

Web コンテンツを用いた訃報映像の自動編集へ向けた 画像選択方法の検討

國代 京花[†] ナック フランク^{††} 井手 一郎[†] 川西 康友[†] 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

^{††} アムステルダム大学 情報学研究所 〒1098 XH オランダ アムステルダム市 サイエンスパーク 904

^{†††} 名古屋大学情報戦略室 〒464-8601 愛知県名古屋市千種区不老町

あらまし 本報告ではWeb コンテンツを用いて訃報を伝える映像（訃報映像）を自動編集する際の画像選択方法を検討する。訃報映像は、幼少期、功績、私生活の3部で編成する。各部に関する重要な事象を表すキーワードをオンライン百科事典 Wikipedia から抽出し、故人に関連する重要な事象に対応する画像を画像検索エンジンを用いて取得する。取得した画像について、その内容を表す Visual Concept に着目し、キーワードに関する一般的な Visual Concept と一致度が高いものを選ぶことで、故人に関する重要な事象を視覚的によく表現した画像を選択する。

キーワード 映像編集, 人物情報, 画像検索

Study on an image selection method towards automatic authoring of video obituary using Web contents

Kyoka KUNISHIRO[†], Frank NACK^{††}, Ichiro IDE[†], Yasutomo KAWANISHI[†],
Daisuke DEGUCHI^{†††}, and Hiroshi MURASE[†]

[†] Grad. School of Information Science, Nagoya Univ. Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan

^{††} Informatics Institute, Univ. of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands

^{†††} Information Strategy Office, Nagoya Univ. Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan

Abstract In this report, we propose an image selection method for automatic authoring of video obituary using Web contents. An obituary video is composed of three parts: childhood, profession, and personal life. We extract keyphrases on important events for each of these parts from the online encyclopedia Wikipedia, and then retrieve images from an image search engine that correspond with the important events of the deceased. Focusing on the Visual Concepts of the obtained images, by selecting images that share the same Visual Concepts that represent the keyphrases in general, we select images that visually represent the important events of the deceased well.

Key words Video authoring, Personal information, Image retrieval

1. はじめに

著名人が死亡すると新聞記事やニュース映像など様々な媒体を通して訃報が伝えられる。その多くは、その人物の生前の功績や様子と共に伝えられる。本研究では訃報を伝える映像を「訃報映像」と呼ぶ。訃報映像は視聴者に対し、故人に関する視聴覚的な情報を伝えられる。

近年、インターネット上の様々な媒体から情報を得ることができる。文字媒体ならば Wikipedia のようなオンライン百科事典から、映像や音声ならば Youtube などの映像共有サービスや

画像検索エンジンから必要な情報を容易に得られる。これらの Web コンテンツは、一般ユーザが盛んに投稿することで、莫大な量の情報が日々寄せられているため、著名な人物が関わった出来事について、広範かつ最新の情報が得られる。しかし、大量の情報からある人物に関係する出来事を網羅的に抽出して編集するのは大変な作業である。そこで我々は、Web 上のコンテンツから故人に関連する出来事を抽出し、訃報映像として必要な内容のものを選別することで、自動編集を行うことを目指している。本報告は、そのために必要な画像の選択方法について検討する。

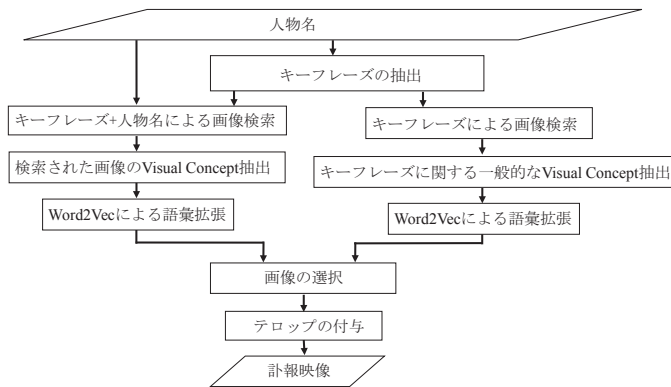


図1 訃報映像生成の処理手順

2. 提案手法

提案手法の処理手順を図1に示す。訃報映像は故人の人生を要約したものであり、幼少期、功績、私生活の3部で編成する。始めに故人に関する特徴的な事象を表現したキーワードをオンライン百科事典 Wikipedia^(注1) とその内容を構造化したデータベース DBpedia [1]^(注2) から抽出する。次に、このキーワードをもとに画像検索をする。一般に画像検索で画像を収集する場合、上位の画像ほどクエリと一致していると考えられる [2] が、クエリと視覚的な内容が一致しない画像が現れることもある。そこで Visual Concept [3] を用いて画像の視覚的な内容を分析し、キーワードの内容と最も一致するものを選択する。最後に、取得した画像に対応するキーワードをテロップとして付与し、スライドショーの形式で映像化する。

2.1 キーワードの抽出

DBpedia は Wikipedia よりも情報を抽出しやすいため、これから生年月日、職業、出生地など定型的な情報を抽出する。しかし、功績や私生活に関する情報の分類や構造化は不完全であるためそれらについては、Wikipedia から直接抽出する。

Wikipedia では、執筆者は強調したい語句をアポストロフィなどを用いた Wiki 固有の記法により強調する。また文章中から張られるリンクのアンカテキストは固有名詞や重要な語句であることが多い。そこでこれらの語句をキーワードとして抽出する。一般に Wikipedia 冒頭の概要では、その人物の特徴が簡潔にまとめられているため、概要内のキーワードを利用する。しかし、概要には功績と私生活の内容が混在して記述されているため、それぞれに該当するキーワードを分類する必要がある。そこで、Wikipedia 記事中の見出しに対してプレートマッチングを適用する。このとき、予め功績と私生活に関連する語句を含む表現をプレートとしてそれぞれ用意しておき、それらと一致した見出しがあれば、その内容は功績と私生活のいずれかに関係する内容として分類する。分類されたそれぞれの内容の中に表れるキーワードと、概要文章内のキーワードとを比較し、一致するものを取得することで功績と私生活のキーワードを分類する。

2.2 画像の選択

キーワードをクエリとした画像検索で得られる画像の中から、その視覚的な内容がキーワードの内容と一致したものを選択したい。まずキーワードのみをクエリとして画像検索を行い、得られた画像から Visual Concept を抽出する。得



図2 実験結果：Albert Einstein より抜粋。
(上段：画像選択なし、下段：画像選択あり)

られた Visual Concept は、そのキーワードに関する一般的な Visual Concept とみなす。次に、人物名とキーワードの両方をクエリとして再び画像検索を行い、検索された画像から Visual Concept を抽出する。これを先ほど得た一般的な Visual Concept と比較し、最も一致する画像を抽出する。このようにして、故人に関するそのキーワードが表す事象の典型的な画像を選択する。ここで Visual Concept の語彙のみでは画像内容を表現しきれないため、Word2Vec [4] を用いて、各 Visual Concept の名称と共起確率が高い語句を用いて語彙を拡張する。具体的には、キーワードに関する一般的な Visual Concept の共起語句と、人物名とキーワードから得た画像の Visual Concept の共起語句のうち、一致するものがあれば、その語句の共起確率と Visual Concept の尤度を乗算して累積する。最終的に、キーワードに最も一致する画像として選択する。

3. 実験

提案手法により、著名な人物の訃報映像を生成した。ここでは、bing 画像検索エンジン^(注3) から画像を収集し、上位4枚を利用した。Word2Vec の共起語句は、共起確率が0.5以上のものを採用した。図2に Albert Einstein の訃報映像に用いた画像の一部を示す。画像選択により、より故人に関する視覚的情報が多く含まれる画像が得られたことが分かる。

4. むすび

本報告では Web 上のコンテンツを用いた訃報映像の自動生成における画像選択方法について検討した。提案手法では画像を選択する際に Visual Concept を用いたが、語彙にないものは検出できない。今後は語彙にない Visual Concept に対しても、より適した画像を選択する手法へと拡張したい。

謝辞 本研究の一部は、科学研究費補助金による。

文献

- [1] C. Bizer, “DBpedia—A crystallization point for the Web of data,” *Web Semantics: Science, Services and Agents on the World Wide Web*, 7(3), pp.154–165, Sept. 2009.
- [2] Y.-G. Jiang, C.-W. Ngo, and J. 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.
- [3] S.K. Divvala, A. Farhadi, and C. Guestrin, “Learning everything about anything: Webly-supervised visual concept learning,” *Proc. 2014 IEEE Conf. on Computer Vision and Pattern Recognition*, pp.3270–3277, June 2014.
- [4] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” *arxiv preprint arxiv:1301.3781*, Sept. 2013

(注1) : <https://en.wikipedia.org/wiki>

(注2) : <http://dbpedia.org/sparql>

(注3) : <https://www.bing.com/?scope=images>