

ニュース映像アーカイブにおける登場人物の顔照合を利用した名寄せ

小笠原 崇[†] 高橋 友和^{†,††} 井手 一郎^{†,†††} 村瀬 洋[†]

[†] 名古屋大学大学院 情報科学研究科 〒464-8603 愛知県名古屋市千種区不老町

^{††} 日本学術振興会

^{†††} 国立情報学研究所 〒101-8430 東京都千代田区一ツ橋 2-1-2

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

あらまし 近年、大量に蓄積された映像データを効率的・効果的に利用するための技術が求められている。ニュース映像アーカイブにおいて、閲覧・検索のキーとなる重要な情報の一つとして映像中の登場人物がある。登場人物をキーとした映像の閲覧や検索を支援する技術は、従来から研究がなされているが、それらはもっぱら「名前」によってのみ人物を区別し、複数の呼称を持つ人物は、その呼称ごとに別々の人物として扱われてきた。本研究では、この問題に対処すべく、複数の呼称を持つ登場人物の同定（いわゆる“名寄せ”）を行う手法の開発を行っている。名寄せは、言語処理や意味解析の高度化により、ある程度は実現できるが、「東国原知事」に対して「そのまんま東」のような通称を名寄せする場合など、言語処理での解決が困難な場面は多々ある。そこで、本手法では各呼称に対応付く人物の顔を利用することで、言語的な変化に影響されない枠組みでの名寄せを試みた。本報告では、5年半に渡り蓄積されたニュース映像アーカイブに対し本手法を適用した。結果から、本手法の名寄せ効果を確認した。

キーワード ニュース映像, アーカイブ, 顔照合, 名寄せ

People Identification in Broadcast News Video Archive by Face Matching

Takashi OGASAWARA[†], Tomokazu TAKAHASHI^{†,††}, Ichiro IDE^{†,†††}, and Hiroshi MURASE[†]

[†] Graduate School of Information Science, Nagoya University Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8603 Japan

^{††} Japan Society for the Promotion of Science

^{†††} National Institute of Informatics 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430 Japan

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

Abstract Recently, techniques to deal with a large amount of video data efficiently and effectively are required. When we search a news video archive for video clips, persons who appear in the clips are one of the most important information. Therefore many works focus on indexing of persons in videos. However, in those works, the persons are identified by the names. In this case, a “name” represents an appellation that changes according to time and circumstances, therefore, a person can be called by several different names or aliases. This is a serious problem because when we want to search for the person by the name. We have tackled the problem by identifying a person who has multiple names with a face matching strategy. This framework does not influenced by linguistic change. Experiments using a news video archive consisting of 940 hours which were recorded over five years demonstrate the effectivity of the proposed.

Key words News Video, Archive, Face Matching, Name Identification

1. ま え が き

1.1 研究の背景

近年、通信技術や記憶装置の発達、また、メディアの多様化に伴い、世の中の映像データは氾濫の一途をたどっている。そ

れに伴い、それら大量の映像データを効率的・効果的に利用するための技術が求められている。

放送映像、中でもニュース映像は資料的価値が特に高い。そのため、意味内容に基づく構造解析や効率的な閲覧のための検索方法など、ニュース映像を対象とした研究は様々な方向から

なされている [4], [8]。

ニュース映像は人間社会の現状を映し出す貴重な情報源である。そのため、その映像中に現れる“人物”に注目することは自然かつ重要であると考えられ、ニュース映像中に現れる人物の名前や顔に着目した研究は、すでにいくつかなされている [2], [5], [7]。しかしながら、“人物”に注目したこれらのデータアーカイブの解析においては、放送映像を対象とした場合に限らず、1つの共通した大きな問題がある。その問題は、人物が専ら“名前”(呼称)によってのみ区別されるために生じる。人物が言及されるときに用いられる呼称は、状況や時流の変化によって異なるため、1人の人物は複数の呼称と対応しうる。それに関わらず、従来の多くの研究では人物が一般に呼称によって区別されるため、複数の呼称を持つ人物は、その呼称ごとに別々の人物として扱われてしまう問題があった。

1.2 呼称の多様性の原因と問題

同一人物に関する呼称変化の例を図1に示す。同一人物に対し、a. から h. に挙げた異なる8つの呼称が与えられている。

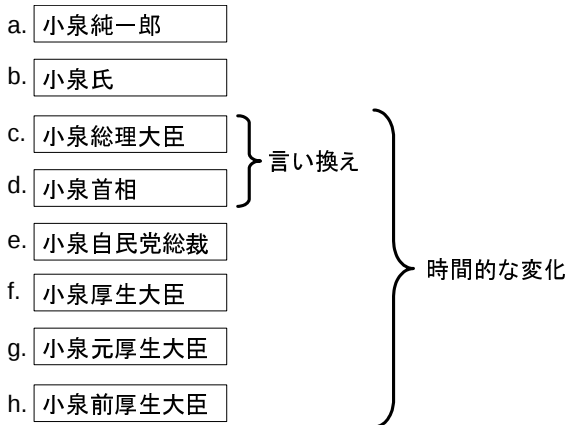


図1 同一人物に関する呼称変化の例

呼称の多様性を引き起こす原因は以下のものが挙げられる。

(1) 敬称・役職等の付与

a. から h. までの呼称は、全て同一の人物を指すが、敬称や役職を付与することによって、多様化することがわかる。敬称・役職が付与された人物の同定を行うためには外部からの社会的な知識が不可欠である。

(2) 同義語の言い換え

c. と d. は、「総理大臣」と「首相」という、本来同一の概念を異なった語で表したために起こった呼称の変化である。このような場合は、あらかじめシソーラス(類語辞書)を用意しておき、同義語の言い換えを行うことで対応付けが可能となる。

(3) 時間的な変化

c. から h. までを全て見ると、同義語の置き換えではなく、本来別の概念を指す名詞を付加したことによって、多様な呼称が現れている。図1に示した人物は、歴史のある時点では「厚生大臣」という役職についていたが、その後、「自民党総裁」や「総理大臣」となったため、このように呼称が変化している。時間的な変化は、図1の例のような役職の変化の他にも、事件を起こしたため逮捕・起訴された場合の「容疑者」や「被告」と

いった称号の付加や、婚姻等による姓の変化、改名、さらにはニックネームなど、様々なものがある。長期にわたり蓄積したデータに対しては、このような変化はことさら大きな問題となり、社会の変化に対応して随時知識を更新しなければならない。

これらの呼称変化が引き起こす問題を、映像の検索を例に1つ挙げる。例えばここに、人物をキーにして映像を検索できるシステムがあり、検索の際には、検索したい人物の呼称をクエリとして与えるとする。このとき、ユーザが与えるクエリが「小泉総理大臣」であったとすると、システムが検索結果として返せるのは「小泉総理大臣」として言及された映像のみであり、他の呼称(例えば、「小泉自民党総裁」、「小泉厚生大臣」)で言及されていた映像は、ユーザが探している“人物”の映像であるに関わらず、検索結果としては挙がってこない。

1.3 提案手法の概要

本研究では、上記の問題を解決するため、同一の“人物”を示す、複数の異なる呼称を同定(“名寄せ”)する手法を提案する。

名寄せの高精度化は、言語処理や意味解析の高度化により、ある程度は解決できるが、1.2節で述べた呼称変化とその原因を考慮すると、呼称を“人物”として同定することの困難さがわかる。特に、(1)「敬称・役職等の付与」や(3)「時間的な変化」については、その呼称を文字列として眺めただけで同定することは困難を極める。例えば、人物の姓に着目して名寄せを行った場合、「小泉総理大臣」に対し名寄せされるべきでない「小泉孝太郎」を名寄せしてしまったり、「田村亮子選手」に対し名寄せされるべき「谷亮子選手」が名寄せされなかったりという問題が生じる。他にも、「東国原知事」に対して「そのまんま東」という通称を名寄せする場合など、言語処理での解決が困難な場面は多々ある。

これらを正しく名寄せするためには、外部から与えられる社会的な知識が不可欠であるが、人手を介さない自動的な処理での名寄せを目指す本研究では、それらの知識は利用できない。そこで、本研究では、各呼称に対し映像中で対応付く顔を利用することで、言語的な変化に影響されない枠組みで名寄せを行う。

まず、ニュース映像アーカイブ中の映像内で言及される各呼称に対し、ニュース映像の各場面でのその呼称に対応付けられる顔を収集する。次に、各呼称に対して収集された複数の顔によって形成されるクラスどうしを照合し、照合の類似度をもとに、呼称の同定を行う。

本手法の類似研究としては、Bergら[1]がWebニュースを素材にして行った、呼称に基づいて収集された顔画像を、顔の類似度を元に再クラスタリングする手法がある。ただし、この手法では、ニュース映像内で対応付けられた呼称と顔の対応付けは、最終的なクラスタリング結果において保証されず、ニュース映像内では全く起こりえなかったラベル付け間違いが起こる可能性がある。これに対し、本手法は、1つの呼称をラベルとして持つクラスどうしを照合しており、呼称と顔の対応付けはニュース映像から得られるもののまま処理を進めるため、そのようなラベル付け間違いは起こらない。

本報告では、以降、2.章で処理の詳細を述べる。3.章で提案

手法の効果を実験によって示し、4. 章でまとめる。

2. 顔を手がかりとした名寄せ

本手法における名寄せは、基準とするある1つの呼称に対して、その呼称が指す人物と同一と見なされる人物を示す他の呼称を列挙する処理である。同一と見なされる人物とは、顔照合で類似度が閾値以上となったものを指す。図2に、本手法における名寄せの概念図を示す。基準となる「安倍幹事長」という呼称に対し、「安倍官房長官」および「安倍官房副長官」という呼称が名寄せされている。

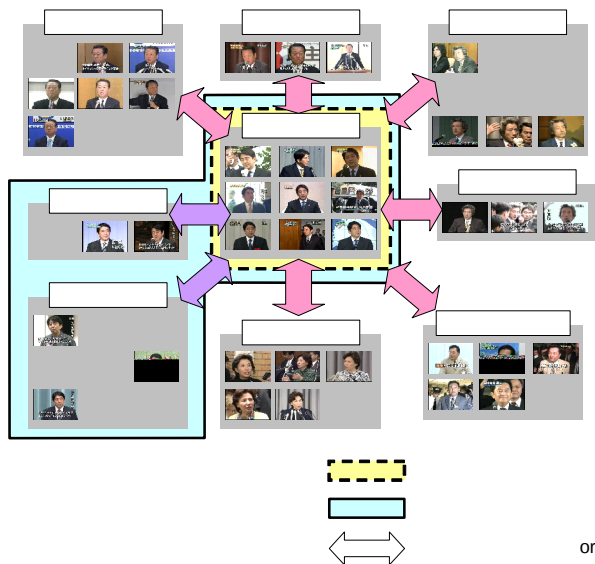


図2 名寄せの概念図

顔の照合は、画像1枚ずつ比較するのではなく、同一の呼称ラベルを与えられた複数枚の画像を1つのクラスとし、そのクラス単位で照合する。2.2節にて詳細を述べるが、このようにクラス単位の照合を行うことで、クラス内に誤った顔が混入した場合にも、その影響を受けにくい名寄せを実現する。

本手法による名寄せは、大まかに、以下に挙げる2つの処理によって行われる。

手順1 各呼称に対応する顔の収集

手順2 収集した顔の照合

以降、2.1節および2.2節で、各処理について述べる。

2.1 呼称に対応する顔の収集

放送局から配信されるニュース映像は、主に画像と音声から構成されているが、近年これらに加え、クローズドキャプション (Closed Caption, CC) と呼ばれる音声を書き下したテキストデータが付与され始めている。ここでは、このクローズドキャプションから得られる人物の名前 (呼称) に対し、各場面において対応する顔を収集する。呼称に対応する顔の収集処理の大まかな流れは図3のようになる。

まず初めに、クローズドキャプションテキストから人物名詞を抽出する。次に、その人物名詞が得られたクローズドキャプションの1文と時間的に同期する動画の区間から、その人物名詞に対応する顔を抽出し、ニュース映像の各場面における顔

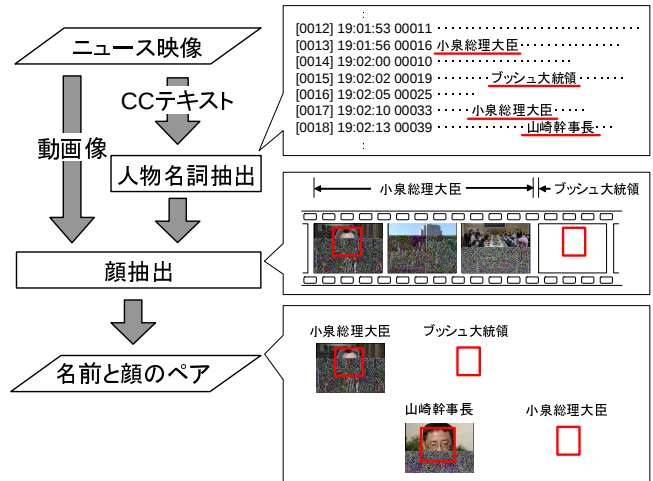


図3 名前に対応する顔の収集

と呼称の対応付けを得る。ただし、1文中に複数の人物名詞が現れる文は、除外する。これは、その文に対応する区間の動画画像から得られる顔が、どの名前と対応付けられるべきかがあいまいとなり、正しい名前と顔の対応が得られない可能性があるためである。よって、本手法でははじめから除外することとした。これは、本手法が、数100～数1,000時間という大規模な映像群を対象とすることを想定しており、いくらかの顔を除外しても、照合に十分なデータ量 (顔の数) が確保されることが期待できるためである。このことは、3.2節で述べる実験においても確認されている。この対応付けのあいまい性を解消するために Satoh らの Name-It システム [7] を利用することもできるが、そのような処理を行わずともデータ量は充足されうる。

以下に、クローズドキャプションテキストからの人物名詞抽出、および、動画からの顔抽出について述べる。

2.1.1 CCテキストからの人物名詞抽出

ニュース映像アーカイブの全てのCCテキストをテキスト解析し、その中に現れる人物名詞を抽出する。本手法で用いる映像は、[3]で既にアーカイブ化されたものを用いており、CCテキストは事前に音声との同期がとられている。

人物名詞抽出には、井手らの手法 [9] を用いており、おおまかに以下に示す手順で行う。

手順1 CCテキスト内の各文に形態素解析を施し (JUMAN [10] 3.61 を使用)、名詞列を抽出する。

手順2 「～さん」「～会長」「～大臣」といった接尾名詞に注目して事前に作成した辞書と、手順1で抽出された各名詞列とを照合し、その語義属性を解析することで人物名詞を抽出する。

2.1.2 動画からの顔抽出

2.1.1で述べた処理で抽出された人物名詞が含まれるクローズドキャプションの1文と時間的に同期する動画の区間から、その人物名詞に対応する顔を抽出する。これにより、ニュース映像の各場面における顔と呼称の対応を得る。ただし、前述のとおり、顔と名前の対応付けのあいまいさを回避するため、1文に複数の人物名詞が含まれる文については、この処理の対象としない。

ここで抽出される顔は、該当する映像区間において最初に以下の条件を満たして検出された顔 1 つである。ただし、これらの条件を満たす顔が画面内に複数ある場合は、最も画面中央に映っている顔を抽出する。

条件 1 画面上に、一定以上の大きさで映っている

条件 2 正面向きの顔

これらの条件は、ニュース映像において、そこで言及されている人物が、概ね以下に挙げる画像的な特徴をもって撮影されていることを念頭においたものである。

特徴 1 主に話題の中心となっている人物が映し出され、話題と全く関係がない人物が故意に映されることはない。

特徴 2 言及される人物を映す場合には、撮影の際に何らかの制約がない限り、その人物が画像中央に大きく映るように撮影される (図 4)。

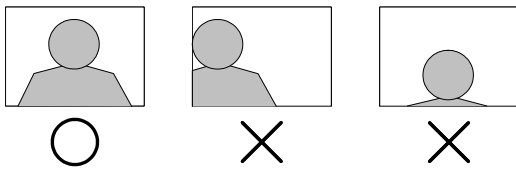


図 4 ニュース映像における人物の撮影

顔の検出には、照明条件やノイズに対してロバストであり、かつ解像度に依存せず高速な検出が可能な Joint-Haar like 特徴 [13] を用いる。これは、複数の Haar-like 特徴 [6] の共起に基づく特徴量であり、Haar-like 特徴を組み合わせて顔の構造に基づいた特徴の共起関係を評価することにより高い識別能力が得られている。

Haar-like 特徴とは、画像における特徴量として、照明条件やノイズの影響を受けやすい各画素の明度値をそのまま用いるのではなく、近接する 2 つの矩形領域の明度差を用いることで得られる特徴量である (図 5)。



図 5 基本的な Haar-like 特徴のセット

2.2 収集した顔の照合

顔の照合は、同一の呼称ラベルを割り当てられている複数の顔を用いてそのラベルの顔を表す部分空間を作り、その部分空間どうしを照合する相互部分空間法 (Mutual subspace method) [11], [12] によって行う (図 6)。

相互部分空間法は、一枚の静止画での照合ではなく、複数の多様な画像群どうしの照合を行うことにより、顔向きや表情の変化、照明変動等に対してロバストな顔照合を行う手法である。撮影環境が一定とは限らないニュース映像に出現する顔の照合にも有効であると考えられる。

本手法において、相互部分空間法による呼称をラベルとして持つクラス対クラスの照合には、撮影環境の変化や表情の変化に対するロバスト性以外にも、もう 1 つの大きな利点がある。

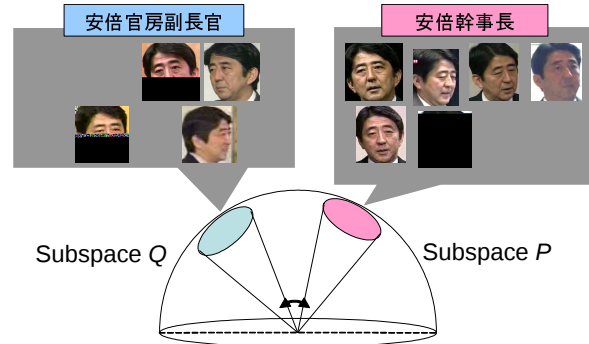


図 6 相互部分空間法 の概念図

それは、2.1 節で述べた呼称に対応する顔の収集の際に、ラベル付けで誤ったために混入する異なる人物の顔に対する効果である。これは、動画像中からの顔抽出時に、本来その呼称と対応付くべきではないものを抽出してしまった場合に発生する。Sato ら [7] や Ikizler [2], Ozkan ら [5] のように、クラス内の各顔を 1 つずつ照合する手法をそのまま名寄せに応用すると、このような、誤った顔の存在が名寄せ結果に大きく影響してしまう。これに対し、本研究では、1 つの呼称に対して 1 つのクラスを作り、クラスどうしの平均的な顔特徴の比較によって名寄せを行うため、誤った顔の占める割合が大きくない限り、それらの影響を軽減できる。

3. 人物の名寄せ実験

本手法による、ニュース映像中の人物の名寄せ効果を確認するために、2. で述べた処理を実際のニュース映像に適用して実験を行った。

3.1 実験映像

実験に用いた映像は、2001 年 3 月から 2006 年 9 月までのおよそ 5 年半に渡り蓄積された 1,906 本 (総計約 940 時間) のニュース番組、「NHK ニュース 7」である。映像の仕様の詳細を表 1 に示す。これらの映像は、国立情報学研究所が構築したニュース映像アーカイブ [3] より提供されたものである。

表 1 実験で用いたニュース映像

ニュース番組	「NHK ニュース 7」
フレームレート	30 [frame/sec]
画素数	352 × 240 [pixel]
期間	2001 年 3 月 16 日 ~ 2006 年 9 月 3 日
本数	1,906 [本]
総時間	939 時間 45 分 26 秒

「NHK ニュース 7」は、通常 1 本 30 分の番組であるが、重大ニュースの特別報道等により時間が伸縮する。

3.2 人物名詞抽出の結果と分析

人物名詞抽出の結果を表 2 に示す。表 2 に示すとおり、実験に用いた延べ約 940 時間のニュース映像には、総計 257,045 文の CC テキストが付与されており、そこから、総計 94,197 個の人物名詞が抽出された。このうち、人物名詞が 1 つのみ含まれる文から得られたものは、53,637 個であった。このことから、2.1 節で述べた顔の収集の際に、「1 文中に複数の人物名詞が現

れる文を処理対象から除外しても、照合に十分な数の顔が得られる」ことが確認された。

文字列としてユニークな数(異なり数)は18,793個であった。これはすなわち、ニュース映像から得られた呼称の数である。総出現回数が94,197であるので、人物名詞1種類あたりの出現回数は平均的におよそ5回となるが、出現頻度にはバラつきがあり、図7に示すような分布となった。

出現回数が5回未満の人物名詞15,888個の内訳を確認したところ、それらは概ね、窃盗事件の容疑者や交通事故の被害者等、日常的に起こる(社会的な重要性の低い)事件・事故の関係者であり、他のニュースで取り上げられることは、まずないと考えられる人物であった。本手法で行う名寄せの対象となる人物は、何度も様々なニュースで取り上げられる人物であると考えられる。そのため、実際に名寄せされるべき呼称は、上記以外のもの、すなわち、ある程度の回数ニュース中に出現した、およそ3,000個であると考えられる。

表2 人物名詞抽出結果

CC テキスト数 [文]	抽出された人物名詞の数 [個]	
	総数	異なり数
257,045	94,197	18,793

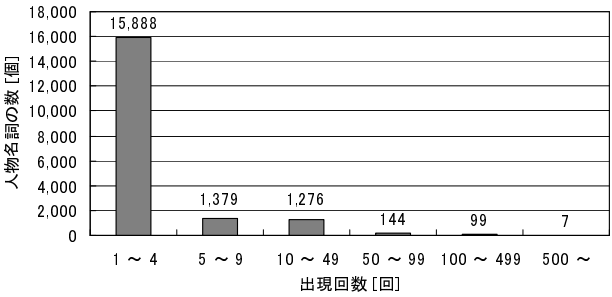


図7 1つの人物名詞の出現回数の分布

3.3 名寄せ結果

ニュース映像から抽出された人物名詞各々に対して名寄せを行った結果の一部を表3に示す。表には、「基準となるラベル」に対して、本手法によって「名寄せされたラベル」が示してある。「名寄せされたラベル」を人手で正誤確認し、正しく名寄せされていると判断したものを太字で、誤って名寄せされていると判断されたものは括弧付きで示してある。

なお、名寄せの際に用いられる顔類似度の閾値は、本手法で用いた相互部分空間法を顔認識器として用いたとき、入力に誤った顔が混入していない前提で最適な認識率を持つ値とした。

全体の名寄せ結果をまとめると、表4のようになった。ただし、人物名詞抽出時の誤抽出(個人を示す人物名詞以外が得られたもの)は事前に人手で除いた。

本手法によって行われた名寄せ総数は、95ペアであり、先に表3で示したのは、このうちの12ペアである。名寄せ総数95ペアのうち、人手にて正誤を判定したところ、正解は37ペア、誤りは58ペアであった。この結果、式(1)で定義される正解

表3 名寄せ結果(一部)

基準となるラベル	名寄せされたラベル
安倍官房副長官	安倍幹事長代理 安倍副長官 安倍幹事長
堀江元社長	堀江前社長 堀江社長 (鈴木議員)
小泉元厚生大臣	小泉総理大臣
パウエル長官	(ブリクス委員長) パウエル国務長官
北側大臣	北側国土交通大臣
三木谷社長	(堀江社長)
松井秀喜選手	(イチロー選手)

表4 名寄せ結果(全体)

名寄せ総数 [ペア]	正解 [ペア]	誤り [ペア]
95	37	58

率は、およそ39%となった。

$$\text{正解率} = \frac{\text{正解}}{\text{名寄せ総数}} \tag{1}$$

3.4 名寄せ結果の考察

表3を見ると、「堀江元社長」に対して「鈴木議員」が名寄せされるなど、誤った名寄せは見られるものの、「小泉元厚生大臣」に対して「小泉総理大臣」が名寄せされるなど、本来ならば社会的な知識が必要な名寄せができており、本手法の有効性が確認できた。

しかしながら、全体として見た正解率は、名寄せの結果として十分な精度とは言えない。そこで、名寄せ誤りの原因について考察する。

名寄せ誤りは、主に、顔照合の精度に起因するものと考えられる。しかし、名寄せを誤った例をいくつか注意深く見てみると、それ以外の原因が浮かび上がってくる。その例として挙げられるのが、表3中の「三木谷社長」と「堀江社長」の組み合わせや、「松井秀樹選手」と「イチロー選手」の組み合わせである。これらの組み合わせは、ニュース映像において、頻繁に共通する話題(トピック)に共起する。具体的には、前者はIT関連の話題で、後者はスポーツ(特に、アメリカのメジャーリーグ)の話題で頻繁に共起する。これらの名寄せ誤りに関しては、顔照合の精度に起因するものの他に、共通したトピックで頻繁に共起したために、顔抽出時に誤って混入する顔が相互に多くなり、その結果が名寄せに影響してしまったことも原因の1つとして考えられる。よって、この問題に対処するためには、人物名詞と顔の対応付けの精度を向上させる必要がある。

また、正解率を高めるために必要な「人物名詞と顔の対応付け」および「顔照合」の精度向上には、現時点ではまだ名寄せしきれていない他のラベルを名寄せできるようになる効果もあると考えられる。正解率を保ったまま名寄せされるラベルの数が増やせれば、映像の検索結果として、ユーザの望む映像をより多く提示することができる。

4. む す び

本報告では、ニュース映像に現れる同一の人物を示す複数の呼称を、自動的に同定(名寄せ)する手法を提案した。これにより、従来、それぞれの呼称が、あたかも別人物のように扱われていた問題の解決を図った。

名寄せの高精度化は、言語処理や意味解析の高度化により、ある程度は解決できるが、例えば、人物の姓に着目すると、「小泉総理大臣」に対し名寄せされるべきでない「小泉孝太郎」を名寄せしてしまったり、「田村亮子選手」に対し名寄せされるべき「谷亮子選手」が名寄せされなかったりという問題が生じる。他にも、「東国原知事」に対して「そのまんま東」のような通称を名寄せする場合など、言語処理での解決が困難な場面は多々ある。そこで、本手法では各呼称に対応付く人物の顔を利用することで、言語的な変化に影響されない枠組みで名寄せを行った。

1,906本(延べ約940時間)のニュース映像に対し、提案手法による自動名寄せ実験を行った結果、正解率は39%と低いながらも、「安倍官房副長官」に対して名寄せされた、「安倍幹事長代理」「安倍副長官」「安倍幹事長」などを例に、従来は社会的な知識がなければ名寄せが困難であったものに対しても、正しく名寄せする処理に成功した。

本手法によって行われる名寄せは、映像の検索や閲覧に応用できる。例えば、ニュース映像を登場人物によって検索するシステムにおいて、本手法を用いることで、従来は得られなかった、クエリとして与えた呼称以外の呼称で言及された映像を得ることができる。

現時点では、本手法の名寄せ精度は、まだ改善の余地がある。名寄せの正解率を高めることができれば、ユーザが検索結果を取捨選択する際の手間の軽減につながる。また、正解率を保ったまま、名寄せされるラベルの数をより多くできれば、ユーザが映像を検索した結果として得られる映像が増え、いわば埋もれた映像をユーザに提示することができる。これらの精度を高めるためには、名寄せ判断に用いる顔照合の高精度化が必要となる。顔照合の高精度化には、顔の類似度計算手法の検討とともに、照合に用いる顔の厳選などが必要となる。

謝 辞

本研究にとって不可欠な技術(2.1.2の顔検出および2.2の顔照合)の提供およびご助言をくださった株式会社東芝研究開発センターに深く感謝する。研究に必要な数多くのデータを提供してくださった情報・システム研究機構国立情報学研究所に感謝する。日頃より熱心に御討論頂く名古屋大学村瀬研究室諸氏に感謝する。本研究の一部は日本学術振興会・文部科学省科学研究費補助金および21世紀COEプログラム「社会情報基盤のための音声・映像の知的統合」による。本研究では、画像処理にMISTライブラリ(<http://mist.suenaga.m.is.nagoya-u.ac.jp/>)を使用した。

文 献

- [1] Tamara L. Berg, Alexander C. Berg, Jaety Edwards,

- Michael Maire, Ryan White, Yee-Whye The, Erik Learned-Miller, and David A. Forsyth. “Names and faces in the news”. *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2004)*, Vol. 2, pp. 848–854, July 2004.
- [2] Nazli Ikizler and Pinar Duygulu. “Person search made easy”. *Proc. International Conference on Image and Video Retrieval (CIVR 2005), Lecture Notes in Computer Science*, Vol. 3568, pp. 578–588, July 2005.
- [3] Norio Katayama, Hiroshi Mo, Ichiro Ide, and Shin’ichi Satoh. “Mining large-scale broadcast video archives towards inter-video structuring”. *Advances in Multimedia Information Processing – PCM2004 5th Pacific Rim Conference on Multimedia, Lecture Notes in Computer Science*, Vol. 3332, pp. 489–496, November 2004.
- [4] Yuichi Nakamura and Takeo Kanade. “Semantic analysis for video contents extraction – Spotting by association in news video”. *Proc. 5th ACM International Conference on Multimedia*, pp. 393–402, November 1997.
- [5] Derya Ozkan and Pinar Duygulu. “Finding people frequently appearing in news”. *Proc. International Conference on Image and Video Retrieval (CIVR 2006), Lecture Notes in Computer Science*, Vol. 4071, pp. 173–182, July 2006.
- [6] Constantine P. Papageorgiou, Michael Oren, and Tomaso Poggio. “A general framework for object detection”. *Proc. International Conference on Computer Vision (ICCV 1998)*, pp. 555–562, January 1998.
- [7] Shin’ichi Satoh, Yuichi Nakamura, and Takeo Kanade. “Name-It: Naming and detecting faces in news videos”. *IEEE MultiMedia*, Vol. 6, No. 1, pp. 22–35, January 1999.
- [8] Howard Wactlar, Takeo Kanade, Michael A. Smith, and Scott M. Stevens. “Intelligent access to digital video: The Informedia project”. *IEEE Computer*, Vol. 29, No. 5, pp. 46–52, May 1996.
- [9] 井手一郎, 浜田玲子, 坂井修一, 田中英彦. “テレビニュース字幕の語義属性解析のための辞書作成”. 電子情報通信学会論文誌(D-II), Vol. J85-D-II, No. 7, pp. 1201–1210, July 2002.
- [10] 京都大学長尾研究室, 東京大学黒橋研究室. “日本語形態素解析システム“JUMAN””. <http://www.kc.t.u-tokyo.ac.jp/nl-resource/juman.html>.
- [11] 前田賢一, 渡辺貞一. “局所的構造を導入したパターンマッチング法”. 電子情報通信学会論文誌(D), Vol. J68-D, No. 3, pp. 345–352, March 1985.
- [12] 山口修, 福井和広. “顔向き表情変化にロバストな顔認識システム“Smartface””. 電子情報通信学会論文誌(D-II), Vol. J84-D-II, No. 6, pp. 1045–1052, June 2001.
- [13] 三田雄志, 金子敏充, 堀修. “顔検出に適した Joint Haar-like 特徴の提案”. 画像の認識・理解シンポジウム(MIRU2005)講演論文集, pp. 104–111, July 2005.