒ずらしを組み合わせた。 これらの手法を用いてコメントを加工した後、それぞれの F 値を算出した。 評価結果を図 3, 図 4, 図 5, 図 6, 図 7 に示す。

図 3 から秒ずらしのみを用いた場合の F 値の変化は動画の内容によって異なり、動画によって、ずらす秒数を適切に変化させる必要があることが分かった。 動画を見る視聴者の属性が、コメントを入力する速度に影響を与えているのではないかと考えられる。 また、図 4~7 から、何も

手を加えない場合に比べ、手法を組み合わせた場合は、いずれも F 値が向上したことが分かる。 一方で、秒ずらしと組み合わせない場合は、手法なしと F 値が同等もしくは低くなることが多い。 しかしながら、秒ずらしと組み合わせることにより、F 値の値が向上し、秒ずらしのみを用いる場合よりも F 値が向上することが分かった。 Game A, Game B, Cooking A の 3 種類の動画の F 値で最も良い値を示したのは「タグ」 & 「w 抽出」, 「ユニーク ID」 & 「w 抽出」, 「タグ」 & 「ユニーク ID」 & 「w 抽出」, の 3 種類であった。 このときの F 値は、「w 抽出」のみを行った場合

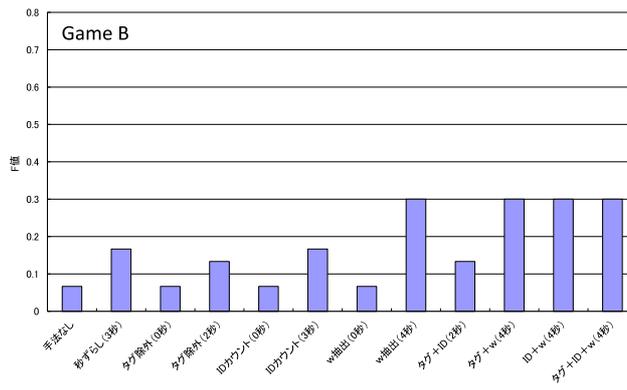


図 5 各手法を用いた場合の Game B の F 値の変化
Fig. 5 Change of F-measure in Game B with each scheme.

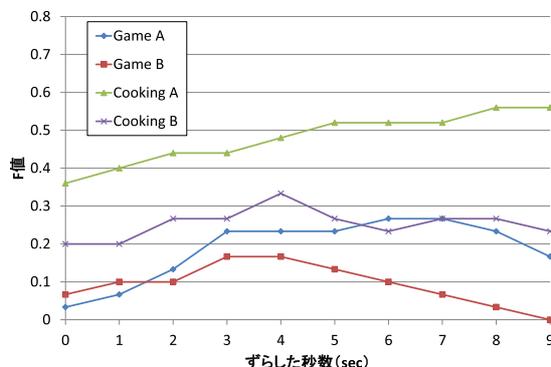


図 3 秒ずらしのみを用いた場合の F 値の変化
Fig. 3 Change of F-measure with shift of time.

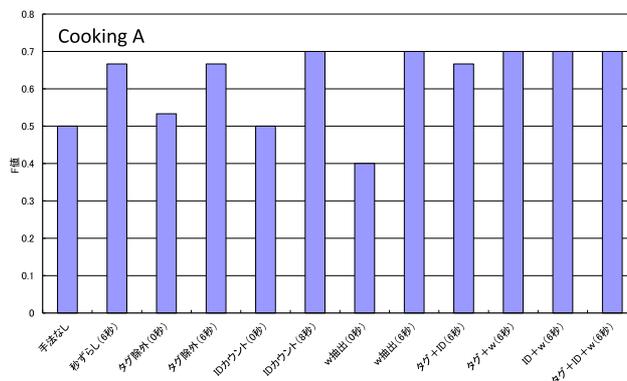


図 6 各手法を用いた場合の Cooking A の F 値の変化
Fig. 6 Change of F-measure in Cooking A with each scheme.

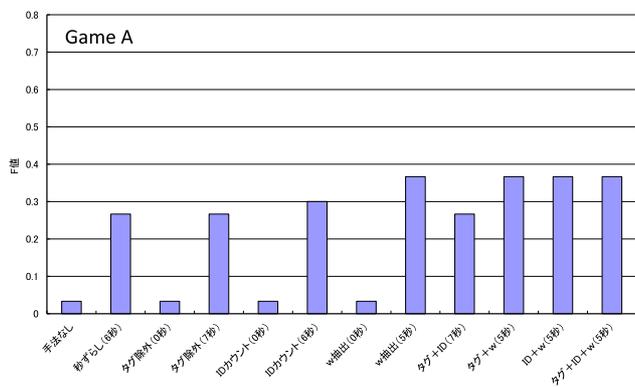


図 4 各手法を用いた場合の Game A の F 値の変化
Fig. 4 Change of F-measure in Game A with each scheme.

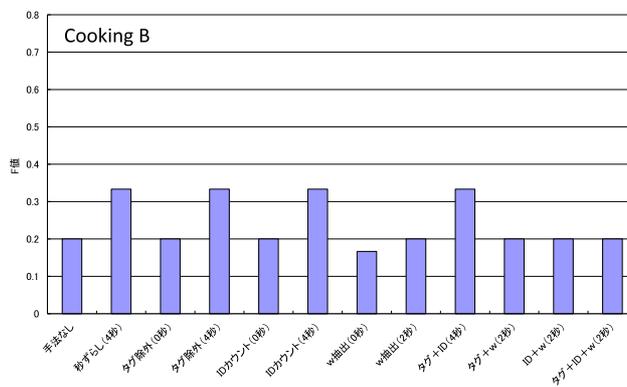


図 7 各手法を用いた場合の Cooking B の F 値の変化
Fig. 7 Change of F-measure in Cooking B with each scheme.

と同一であることから、手法すべての中では「w を含まないコメントを除外する手法」が最も有効ではないかと考えられる。Cooking A が他の動画よりも良い値を示している理由としては、それぞれの料理を作るシーンにエンターテインメント性があるためインパクトがあり、「w」を含んだコメントが、紹介動画として必要なシーンに多かったためだと考えられる。

一方 Cooking B の動画では、取得件数を 30 件に固定した場合「w を含まないコメントを除外する手法」は最も良い値ではなく、「ユニーク ID」と「タグ」の手法で最も良い値になることが分かった。この理由として考えられるのは、Cooking B の動画の特徴はエンターテインメント性が低く、面白いというシーンに対するリアクションとしての「w」を含んだコメントがあまり行われなためと考察する。そのため、「w」を含んだコメントのみを抽出し、その手法を用いてシーン検出を行うも、適切なシーンが検出されなかったと考えられる。実際に 1 秒あたりの平均の w を含むコメント数を見てみると、Game A: 4.58 コメント/秒、Game B: 8.09 コメント/秒、Cooking A: 3.82 コメント/秒、Cooking B: 2.722 コメント/秒となった。そのため、w を含むコメント数が多い動画には効果がある傾向があり、少ない動画では効果が薄いと考えられる。

このように、コメントから適切なシーン抽出が難しい場合、三浦ら [4], [5] の研究のように映像のパターンを分析する、映像・音声を用いた動画要約との併用が有効である。一方で、コメントから適切なシーン抽出が行える場合は、映像・音声解析といった時間がかかる処理が必要ないため、高速に紹介動画が生成できると考えられる。

7. 考察

提案手法により作成した動画の F 値は、Cooking A の動画については高い値を示した。しかし、その他の動画については 0.3~0.5 以下であり、今後 F 値向上に向けてさらなる検討を行う必要がある。特に、除外できなかったシーンについて分析を行い、それらのシーンに含まれるコメントを分析する。さらに、「w」を含まない動画については、形態素解析などを用いてコメントに含まれる単語を抽出し、「w」以外の単語または文節の意味を用いてシーンを分析する必要がある。

また、コメントが行われた時間をずらす手法では、動画ごとに F 値が最も高い時間が異なっていた。しかし、今回の動画作成においてずらす時間について最も適切な時間が分かったという前提のもと、最も F 値が高い時間にずらし動画作成を行った。そのため、今後はコメントをずらす時間を推測する必要がある。

今回の実験では、ゲームと料理のジャンルについてのみを対象に紹介動画作成アルゴリズムの検討を行った。しかし、実験結果からも、ジャンルが紹介動画作成アルゴリ

ムの有効性に与える可能性があることが分かる。そのため、ゲームと料理以外のジャンルの動画について有効性は確認できていない。また、同一ジャンルであっても、動画の構成によって F 値の変動は大きかった。そのため、今後は動画のジャンルや動画の構成に応じた紹介動画作成アルゴリズムの検討を行う必要がある。

本研究では、不必要なシーンを除外する手法について検討を行ったが、コメントが少なく必要なシーンの抽出手法については着目しなかった。そこで、コメントが少なく必要なシーンを抽出する手法について検討する必要がある。検討内容としては、コメントが少ない箇所を抽出することはコメントの件数が少なく着目することは難しいと考えられる。そこで、前後の時間に行われたコメントから必要か不要かを判断するなどの手法が考えられる。前後の時間によつてどのようなコメントが行われた場合、そのシーンが必要になるかを今後検討していく。

本稿で提案した紹介動画作成アルゴリズムはニコニコ動画のコメントを用いることを前提としている。そのため、ニコニコ動画の文化である「w」を含むコメントを用いる手法や、タグを用いる手法は他の動画共有サービスではそのままでは利用できない可能性がある。そのため、他の動画共有サービスで「w」に相当する文字や、タグに相当する機能と対応付ける必要がある。一方で、ユニーク ID を用いる手法は、コメント可能な動画共有サービスであれば、コメントに ID が付いている場合が多いため、そのまま利用できる可能性が高いと考えられる。

8. おわりに

本稿では、時系列ごとにコメントが可能なサービスを用いた、視聴者が元動画の視聴の判断を支援するための紹介動画作成手法を提案した。調査実験により、紹介動画に必要なシーンを明らかにし、それらのシーンと視聴者コメントとの関係から、視聴者コメントを用いた紹介動画作成アルゴリズムをいくつか検討した。評価を行った結果、「w を含まないコメントを除外する手法」が最も紹介動画に必要なシーンを抽出できることが分かった。

今後は、F 値のさらなる向上を目指して視聴者コメントの分析を行う。また、提案した手法では、不必要なシーンを除外する手法について検討を行ったが、コメントが少なく必要なシーンの抽出手法については着目しなかった。そこで、コメントが少なく必要なシーンを抽出する手法について検討を行う。

参考文献

- [1] YouTube, available from (<http://www.youtube.com/>), (accessed 2011-05).
- [2] ニコニコ動画, 入手先 (<http://www.nicovideo.jp/>), (参照 2011-05).

- [3] 三浦宏一, 浜田玲子, 井手一郎, 坂井修一, 田中英彦: 料理映像の特徴を利用した要約手法の検討, 電子情報通信学会技術研究報告 PRMU, パターン認識・メディア理解, Vol.102, No.155, pp.15-20 (2002).
- [4] 三浦宏一, 浜田玲子, 井手一郎, 坂井修一, 田中英彦: 動きに基づく料理映像の自動要約, 情報処理学会論文誌: コンピュータビジョンとイメージメディア, Vol.44, No.SIG9(CVIM7), pp.21-29 (2003).
- [5] Fleischman, M., Roy, B. and Roy, D.: Temporal Feature Induction for Baseball Highlight Classification, *Proc. ACM Multimedia*, Augsburg, Germany (2007).
- [6] Takahashi, Y., Niita, N. and Babaguch, N.: VIDEO SUMMARIZATION FOR LARGE SPORTS VIDEO ARCHIVES, *Proc. IEEE ICME 2005*, pp.1170-1103 (2005).
- [7] 宮森 恒, 中村聡史, 田中克己: 番組実況チャットを利用したテレビ番組のメタデータ自動抽出方式, 情報処理学会論文誌: データベース, Vol.46, No.SIG18(TOD28), pp.59-71 (2005)
- [8] 青木秀憲, 宮下 芳: ニコニコ動画における映像要約とサビ検出の試み, 情報処理学会研究報告, 2008-HCI-128/2008-MUS-75, Vol.2008, No.50, pp.37-42 (2008).
- [9] Tang, L.-X., Mei, T. and Hua, X.-S.: Near-Lossless Video Summarization, *ACM MM'08*, pp.352-360 (2009).
- [10] 橋本隆子, 白田由香利, 真野博子, 飯沢篤志: TV 受信端末におけるダイジェスト視聴システム, 情報処理学会論文誌: データベース, Vol.41, No.SIG3(TOD6), pp.71-84 (2000).
- [11] レー ヒェウハン, ルートラットデーチャークン ティティポー, 渡部徹太郎, 横田治夫: 講義講演ビデオからダイジェスト自動作成のための重要シーン抽出手法の評価, 電子情報通信学会第 19 回データ工学ワークショップ (DEWS2008) 論文集, pp.E4-1 (2008).
- [12] Hoashi, K., Sugano, M., Naito, M., Matsumoto, K. and Sugaya, F.: Video Story Segmentation based on Content-independent Low-level Features, 電子情報通信学会技術研究報告 PRMU, パターン認識・メディア理解, Vol.105, No.118, pp.43-48 (2005).



磯貝 佳輝 (学生会員)

岩手県立大学. 平成 21 年岩手県立大学ソフトウェア情報学部卒業. 平成 21 年 4 月から同大学大学院ソフトウェア情報学研究科博士前期課程でインターネット放送の研究に取り組む. 平成 23 年 3 月同博士前期課程修了.



村山 優子 (正会員)

岩手県立大学. 津田塾大学学芸学部数学科卒業. 三菱銀行および横河ヒューレット・パカード社に勤務. 昭和 59 年 University College London 大学院理学部計算機科学科修士課程修了. 平成 2 年同大学院博士課程修了. Ph.D.

(ロンドン大学). 慶應義塾大学環境情報学部非常勤講師を経て, 平成 6 年 4 月より広島市立大学情報科学部情報工学科講師, 平成 10 年 4 月より岩手県立大学ソフトウェア情報学部助教授. 平成 14 年 4 月より教授. 現在に至る. インターネット, ネットワークセキュリティ, 安心およびトラストの研究に従事. IEEE, ACM, 電子情報通信学会, 映像情報メディア学会, 日本 OR 学会, 情報知識学会各会員.



齊藤 義仰 (正会員)

岩手県立大学. 平成 18 年静岡大学大学院理工学研究科博士課程修了. 博士(情報学). 平成 16 年から平成 19 年まで独立行政法人情報通信研究機構(NICT) 特別研究員・専攻研究員で次世代無線ネットワークの研究に従事.

平成 19 年 10 月より岩手県立大学ソフトウェア情報学部講師. 平成 23 年 10 月より准教授. 現在に至る. インターネット放送, インタラクティブ TV の研究に従事. IEEE, ACM, 電子情報通信学会各会員.