

人 関係に基づくニュース映 の検索と閲

井 手 一 郎 佐 藤 真 一

国立情報学研究所

〒 101-8430 東京都千代田区一ツ橋 2-1-2

TEL: (03)4212-2585 FAX: (03)3556-1916
{ide,satoh}@nii.ac.jp

従来のニュース映 の は、付与された 引 を し、 を閲覧するだけのものが 流であった。今は、より な映 の利用を するために、量のニュース映 から される 化された に づく と閲覧を えて いる。本発表では、 量のニュース映 の話 に する 物名の に 目することにより 物 に する を し、それに づいて映 の と閲覧を なう 法を 案する。

ニュース映 , , 閲覧, 物 , キャプション, 物名

Retrieval and browsing of news video by human relationship

ICHIRO IDE

SHIN'ICHI SATOH

NATIONAL INSTITUTE OF INFORMATICS

2-1-2 HITOTSUBASHI, CHIYODA-KU, TOKYO 101-8430, JAPAN

TEL: +81-3-4212-2585 FAX: +81-3-3556-1916

{ide,satoh}@nii.ac.jp

Abstract Most conventional news video retrieval and browsing was based on simple keyword retrieval and browsing of the retrieved segment. We are considering a retrieval and browsing method that utilizes structured information derived from the target video contents. In this presentation, we will introduce a retrieval and browsing method based on human relationship that is derived from cooccurrence of human names in large news video corpus.

key words news video, retrieval, browsing, human relationship, caption, human name

1 はじめに

1.1 研究の背景と目的

近年，放 される映 量が 加するに伴い，放 された映 の検 や 利用への要求が まっており，映 の閲覧や 引付けに関する研究が んに われている。な かでも，資料的価値が いことから，ニュース映 を扱うことが注目されている。

本 では， 量に されたニュース映 の検 と閲覧を うために，映 コーパスから抽出された 報を用いる手法を 案する。具 的には，ニュース映 コーパス の音 の書き下しテキストから 物名の共起 報を集計し，得られた共起関係を 物関係とみなして可 化し， 果的な検 や閲覧を実現することを目 す。

1.2 用語の定義

以下のように映 に関連する用語を 義する：

- 映像

映 とは，動画 ・音 ・文字放 データからなる，一まとまりのマルチメディア 報のこととする。

- フレーム

フレームとは，動画 の 小 単位である
画 のこと。NTSC放 方式では，動画 は毎秒
約 30 フレームで される。

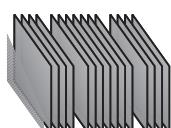
- ショット・カット

ショットとは，画 的に連 なフレームの集 のこ
と。カットとは，ショット間の不連 な点のこと。

- シーン（話題）

シーンとは，意味的に連 するショットの集 の
こと。ニュース映 においては，シーンは話題に
応する。

これに関連して， 1に映 の物理的 を示す。

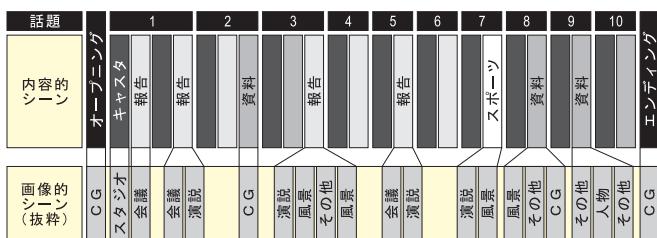


的 や動画 の様々な画 特徴量に する分 を つ
ている。しかし、本研究ではそのような映 から得られ
る表面的な特徴を分 にするのではなく、映 の意
味的内容に踏み んだ 報抽出を目 す。こ

5. が 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、ニュース映像の画的特徴を利用して行う。一般に、ニュース映像は5に示すような構造をとり、話題の冒頭にはキャスターが登場する「スタジオショット」が出現する。ここでは、この特徴を利用して、スタジオショットを画的に検出することにより、話題境界の検出を行う。



5: ニュース映像の例

スタジオショットの特徴は、以下の通りである。

- きめの顔領域（キャスターの顔）が一つ
- 背景領域（スタジオ）の画的特徴の変動は少ない

このような特徴をもとに、以下の手順でスタジオショットの検出を行う。

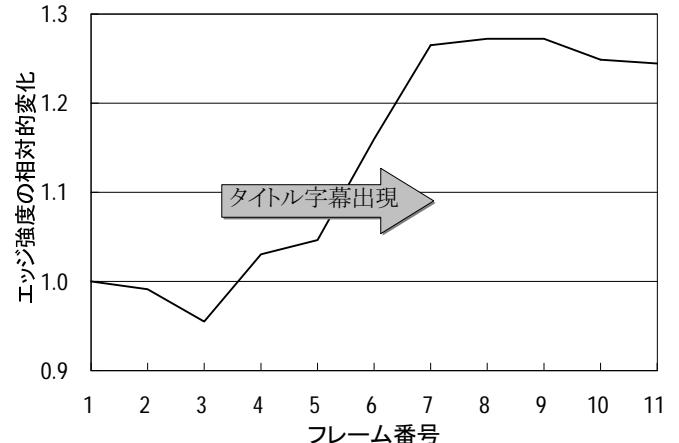
1. きめの顔領域が一つあるショット（人物ショット）を検出
2. 人物ショットの表フレーム（ショットの頭フレームを用）のヒストグラムを特微量ベクトルとみなし、特微量ベクトル空間でもっとも稠密なクラスタを形成するものをスタジオショットと判断

しかし、このようにして検出したスタジオショットは、話題の途切れに出現することもある。そこで、話題の冒頭のスタジオショットのもう一つの特徴として、以下の特徴を加味する。

- タイトル字幕が出現

タイトル字幕の出現を検出するために、さらに以下の手順を付け加える。

3. フレームのエッジ検出を行い、エッジ強度の変動を検出



6: タイトル字幕出現に伴うエッジ強度の相対的变化

これは、タイトル字幕の出現後で、エッジ強度が6に示すように急激に変化することを利用している。

以上の手順に従い、75分間（計468ショット）のニュース映像にて検出実験を行った[3]。この実験で、顔領域検出には修正HSI表系[4]を利用した顔領域抽出とテンプレートマッチングを利用した手法を用いた。また、カット検出はしくわれたものとした。その結果、表1に示すように良好な結果が得られた。なお、スタジオショット（冒頭）における検出漏れは、話題の冒頭にもかかわらず、タイトル字幕が出現しなかったものである。

表 1: スタジオショット検出実験の結果

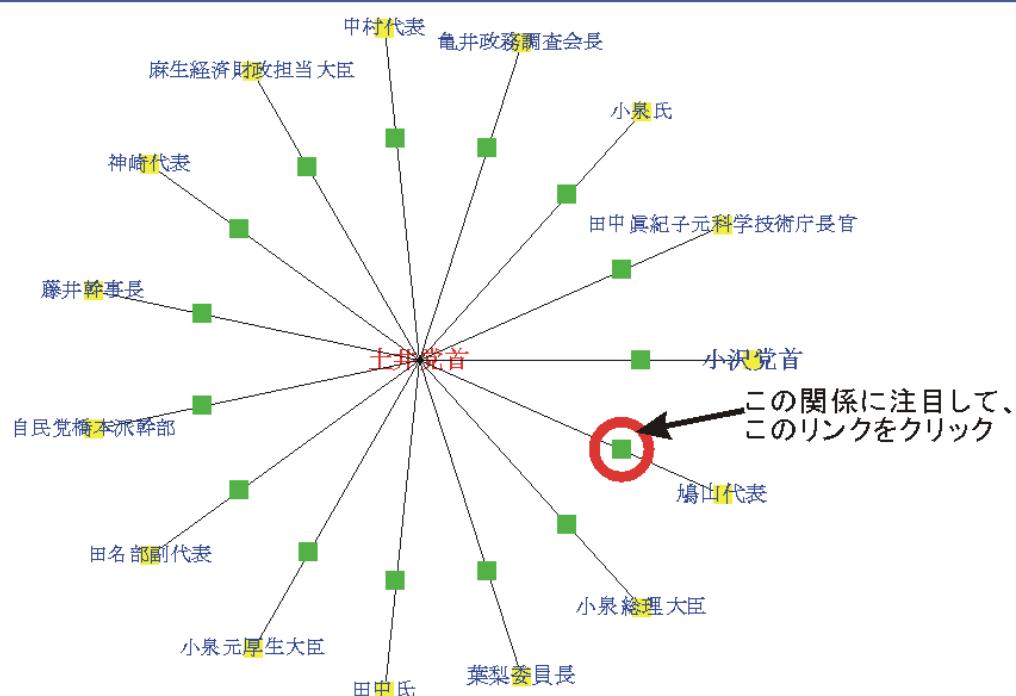
ショット	正検出	誤検出	検出漏れ	再現率	適合率
スタジオ	53	0	0	100%	100%
リ（冒頭）	37	0	6	86%	100%

3.4 人物関係の抽出

人物関係の抽出は、3.2で検出した物名の共起を、3.3で検出した話題内においても、話題にして集計することにより得られる。このようにして抽出した人物関係は、4に示したドラマにおける登場人物間に示されるような人物関係の意味付け（「友」「片思い」など）を今までにはいたらないが、どのような話題で共起しているかによって関係を把握することはできるだろう。

3.5 検索・閲覧インターフェース

以上の手順に従って抽出されたニュース映像の登場人物の人物関係を手がかりに映像を検索・閲覧するイン



Applet started.

7: 物関係の可 化： 物関係閲覧インターフェース

タフェースを実 し， 供する。 7に 物関係を可 化した 物関係の閲覧インターフェースを示す。ここでは， 央に する 物に注目して， 円周 に配 した 物 との一階 の 物関係を可 化している。 物関係をた どるためのインターフェースとして， 円周 に配 した 物名の 央にある 角形をクリックすることにより， 注 目する 物を 替えることができる。 また， 関係の強 さ (同一話題内の共起頻度の サ) に比例して 物名を きく表示し， 右回りに順に並べてある。

次に， 7の 物関係のうち， 特 のものに注目して 映 を閲覧するためのインターフェースを 8に示す。 この映 閲覧インターフェースは， 7の 物関係閲覧イン タフェースの 物間のリンク (の 物と円周 の 物を するスポーク の 分) にある 角形 (この は 7 の○で囲まれた 角形) をクリックすることにより表示される。 このインターフェースにより， 注 目する 物関係が得られた映 を閲覧できるだけでなく， それがどのような 物関係であったかを動画 および 文字放 字幕の 報から得ることができる。 表示されている 表画 は， 各々の話題において円周 の 物名 (この例の は「鳩 表」) が文字放 字幕 に出現した時 のフレームである。



8: 映 閲覧インターフェース： 7 の○で囲まれた リンクをクリックして表示される結果を例示

4 実験

述べた手順および手法に基づき、実のニュース映像から物関係を抽出し、検査・閲覧インターフェースをした。実験には、平成13年3月16日から5月31日の期間内の53日分の特集ニュース映像（毎日30分）を用いた。この映像群に含まれる文字放字幕は計2,807文、話題は計560であった。また、今回の実験では、話題境界の検出はほぼ確に自動化できることをふまえ、3.3で述べた手法を用いずに、手で切り分けた結果を利用した。その結果、のべ2,799件の共起、異なり292件の物関係が得られた。

得られた物関係を見ると、表記が異なる同一物の共起も物関係とみなして抽出されていた。これには単純に同義語による異表記（「小首」と「小理」など）のほか、同義語ではないが意味的に同義である異表記（「小」と「小元」など）も含まれる。このような異表記の名寄せを行うことは膨大な背景知識を必要とするため難であるが、これも物関係の一種（等価(is-a)な関係）とみなすことにより、検査の問合せ語の表記的・意味的様子を吸収することができるとえられる。

5 おわりに

本では、規模ニュース映像コーパスから、話題の共起を手がかりとして物関係を抽出し、それを引とした検査・閲覧手法を案し、実験した結果を介绍了。

後は、より規模な映像コーパスを以てし、さらに画面の物の顔の利用方法などの検討をめ、より映像内容にした検査・閲覧手法の実現を目指す。また、共起がじた話題の種類（政治、事件など）を判別することにより、物関係の把な意味付け（物間のリンクの意味）と可視化についても検討していくたい。

参考文献

- [1] 安藤男、井亭、小林夫、本間一、後藤淳、島、小健、藤庄衛、尾和穂、木寛之、井篤、井淳、田英輝、都木徹、宮栄一、磯野春雄：“音認識を利用した放送用ニュース字幕システム”，学論、vol.J84-D-II, no.6, pp.877-887 (June 2001).
- [2] 井手一郎、田英彦：“末尾の名に目したテレビニュース字幕の語義解釈”，処学論、vol.39, no.8,

pp.2543-2546 (Aug. 1998).

- [3] 井手一郎、本、浜田玲、田英彦：“ショット分類に基づく映像への自動的引付け手法”，学論、vol.J82-D-II, no.10, pp.1543-1551 (Oct. 1999).
- [4] 橋、藤本研、納、南敏：“顔領域抽出に有効な修正HSV表示系の提案”，テレビ誌、vol.49, no.6, pp.787-797 (June 1995)
- [5] 藤一、于、孟洋、片紀、浜田喬：“映像インデクシングの実的研究のための映像データベースシステムの試行”，2001年学術SD-5-9, vol. 報システム2, pp.367-368 (March 2001).
- [6] S. Satoh, Y. Nakamura, and T. Kanade: “Name-It: Naming and detecting faces in news videos”, IEEE Multimedia, vol.6, no.1, pp.22-35 (March 1999).
- [7] 藤隆、児島治彦、阿久明、外佳：“映像コーパスの構造と分類”，学論、vol.J82-D-II, no.10, pp.1552-1560 (Oct. 1999).
- [8] B. M. Sundheim: “Named entity task definition, version 2.1”, Proc. Sixth Message Understanding Conf. (MUC-6), pp.317-332 (Nov. 1995).
- [9] “付録H:NEの定義”，IREXワークショップ予稿集，pp.264-273 (Sep. 1999).
- [10] 技術研究報処理開発機(RWCP):“RWCTextCorpus 1版”.
- [11] 京都大学学院報学研究科知能報学言語メディア研究室：“京都大学テキストコーパス2.0版”，<http://www-lab25.kuee.kyoto-u.ac.jp/nl-resource/corpus.html>より手.
- [12] 京都大学学院報学研究科知能報学言語メディア研究室：“日本語形態解説システムJUMAN 3.6版”，<http://www-lab25.kuee.kyoto-u.ac.jp/nl-resource/juman.html>より手.
- [13] 語研究所：“「分類語彙表」形式による語彙分類表（補版）[補ニタ用化データ】”.