

SNS 利用者の反応を用いた単語スコアに基づく 複数ニュース映像の要約

加藤 光佑[†] 井手 一郎[†] 出口 大輔^{††,†} 村瀬 洋[†]

[†] 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

^{††} 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: [†]katok@murase.m.is.nagoya-u.ac.jp, [†]{ide,murase}@is.nagoya-u.ac.jp, ^{††}ddeguchi@nagoya-u.jp

あらまし ニュース映像は社会に関する資料的な価値が高く、アーカイビングが重要視されている。アーカイブされたニュース映像を資料として活用する場合、話題の流れを辿り、内容を総合的に理解することが重要である。それを実現するために、強く関連するニュースストーリーを時系列に連鎖した構造（トピックスレッド構造）を抽出する手法が提案されている。トピックスレッド構造では、起点となるストーリーから関連するストーリーを辿ってその顛末を知ることができる。しかし、多様なトピックのストーリーが含まれる場合、利用者はトピックスレッド構造中のストーリーを逐一視聴する必要があり、負荷が高い。一方、近年ソーシャルネットワーキングサービス（SNS）が爆発的に普及している。SNS 利用者は放送映像を視聴しながら自分の意見や興味・関心がある情報の受発信をすることも多い。そこで、我々は視聴者視点で映像を編集することを考え、SNS 利用者の反応を用いた単語スコアに基づいた複数ニュース映像要約手法を提案する。

キーワード ニュース映像、ニュース映像アーカイブ、SNS

Summarization of multiple news videos based on term scores using responses of SNS users

Kosuke KATO[†], Ichiro IDE[†], Daisuke DEGUCHI^{††,†}, and Hiroshi MURASE[†]

[†] Nagoya University, Graduate School of Information Science

Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

^{††} Nagoya University, Information and Communications Headquarters

Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

E-mail: [†]katok@murase.m.is.nagoya-u.ac.jp, [†]{ide,murase}@is.nagoya-u.ac.jp, ^{††}ddeguchi@nagoya-u.jp

Abstract Archiving news videos is regarded as important since they are valuable as sources of important social information. When exploiting archived news videos as information sources, it is important to track the flow of topics to understand their contents comprehensively. Concerning it, a method that structures the chronological semantic relations between news stories, namely the “topic thread structure” has been proposed. It allows the understanding of the circumstance of the topic by tracking related stories one after another from the initial story. However, it imposes a user to watch many stories along the topic thread structure when it contains various topics. On the other hand, social networking services (SNS) have become very popular. SNS users often send and receive information in which they are interested while watching broadcast video. Thus, we propose a method that summarizes news videos based on term scores using responses of SNS users, in order to edit it from the viewpoint of the viewers.

Key words News video, news video archive, SNS

1. はじめに

近年、記憶装置の大容量化に伴い、放送映像を大量に蓄積

できるようになった。放送映像にはスポーツ、ドラマ、バラエティ、アニメーションなど様々な種類があるが、そのなかでもニュース映像は実世界の出来事を記録したものであるため、資

料的な価値が高い。そのため、ニュース映像のアーカイブが重要視されている [1]。アーカイブされたニュース映像を資料として活用する場合、調査対象のトピックに関するニュース映像群の一部を見て理解するだけでなく、トピックの流れを辿り、内容を総合的に理解することが重要である。例えば、ある殺人事件について調査する場合、事件の発端から容疑者の割り出し、容疑者の逮捕、動機の究明、容疑者の起訴、裁判の判決、というトピックの流れを辿ることができれば、事件の全容を理解することができる。しかし、膨大な量のニュース映像が蓄積されているため、目的の事象に関連するものを人手で検索、追跡するのは非常に手間を要する。そこで、ニュース映像におけるトピックの検索・追跡に対する計算機の支援が必要である。

このような背景をうけ、ニュース映像の閲覧を支援することを目的として、ニュース映像の意味を解析して時系列に構造化する研究が行われている。Duygulu ら [2] は特定のトピックに強く関連するストーリー（ニュース映像の意味的な最小単位 [3]）を直線状に時系列に連ねる方法を提案した。しかし、この方法では大規模なニュース映像アーカイブを対象として利用者が追跡、理解したい場合に、膨大な数のストーリーを逐一直線状に辿らなければならない。映像を視聴しながら、数 10 個連なったストーリーの内容を逐一把握するのは非常に負荷が高いため、実用性に問題がある。これに対し Wu ら [4] は、特定のトピックに関連するストーリーを集めたクラスタにおいて、時系列の前後関係と話題の変化に応じて 2 分グラフを構築する手法を提案した。しかし、この方法でも、新規ストーリー同士の関係は時系列の前後関係のみであり、同時並行して進む複数のトピックの流れを表現できない問題がある。そこで井手らは同時並行して進む流れを表現する時系列意味構造（トピックスレッド構造）を抽出する手法 [5] と、それを利用した閲覧インタフェースである mediaWalker [6] を提案した。mediaWalker の概観を図 1 に示す。

トピックスレッド構造は単純な木構造に比べ分岐が少なく、重複ノードが存在しないため、効率的に映像群を閲覧できる。トピックスレッド構造では、起点となるストーリーから関連するストーリーを辿ってその顛末を知ることができる。井手らの報告 [5] によると、122 日分のニュースに出現する 1,431 ストーリーを起点として、探索期間を 100 日としてトピックスレッド構造を構築した場合、2 ストーリー以上からなるトピックスレッド構造が 437 件得られた。それらを分析すると、探索期間 100 日間のニュース映像の総映像長 47.7 時間に対して、トピックスレッド構造を構築することによって平均で総映像長の 2.6% の 1.22 時間、最悪で 22% の約 10.6 時間の映像を閲覧すれば網羅的に関連映像を閲覧できる。そのなかから更に、利用者の関心がある話題の流れに沿ったトピックスレッドを選択できれば、平均で総映像長の 1.6% の 46 分、最悪で 11% の約 5.3 時間を閲覧すれば、関連映像の内容を把握できる、とされている。しかし、実際には平均 46 分も映像を視聴するのは大変である。

一方で、近年、ソーシャルネットワーキングサービス (SNS) が爆発的に普及している。SNS では利用者間の社会的な繋がりを支援する機能が提供され、多くの利用者はその機能を利用し

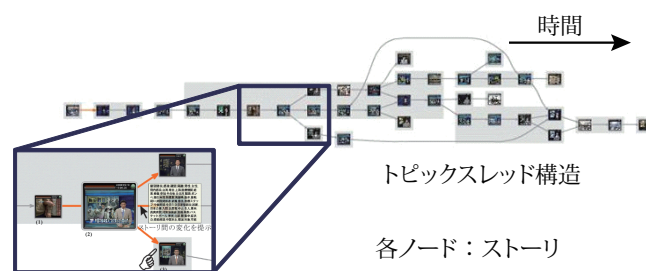


図 1 トピックスレッド構造に基づくニュース映像閲覧インタフェース: mediaWalker

て、特に自分の興味・関心がある情報の受発信を行っている [7]。放送映像を視聴しながら、いわゆる「セカンドスクリーン」で SNS 上に投稿するユーザも増えている。そこで、放送映像中のある対象に関する SNS 上のコメントや評価（以下、これらをまとめて SNS レスポンスと呼ぶ）を解析することによって、その対象への興味・関心の度合いや意見の種類について、利用者視点で抽出できると考えられる。そこで本報告では、SNS における利用者の反応を用いた単語スコアに基づく複数ニュース映像の要約方法を提案する。

SNS を利用した映像要約手法として、Kobayashi ら [8] や Doman ら [9] の手法が挙げられる。Kobayashi らの手法では、マイクロブログサービス Twitter における投稿から投稿者が応援するチームに関する属性を判定し、同一チームを応援する視聴者の実況書き込みからスポーツ映像の自動要約を行った。Doman らの手法では、スポーツ中継に対する Twitter への投稿、リツイート、連続文字の数を用いることによって言語によらないスポーツ映像の自動要約を行った。これらの研究では、いずれも投稿数がある程度期待できるスポーツ映像が対象であり、投稿数が少ないニュース映像の要約には有効に適用できないと考えられる。

2. SNS 利用者の反応を用いた単語スコアに基づく複数ニュース映像の要約

SNS 利用者の反応に基づくストーリー系列の選択手法について、その処理の流れを図 2 に示す。以下、各処理について説明する。

2.1 トピックスレッド構造の構築 [5]

アーカイブ中の指定されたストーリーを起点とするトピックスレッド構造を構築する。トピックスレッド構造上で隣接しているストーリー同士は、意味的に関連が強く、時系列上の前後関係が維持されるという性質を持っている。

トピックスレッド構造はストーリー分割とトピックスレッド構造の構築という 2 段階の処理により構築する。

2.1.1 ストーリーの分割

- (1) 音声書き下しテキストである文字放送字幕（以下、CC と呼ぶ）の各文を形態素解析し、名詞列を抽出する。
- (2) 名詞列の語義属性（一般、人物、場所・組織、時相）を解析し、語義属性ごとに 4 つのキーワードベクトルを作成する。

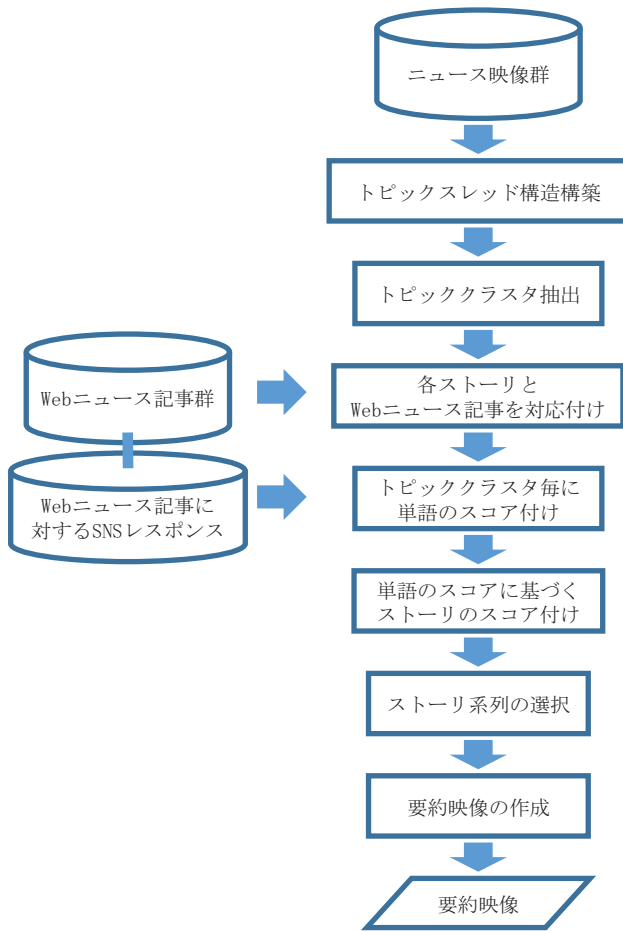


図2 提案手法における処理の流れ

(3) 連続する2文が同一ストーリーに属するかを判定するため、語義属性別に両文の前後 w 文を結合したキーワードベクトル間の距離を余弦尺度で評価する。ここで、 w を1文から10文まで変え、最後に各 w における距離の最大値を類似度とする。

(4) 各語義属性別の類似度を重み付き和の形で統合し、しきい値を下回る場合にストーリー境界が存在すると判定する。

(5) 分割された各ストーリーに対して再度キーワードベクトルを作り、隣接するストーリー間の類似度がしきい値を上回れば再結合する。

2.1.2 トピックスレッド構造の構築

(1) ストーリー間のキーワードベクトルの類似度を評価し、しきい値を上回れば強く関連しているとみなす。

(2) 次の条件を満たしながら、各ストーリーを根とする単純なストーリー木を展開する。

- 子供は親と強く関連し、必ず親よりも新しい。
- 兄弟は必ず年少の方が新しい。

(3) 次に、 T_r 中の全ての部分木 $T_S(i)$ に対して、年長の節点に等価な部分木 $T_S(j)$ が存在するとき、次の操作を施す。

- 除去: $T_S(j)$ が $T_S(i)$ の兄の子孫であるとき、 $T_S(i)$ を除去。
- 統合: $T_S(j)$ が $T_S(i)$ の先祖(親を除く)であるとき、 $T_S(i)$ を $T_S(j)$ と統合。

2.2 トピッククラスタの抽出 [5]

トピックスレッド構造から、トピックのまとめりであるトピッククラスタを抽出する。以下、その方法を説明する。

(1) 起点ストーリーをクラスタ開始ノード ($C_0 = S_0$)、かつ注目ストーリー ($S = S_0$) に指定。

(2) 注目ストーリー S の子を $S_c(j) (j = 1, \dots, J)$ とする。このときストーリー C_0 と $S_c(j)$ のキーワードベクトルの類似度がしきい値以下ならば、 $S_c(j)$ を新たなクラスタ開始ノードに設定 ($C_0 = S_c(j)$)。

(3) 全ての葉ノードに到達するまでトピックスレッド構造を走査して注目ストーリー S をずらしていき、手順(2)を再帰的に適用。

(4) 最後に、各クラスタ開始ノードからトピックスレッド構造をたどって次のクラスタ開始ノードまでに含まれる全てのストーリーを同一トピッククラスタに属するとする。

2.3 ストーリーと Web ニュース記事の対応付け [10]

ニュース映像に直接 SNS 利用者の反応を対応付けるのは難しい。そこで、Web ニュース記事とニュース映像を対応付け、Web ニュース記事に対する SNS レスポンスをそのニュース映像に対する反応とみなす。そのために、トピックスレッド構造に含まれる全てのストーリーについて、内容がほぼ同じ Web ニュース記事を対応付ける。

まず、Web ニュース記事の本文を抽出する。この本文とストーリーに付属する CC を入力文書として、Web ニュース記事とストーリーが対応付いているか否かを判定する。以下にこの手順を示す。

(1) 各入力文書を形態素解析する。

(2) 名詞と未知語をキーワードとして抽出し、出現頻度ベクトルを作成する。

(3) 両者の出現頻度ベクトルの余弦距離を算出する。

(4) 余弦距離がしきい値以上ならば両入力文書を対応付ける。

2.4 単語のスコア付け

トピッククラスタ毎に、ストーリーに付属する CC に含まれる全単語をスコア付ける。このとき、視聴者が注目した可能性がある単語を重要視するために、SNS レスポンスの数をを用いる。

具体的には、トピッククラスタについて以下の手順で単語にスコア付ける。

(1) トピッククラスタに含まれる全てのストーリーに付属する CC を形態素解析する。

(2) 名詞と未知語をキーワードとして抽出し、各ストーリーにおいてキーワードの tf-idf ベクトルを作成する。

(3) Web ニュース記事に対する SNS レスポンスの数を N とするとき、 i 番目のストーリーにおける単語 j の単語スコア $W_{\text{story}}(i, j)$ を以下のように定義する。

$$W_{\text{story}}(i, j) = \begin{cases} W_{\text{CC}}(i, j)N & (W_{\text{article}}(i, j) \neq 0) \\ W_{\text{CC}}(i, j) & (\text{otherwise}) \end{cases} \quad (1)$$

ここで、 $W_{\text{CC}}(i, j)$ は i 番目のストーリー中の単語 j に関する tf-idf 値を、 $W_{\text{article}}(i, j)$ は i 番目のストーリーに対応づいた Web

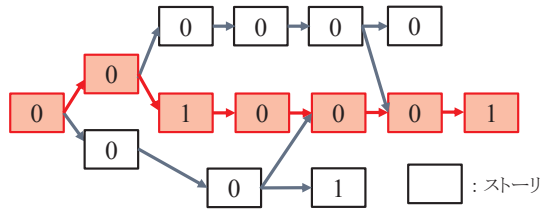


図3 ストーリー系列選択の様子

ニュース記事中の単語 j の頻度を表す。

また、 k 番目のトピッククラスタにおける単語 j のスコア $W_{\text{cluster}}(k, j)$ を以下のように定義する。

$$W_{\text{cluster}}(k, j) = \sum_{i \in C_k} W_{\text{story}}(i, j) \quad (2)$$

ここで、 C_k を k 番目のトピッククラスタに属するストーリー番号の集合とする。

2.5 ストーリーのスコア付け

スコアが高い単語が多く含まれていれば、そのストーリーには視聴者が注目する対象が多く含まれていると考えられるため、単語スコアを基にストーリーにスコア付けする。

まず、句点を手掛かりに、ストーリーに付属する CC を文に分割する。次に、長い文のスコアが高くなるように、文に含まれる単語のスコアの平均をその文のスコア（文スコア）とする。最後に、文を多く含むストーリーのスコアが高くなるように、ストーリーに含まれる文スコアの平均をそのストーリーのスコア（ストーリースコア）とする。

2.6 ストーリー系列の選択

トピックスレッドに含まれるストーリースコアの平均が最も高いものを選択すると、短いトピックスレッドが選択される傾向がある。そこで、ストーリースコアを 2 値化したものを新たなストーリースコアとし、トピックスレッドに含まれる 2 値化したストーリースコアの和が最大となるものを選択する。その様子を図 3 に示す。

2.7 要約映像の作成

通常、ニュース番組では、アナウンサーが説明する内容に合わせた映像が用いられるため、CC 中の単語に対応する映像にはその単語に関連する物体や状況が映っていると仮定する。

以下にストーリー系列中の各文から、ストーリー中の文スコアを基に要約映像を作成する手法について説明する。

(1) 各クラスタに含まれる各文について、文スコアが高い順に並べ替える。

(2) 文に対応する映像長の合計が t 秒以内となるように文を選択する。まず、 k 番目のクラスタの最大映像長 t_k を以下のように決定する。

$$t_k = t \frac{|C_k|}{\sum_i |C_i|} \quad (3)$$

各クラスタにおいてクラスタの最大映像長を超えないように文スコアが高いものから選択する。このとき、「次は」、「続いて」、「まず」から始まる文と「か」で終わる文は候補から除外する。

(3) 各クラスタで選択された文を時系列に並べ替える。



(a) クラスタ番号を表示したフレーム (b) 字幕スーパを切り出した画像

図4 クラスタの変わり目に挿入されるクラスタ番号を表示したフレームと、ストーリーの変わり目に挿入される字幕スーパを切り出した画像の例

表1 実験に用いたトピックスレッド構造の起点ストーリーとその内容

構造	起点ストーリー	起点ストーリーの内容
1	2014年 2月 21日	ウクライナで警察とデモ隊が衝突
2	2014年 5月 15日	集団的自衛権に関する与党協議
3	2014年 5月 22日	北朝鮮拉致問題に対する制裁措置の検討
4	2014年 9月 12日	スペインのカタルーニャ州やイギリスのスコットランドで独立の声

(4) 各文に対応する映像を切り出し、文の順序で全て結合することで、1つの映像とする。このとき、クラスタの変わり目では、クラスタ番号を表示したフレームを挿入し、ストーリーの変わり目では、ストーリーの概要を示す字幕スーパを切り出した画像と、そのテキストを人工的に読み上げた音声を挿入する。これらの画像の例を図 4 に示す。

3. 実験

3.1 データセット

本実験では、入力映像として NHK ニュース 7 (毎晩 19:00～19:30 放送) を使用し、事前に構築したトピックスレッド構造のなかから 4 個を実験に用いた。表 1 に 4 個のトピックスレッド構造の起点ストーリーの内容を示し、表 2 に実験に使用したトピックスレッド構造の規模を示す。また、ストーリーに対応付ける Web ニュース記事として「NHK NEWSWEB」[11]、SNS レスポンスとしてソーシャルネットワーキングサービス Facebook [12] の「いいね!」、「シェア」、「コメント」の数を用いた。

ストーリー系列の選択の際に用いたしきい値は、トピックスレッド構造に含まれる全ストーリースコアの平均と標準偏差の和とした。

以下、表 1 に示した 4 つの起点ストーリーを基に生成されたトピックスレッド構造またはそれに含まれる映像セットをデータ 1～データ 4 と呼ぶ。

3.2 ストーリー系列の選択方法に関する評価

本節では、要約映像を作成する際の映像候補となるストーリー系列の選択方法に関する評価実験について説明する。

3.2.1 ストーリー系列の作成

実験に用いたストーリー系列の作成方法を以下に示す。

- 提案手法: tf-idf と SNS レスポンスを用いた単語スコア

表 2 実験に用いたトピックスレッド構造の規模

	構造 1	構造 2	構造 3	構造 4
スレッド数	67	383	223	4
ストーリー数	40	50	41	8
クラス数	15	15	11	1
1 ストーリーあたりに対応付いた Web ニュース記事の平均	3.4	2.9	2.8	8.0

表 3 基準 1 に対する各手法の得票結果

手法	データ 1	データ 2	データ 3	データ 4	平均
提案手法	1	5	7	6	4.75
比較手法 1	2	2		1	3.00
比較手法 2	7	3	3	3	4.00

表 4 基準 2 に対する各手法の得票結果

手法	データ 1	データ 2	データ 3	データ 4	平均
提案手法	4	9	8	6	6.75
比較手法 1	4	1		4	4.25
比較手法 2	2	0	2	0	1.00

に基づいて選択されたトピックスレッド

- 比較手法 1 : tf-idf のみを用いた単語スコアに基づいて選択されたトピックスレッド

- 比較手法 2 : 起点ストーリーとの類似度がしきい値以上のストーリーを単純に時系列に並べたもの

3.2.2 被験者実験

20 代の男女 17 名による被験者実験により評価を行った。実験に用いる 4 つのデータそれぞれに対して 10 名が評価するようにした。実験方法の詳細を以下に示す。被験者に以下の 2 つの基準に基づいて、より良いと考えるストーリー系列を選択してもらった。

- 基準 1 : ストーリーを 1 つずつ時系列に見たときに、ストーリー間の繋がりがよいもの
- 基準 2 : 時系列に見たときに全体としてトピックの流れがよいもの。

また、ストーリー系列を見てもらい、一連の内容を理解するうえで必要不可欠でないとされるストーリーを別途挙げてもらった。

3.2.3 実験結果と考察

被験者実験の基準 1 に対する結果を表 3、基準 2 に対する結果を表 4 に示す。なお、各々について、データ 3 においては、提案手法と比較手法 1 により選ばれたストーリー系列が同じであったため、まとめて示した。

表 3 より、基準 1 では 4 つのデータのうち 3 つ、また表 4 より、基準 2 では 4 つのデータ全てにおいて、提案手法が最多得票であった。

各データにおいて、ストーリー系列に含まれる全ストーリー数に対して、被験者が内容把握に必要不可欠でないと挙げてストーリー数の平均の割合を表 5 に示す。基準 1 の得票数の順位は、概ね各データにおける必要不可欠でないとされたストーリーの割合と同じ傾向を示している。

基準 1 では、データ 1 において提案手法と比較手法 1 の得票

表 5 必要不可欠でないとされたストーリーの割合

手法	データ 1	データ 2	データ 3	データ 4	平均
提案手法	0.360	0.177	0.167	0.120	0.206
比較手法 1	0.350	0.208		0.160	0.221
比較手法 2	0.225	0.167	0.297	0.214	0.225

が少なかった。これは、必要不可欠でないストーリーの割合が比較手法 2 よりも多かったことが原因と考えられる。また、提案手法において、被験者全員が必要不可欠でないストーリーとして挙げたものがあつた。それは 2014 年 5 月 14 日のストーリーで、内容は「宇宙飛行士の若田光一さんが地球に帰還した」という内容であった。このストーリーの前のストーリーは「親ロシア派とウクライナ軍の双方の動きが活発化した」という内容、後のストーリーは「アメリカの国防長官が地对空ミサイルはロシアが提供したと発言した」という内容であった。明らかに内容が異なるストーリーが系列中に混入したことによりストーリーの繋がりがよくないと判断されたと考えられる。この原因は、語義属性毎にキーワードベクトルの余弦尺度を計算した際に、重みが大きい語義属性(場所・組織)で、値が大きくなったためである。

一方、基準 2 では、4 つのデータ全てにおいて、提案手法が最多得票であったのはトピックの変遷を表すトピックスレッドを選択した効果だと考えられる。比較手法 2 は起点ストーリーに類似したストーリーを全て時系列に繋いだものなので、話題の切り替えが頻繁に発生、トピックの流れがよいという基準 2 では得票数が少なかったと考えられる。提案手法と比較手法 1 の違いは、SNS レスポンスの有無であるが、提案手法の方がよい結果であることから、視聴者の関心があるトピックの流れは細かなトピックの流れではなく、トピックの流れが全体的によりものであると考えられる。

3.3 要約映像の作成方法に関する評価

本節では、ストーリー系列から要約映像を作成する手法に関する評価実験について説明する。

3.3.1 被験者実験

被験者実験 1 と同じ 20 代の男女 17 名による被験者実験を行った。被験者には、被験者実験 1 で割り当てたものと同じデータを本実験でも割り当てた。実験方法の詳細を以下に示す。以下に示す基準で要約映像の作成手法を選択してもらった。

- 基準 : ストーリー系列を最も適切に要約していると思うもの。

実験に用いた要約映像の作成方法を以下に示す。要約映像の基となるストーリー系列は提案手法で選択されたものを用いた。

- 提案手法 : tf-idf と SNS レスポンスを用いた文スコアに基づいて選択した文に対応する映像を切り出し、繋ぎ合わせたもの。総映像長 $t = 180$ [秒] とした。

- 比較手法 1 : tf-idf のみを用いた文スコアに基づいて選択した文に対応する映像を切り出し、繋ぎ合わせたもの。総映像長 $t = 180$ [秒] とした。

- 比較手法 2 : 文スコアなどを用いず、各ストーリーのアンカショットのみを繋ぎ合わせたもの。

提案手法と比較手法 1 では、要約に用いる文の選択までは自動

表 6 比較手法 2 で作成された映像の長さ

	データ 1	データ 2	データ 3	データ 4
映像の長さ [秒]	105	335	368	117

表 7 要約映像の作成方法に関する被験者の評価

手法	データ 1	データ 2	データ 3	データ 4	平均
提案手法	3	6	3	6	4.50
比較手法 1	4	4	6	1	3.75
比較手法 2	3	0	1	3	1.75

で行い、字幕スーパの抽出、字幕スーパの読み上げ音声とクラスタ番号を表示するフレームの挿入は人手で行った。比較手法 2 では、ストーリー冒頭のアンカショットを全区間繋ぎ合わせた。

3.3.2 実験結果と考察

比較手法 2 によって作成された映像長を表 6 に、被験者による評価結果を表 7 に示す。

表 7 より、アンカショットを繋ぎ合わせただけの比較手法 2 よりも、提案手法や比較手法 1 のように文スコアが高い文を選択して、それに対応する映像を切り出し、繋ぎ合わせた要約手法の方が得票が多かった。その理由として、比較手法 2 は「映像に動きがなく、見ていて眠くなる」、「内容が冗長である」、「概要のみで詳しい内容があまり伝わらない」という意見があった。アンカショットはアナウンサーがニュース原稿を読むスタジオのショットであるので、話題に挙がっている人物や場所、状況が映像に映らない。また、アンカショットはストーリー冒頭に出現するが、ストーリー冒頭の文はリード文と呼ばれ、そのストーリー全体の要約となっていることが多い。一般にはリード文に続く文から具体的な内容の説明がなされるため、アンカショットのみを繋いだ映像では詳しい内容があまり伝わらないことが考えられる。

次に、提案手法と比較手法 1 では、平均して提案手法の方が得票率が高かった。被験者に回答してもらった、手法を選んだ理由あるいは選ばなかった理由は、「アナウンサーが映っている部分よりも他の場面が良い」というものがあった。

上記の理由から、アナウンサーが話している内容を映像で映しているという仮定が必ずしも正しくなかったことが分かった。実際に、スタジオでアナウンサーや有識者が地図やフリップを交えて状況を説明するシーンが要約映像に含まれていた。現在は単語スコアを基に文スコアを計算し、それに基づいて要約映像に含める文を選択しているが、文スコアだけでなく画像の特徴もスコアに含めることで、よりよい要約映像が作成できると考えられる。例えば、画像特徴として画像に映っているコンセプト（人物、場所、自動車、動物など）を検出し、文に含まれる単語と対応するコンセプトが含まれる確率を画像特徴のスコアとして組み合わせることが考えられる。

4. む す び

本報告では、SNS 利用者の反応を用いた単語スコアに基づいて複数ニュース映像の要約手法を提案した。提案手法では、トピックスレッド構造上から単語のスコアをもとにトピックス

スレッドを選択し、そのトピックスレッドに含まれるスコアが高い文に対応する映像を結合することで要約映像を作成した。

今後は、選択されたストーリー系列から映像を切り出す際に、画像に映っているコンセプトとの対応を考慮して映像を選択することが考えられる。

謝 辞

本研究の一部は国立情報学研究所との共同研究及び科研費による。

文 献

- [1] 井手一郎, “ニュース映像の検索”, 映像情報メディア学誌, Vol.64, No.3, pp.306–311, Mar. 2010.
- [2] Pinar Duygulu, Jia-Yu Pan, and David A. Forsyth, “Towards auto-documentary: Tracking the evolution of news stories”, Proc. 12th ACM Int. Conf. on Multimedia, pp.820–827, Oct. 2004.
- [3] Jonathan G. Fiscus and George R. Doddington, “Topic detection and tracking evaluation overview”, In James Allan, ed. Topic detection and tracking: Event-based information organization, Kluwer Academic Publishers, pp.17–32, 2002.
- [4] Xiao Wu, Chong-Wah Ngo, and Qing Li, “Threading and auto documenting news videos”, IEEE Signal Processing Mag., Vol.23, No.2, pp.59–68, Mar. 2006.
- [5] 井手一郎, 木下智義, 高橋友和, 孟洋, 片山紀生, 村瀬洋, “大量ニュース映像を対象とした時系列意味構造に基づく情報編纂手法の提案”, 人工知能学論, Vol.23, No.5, pp.282–292, Sept. 2008.
- [6] Ichiro Ide, Tomoyoshi Kinoshita, Tomokazu Takahashi, Shin’ichi Satoh, and Hiroshi Murase, “mediaWalker: A video archive explorer based on time-series semantic structure”, Proc. 15th ACM Int. Conf. on Multimedia, pp.162–163, Sept. 2007.
- [7] 総務省, 平成 23 年版 情報通信白書, <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h23/html/nc213120.html> [2014/2/7 アクセス].
- [8] Takashi Kobayashi, Tomokazu Takahashi, Daisuke Deguchi, Ichiro Ide, and Hiroshi Murase, “Detection of biased broadcast sports video highlights by attribute-based tweets analysis”, Proc. 19th Int. Conf. on Multimedia Modeling, Lecture Note in Computer Science, Vol.7733, pp.364–373, Jan. 2013.
- [9] Keisuke Doman, Taishi Tomita, Ichiro Ide, Daisuke Deguchi, and Hiroshi Murase, “Event detection based on Twitter enthusiasm degree for generating a sports highlight video”, Proc. 22nd ACM Int. Conf. on Multimedia, pp. 949–952, Nov. 2014.
- [10] 加藤光佑, 井手一郎, 出口大輔, 村瀬洋, “SNS を用いた映像構造化のためのニュースストーリーと Web ニュース記事の対応付け手法”, 信学技報, MVE2013-112, Mar. 2014.
- [11] NHK NEWSWEB, <http://www3.nhk.or.jp/news/> [2015/2/8 アクセス].
- [12] Facebook, <https://www.facebook.com/> [2015/2/8 アクセス].