

ニュース映像からの人物関係の抽出と索引付けへの利用

6 L-4

井手 一郎, 佐藤 真一

{ide,satoh}@nii.ac.jp

国立情報学研究所*

1 はじめに

1.1 研究の背景と目的

近年、放送される映像量の増加に伴い、映像の検索や再利用の要求が高まっており、映像の閲覧や索引付けについて盛んに研究されている。なかでも、資料的価値の高さから、ニュース映像を扱うことが注目されている。本発表では、大量に蓄積されたニュース映像から、話題中に登場する人物の人物関係に関する知識を抽出して索引付けを行い、検索や閲覧に利用する手法を提案する。

1.2 用語の定義

以下のように映像に関する用語を定義する。

- 映像： 動画像・音声・文字放送データからなる、一まとまりのマルチメディア情報。
- フレーム： 動画像の最小構成単位である静止画像。
- ショット： 画像的に連続なフレームの集合。カットとは、ショット間の不連続な点。
- シーン（話題）： 意味的に連続するショットの集合。ニュース映像においては、シーンは話題に対応。

2 大規模映像コーパス

2.1 映像自動獲得・蓄積システム

大量のニュース映像から情報抽出を行うために、映像自動獲得・蓄積システムを構築している。このシステムは、動画像・主音声と文字放送字幕（音声書き下しテキスト）[1]を自動的に獲得・蓄積し、文字放送字幕が付与されている特定のニュース番組の映像からなる大規模映像コーパス（年間 180 時間程度の規模）を構築することができる。

2.2 コーパスからの情報抽出

映像コーパスからの情報抽出に関する研究としては、佐藤らによるもの [5] がある。この研究では、映像の物理的構成や動画像の様々な画像特徴量に対する分析を行っている。しかし、本研究ではそのような映像の表面的な特徴ではなく、意味的内容に踏み込んだ情報抽出を目指す。

ここでは、ニュース映像の内容理解に重要と思われる、いわゆる 5W1H 情報のうち、特に登場人物（Who）に注目した情報抽出を行う。

3 人物関係に基づく検索・閲覧

3.1 概要

従来から、ニュース映像を対象として人物に注目する索引付け手法 [3] や、検索・閲覧手法 [4] は存在する。これらは、動画像中の人物像と人物名をいわば一次的に対応付ける手法である。

それに対して、本研究では、登場人物関係という二次的な情報と映像を対応付けることを考える。ここで人物関係とは、テレビ雑誌などに掲載されるテレビドラマの登場人物関係図のようなものである。このように関係を可視化することで、特定の人物とその周囲の人物関係をたどりながら映像内容を追跡することができ、話題内容に則した検索・閲覧が可能になると思われる。

具体的には以下の手順で人物関係を抽出する。

1. 音声書き下しテキストから人物名を検出
2. 動画像から話題境界を検出
3. 話題内の人物名の共起から人物関係を抽出

3.2 人物名の検出

テキストから人物名を検出する手法は、固有表現抽出 (Named entity task) の一部として研究されている。本研究では、比較的簡単な処理で実現できる、名詞句の末尾の名詞に着目した語義属性解析手法 [2] を採用した。以下にこの手法による解析の例を示す。

小泉首相	=	小泉+首相	人物	→	人物
小泉町	=	小泉+町	場所・組織	→	場所・組織

このような解析のために、名詞句の末尾に存在して人物を表すような名詞を集めた辞書を予め作成しておいた。

以下の手順でこの手法を用いて人物名を検出する。

1. テキストを形態素解析
日本語形態素解析システム JUMAN[6] を利用
2. 名詞列（名詞句のうち名詞が連続するもの）を検出
3. 固有名詞で始まるものを選別
4. 末尾の名詞に着目して語義属性を解析
5. 属性が人物であると判断されたものを検出

3.3 話題境界の検出

話題境界の検出は、ニュース映像の構成を利用して行う。一般にニュース映像では、話題の冒頭にタイトル字幕付きのキャスタが写っているショット（スタジオショット）が現れる。これを利用して、字幕付きのスタジオショットを画像的に検出することで、話題境界を検出する。

以下の手順で字幕付きスタジオショットを検出する。

* "Retrieval of human relationship from news video and its application to indexing"
Ichiro Ide, Shin'ichi Satoh / National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan

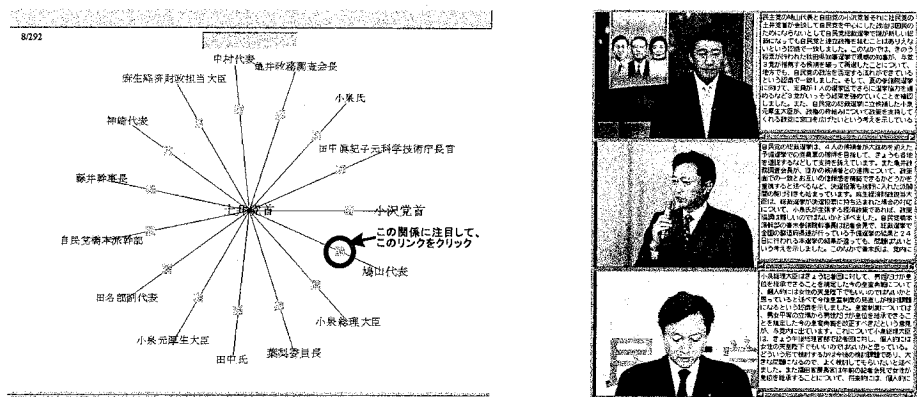


図 1: (左) 人物関係の可視化: 人物関係閲覧インターフェイス (右) 映像閲覧インターフェイス

1. 大きめの顔が一つ存在する「人物ショット」を検出
2. 人物ショットの代表フレーム(ショットの先頭フレーム)の色ヒストグラムを特徴量とみなし、特徴量空間中で最稠密なクラスをスタジオショットと判断
3. タイトル字幕を検出するために、フレーム全体のエッジ強度の変動を検出

以上の手順を 75 分間のニュース映像に適用した結果を表 1 に示す。検出洩れの原因は、話題冒頭にもかかわらず、タイトル字幕が出現しなかったためである。

表 1: 話題冒頭のスタジオショット検出実験の結果

正検出	誤検出	検出洩れ	再現率	適合率
37	0	6	86%	100%

3.4 人物関係の抽出

人物関係の抽出は、3.2 で検出した人物名の共起を、3.3 で検出した話題内で数え、全話題に対して集計して行う。

3.5 検索・閲覧インターフェイス

人物関係を可視化し、検索・閲覧のためのインターフェイスとして提供する(図 1(左))。中央に注目する人物を、円周上にその人物と関係がある人物を共起頻度順に(頻度は文字の大きさに反映)右回りに配してある。また、円周上の人物をクリックすると、注目する人物を切替えられる。

次に、特定の人物関係に基づいて映像を閲覧するためのインターフェイスを図 1(右)に示す。このインターフェイスは、図 1(左)の人物間のリンクをクリックすることで出現する。これにより、注目する人物関係がどのような関係であるか、画像と文字放送字幕の両者から理解できる。

4 実験

実際に、53 日分のニュース映像(文字放送字幕 2,807 文、話題 560 件)から人物関係を抽出した。話題境界の判定は本実験では人手で行ったが、3.3 で示したようにほ

ぼ完全に自動化できる。その結果、のべ 2,799 件の人物の共起、異なり数 292 件の人物関係が得られた。

得られた人物関係を見ると、同一人物に対する異表記も人物関係とみなされていた。異表記には同義語によるものと意味的に同義であるものがあるが、これらの名寄せを行うためには膨大な背景知識が必要となるため、このように人物関係の一種として等価(is-a)な関係とみなすことで、検索の際の問合せの表記的・意味的多様性を吸収することができる利点があると考えられる。

5 おわりに

本発表では、大規模ニュース映像コーパスから人物関係を抽出し、それを索引とした映像検索・閲覧手法を提案し、実装した結果を報告した。

今後は、より大規模な映像コーパスを対象とし、さらに画像中の人物の顔の利用方法の検討を進め、より映像内容に則した検索・閲覧手法の実現を目指す。また、人物関係が共起した話題の内容理解を行い、人物関係の種類の判別と可視化についても検討していく。

参考文献

- [1] 安藤, 今井, 小林, 本間, 後藤, 清山, 三島, 小早川, 佐藤, 尾上, 世木, 今井, 松井, 中村, 田中, 都木, 宮坂, 磯野: “音声認識を利用した放送用ニュース字幕製作システム”, 信学論, vol.J84-D-II, no.6, pp.877-887 (2001).
- [2] 井手, 田中: “末尾の名詞に着目したテレビニュース字幕の語義解析”, 情処学論, vol.39, no.8, pp.2543-2546 (1998).
- [3] 井手, 山本, 浜田, 田中: “ショット分類に基づく映像への自動的索引付け手法”, 信学論, vol.J82-D-II, no.10, pp.1543-1551 (1999).
- [4] Satoh, Nakamura, and Kanade: “Name-It: Naming and detecting faces in news videos”, *IEEE Multimedia*, vol.6, no.1, pp.22-35 (1999).
- [5] 佐藤, 児島, 阿久津, 外村: “映像コーパスの構築と分析”, 信学論, vol.J82-D-II, no.10, pp.1552-1560 (1999).
- [6] 京都大学大学院情報学研究所知能情報学専攻言語メディア研究室: “日本語形態素解析システム JUMAN 第 3.6 版”.