

ニュース映像要約のための ニュース内容に沿ったシーンの選択に関する検討

張 よう[†] 谷繁龍之介[†] 道満 恵介^{††,†} 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
^{††} 中京大学 工学部, 〒 470-0393 愛知県豊田市貝津町床立 101
^{†††} 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: †zhangy@murase.m.is.nagoya-u.ac.jp, †{ide,murase}@is.nagoya-u.ac.jp

あらまし 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。しかし、資料として活用する際、あるトピックに関する映像を全て閲覧するには時間を要するため、ニュース映像を要約する研究が必要である。本報告では、その基礎技術として、ニュース内容に沿ったシーンの選択手法について検討する。本手法では、ニュース映像の画像内容を分析することで、ニュース内容に沿ったシーンを選択する。実際のニュース映像に対する評価実験を行い、本手法によりニュース内容と関連性が高いシーンを選択できることを確認した。

キーワード ニュース映像, 要約, Visual Concept, シーン

A study on selecting scenes that match the news contents for news video summarization

Ye ZHANG[†], Ryunosuke TANISHIGE[†], Keisuke DOMAN^{††,†}, Ichiro IDE[†],
Daisuke DEGUCHI^{†††,†}, and Hiroshi MURASE[†]

[†] Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan
^{††} School of Engineering, Chukyo University
Tokodachi, Kaizu-cho, Toyota-shi, Aichi, 470-0393 Japan
^{†††} Information and Communications Headquarters, Nagoya University
Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan
E-mail: †zhangy@murase.m.is.nagoya-u.ac.jp, †{ide,murase}@is.nagoya-u.ac.jp

Abstract Since news videos record daily events, they are valuable sources of real world information. However, when using them as information sources, it is difficult to watch all of the news videos about a topic reported during a long period. Therefore, it is necessary to develop a method for summarizing news videos. Thus, this report studies a method for selecting scenes that match the news contents for news video summarization. The method selects scenes related to news contents by analyzing image features to understand the visual concepts of an image. We confirmed that it was possible to select scenes related to news contents with this method.

Key words News video, summary, Visual Concept, scenes

1. はじめに

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。しかし、ニュース映像を資料として活用

する場合、あるトピックに関する長年に渡る映像を全て閲覧することは時間的に困難である。そこで、長いニュース映像を短く要約する技術への需要がある。

一方、画像処理の分野で、Visual Concept という概念が提

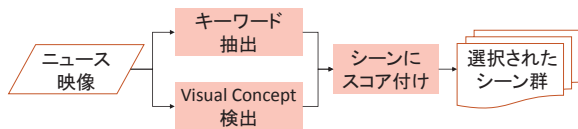


図1 提案手法の処理の流れ

案されている [1]. Visual Concept は、人物や物体、自然など、画像の意味的内容を表すものである。我々は、Visual Concept を用いてニュース映像の画像の意味的内容を理解することで、ニュース内容と関連性が高いシーンを選択することが可能であると考えた。

本報告では、ニュース映像を自動要約するための基礎技術として、ニュース内容に沿ったシーンの選択手法について検討する。

2. ニュース内容に沿ったシーンの選択

提案手法の処理の流れを図1に示す。ニュース映像を入力とし、それに付属する CC (Closed-Caption; 文字放送字幕) からキーワードを抽出する。また、映像からは Visual Concept を検出し、抽出したキーワードとの対応からシーンのスコア付けを行う。以下、順に説明する。

2.1 キーワード抽出

ニュース映像に付属する CC を形態素解析し、品詞判別を行い、一般名詞をキーワードとして抽出する。そして、tf-idf により各キーワードの重要度を算出する。

2.2 Visual Concept 検出

入力のニュース映像をシーンに分割する。本研究では、連続した2枚のフレームの HSV ヒストグラムの差分を用いて、シーンの開始フレームを検出する。このフレームから画像特徴を抽出し、複数の Visual Concept の識別器を用いて、Visual Concept の検出を行う。画像特徴は500次元の SIFT 特徴の Bag of Features 表現を用いる。

2.3 シーンのスコア付け

抽出したキーワードと検出した Visual Concept の結果を用いて、関連度を算出する。算出した関連度により各シーンにスコアを付ける。以下に手順を示す。

- (1) 日本語 WordNet [2] から Visual Concept の同義語を取得し、同義語リストを作成する。
- (2) 代表フレームから検出した全ての Visual Concept またはその同義語に対応するキーワードの重要度の総和を、該当シーンのスコアとする。
- (3) スコアが上位のシーンを選択する。

3. 実験

Visual Concept 検出を用いたシーンの選択方法の有効性を評価するために実験を行った。提案手法で用いる Visual Concept には、10年間の NHK News7 中の出現頻度が高い単語の調査結果を参考に57種類を用意した。これらの Visual Concept を、物体、場所、自然、人、活動の5種類に大別した。種類別の識別器の平均精度を表1に示す。

シーン選択の評価実験には、NHK News7 の映像3本を用い

表1 Visual Concept 識別器の精度 (再現率)

	Visual Concept の種類					
	物体	場所	自然	人	活動	平均
再現率	0.87	0.77	0.93	0.85	0.93	0.81

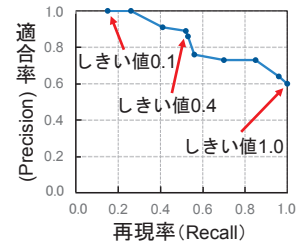


図2 シーン選択の結果

ロシアと中東諸国を訪問していた安倍総理大臣。一連の日程を終えて帰国しました。政府専用機で羽田空港に到着した安倍総理大臣。一連の首脳外交のうち、ロシアのプーチン大統領との首脳会談では...

図3 実験に用いた映像の CC の例



(a) スコアが高い例 (b) スコアが低い例

図4 図3のCCに対するスコアの例

た。選択するシーンの割合のしきい値を変化させたときの適合率と再現率の関係を図2に示す。本研究はニュース映像の自動要約を目的としているため、適合率が重要である。実験結果より、しきい値は0.4前後(スコアが上位40%のシーンを選択)が良いと考えられる。図3は本実験に用いた映像のCCの例である。図4は、図3のCCに対して各シーンにスコア付けを行った結果例である。この結果から、提案手法によりニュース内容に沿ったシーンを選択することが可能であると考えられる。

4. まとめ

本報告では、ニュース映像の自動要約を目的として、ニュース内容に沿ったシーンの選択手法について検討した。今後の課題として、音声情報を考慮した要約手法の検討等が挙げられる。

謝辞

本研究の一部は科研費及び国立情報学研究所との共同研究による。

文献

- [1] Yu-Gang Jiang, Chong-Wah Ngo and Jun Yang, "Towards optimal bag-of-features for object categorization and semantic video retrieval," Proc. 6th ACM Int. Conf. on Image and Video Retrieval, pp.494-501, July 2007.
- [2] Francis Bond, Timothy Baldwin, Richard Fothergill and Kiyotaka Uchimoto, "Japanese SemCor: A sense-tagged corpus of Japanese," Proc. 6th Int. Conf. of the Global WordNet Association, pp.9-16, Jan. 2012.