

Web 画像ニュースの顔と人物名の対応付け

北 原 章 雄[†] 柳 井 啓 司[†]

本研究では、Web 上にある写真画像付きニュースから顔と人物名を抽出しそれらを正しく対応付けることを試みた。顔検出器を用いてニュース画像から人物の正面顔画像を切り出すと同時に、形態素解析によってニュース記事テキストから人物名を抽出した。画像特徴として正面顔認識に有効な固有顔法を用い、k-means 法によるクラスタリングで顔と人物名の対応付け実験を行った結果、適合率 62%を実現した。

Associating Faces and Names in Web Photo News

AKIO KITAHARA[†] and KEIJI YANAI[†]

We propose a system which extracts faces and person names from news articles with photographs on the Web and associates them automatically. The system detects face images from news photos with a face detector included in the OpenCV library (open source image recognition library), and extracts person names from news text with a Japanese morphological analyzer “Chasen”. It uses the eigenface representation as image features of extracted faces, and associates faces with names by the k-means clustering. In the experiment, we obtained the 62% precision rate regarding association of faces and names.

1. はじめに

1.1 背景

現在、Web 上には毎日大量のニュース記事が配信されており、その速報性から新聞やテレビニュースよりも Web から情報を収集している人も多い。これらの情報を保存しデータベース化しようとした場合、新聞は保管するための広い場所が必要である。テレビニュースでは、ビデオにより保存しようとするビデオテープの保管場所が必要であるし、計算機で動画ファイルとして保存するにしても非常に膨大な HDD 容量を必要としてしまう。これに対し Web ニュースは、広い保管場所もテレビニュースほどの膨大な HDD 容量も必要とすることなく情報を保存しておくことができる。また、Web ニュースはインデックス付けなどが簡単に行えるため、新聞やテレビニュースに比べて情報検索が遥かに容易である。これらの視点から、Web ニュースは簡単に蓄積することができ、有用なニュースデータベースとして利用可能であるといえる。

今日では、Web ニュースはテキスト記事だけでなく画像を含む場合が多くなっている。画像は文字だけ

ではわかりにくい内容を視覚的に捉えることができるため、記事内容のさらなる理解の手助けとなる。これらの画像の中には人物が写っている場合が多い。しかしながら、画像中のどの人物が記事本文に出てくる人物なのかは一目見てわからないことがある。そのような場合、画像の人物が誰なのか記事中から探すことになる。画像中の人物が 1 人や 2 人ならこの作業は簡単だが、複数人となると容易ではない。そこで本研究では画像中の人物の人物名を自動で提示できるように、画像内の人物の顔と記事内の名前を対応付けることを試みた。

1.2 目的

本研究の目的は、Web 画像ニュースデータベースを有効活用することである。そのために画像中の人物に注目して、画像内の顔と記事内の人物名を抽出し、抽出された顔と人物名の対応付けを行う。具体的には、当研究室の約 1 年半分の Web ニュースデータベースから顔と人物名を抽出し、それらを対応付ける実験を行った。

2. 関連研究

Web ニュース画像から顔を切り出すためには顔認識処理が必要である。顔認識処理の研究として、M.Turk らによる固有顔がある¹⁾。これは濃淡画像の画素値を

[†] 電気通信大学大学院 電気通信学研究科 情報工学専攻
Department of Computer Science, The University of
Electro-Communications

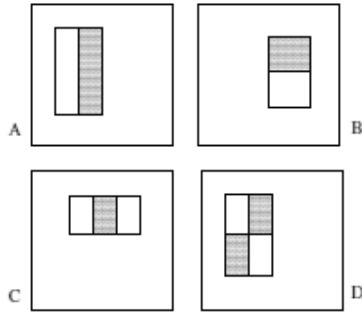


図 1 矩形特徴

ベクトルの要素とみなして、そのベクトルを主成分分析 (PCA) により圧縮し、圧縮されたベクトルを画像の特徴量とする方法である。この方法により正面顔に限定すれば高精度で顔認識が行えるということが示されている。本研究では顔と人物名を対応付けさせる際の顔の画像特徴量として、この固有顔を利用する。固有顔については 3. 2 で詳しく述べる。

他の顔認識研究として、P.Viola らによる研究がある²⁾。Viola らは、画像の局所領域の矩形特徴を利用した弱識別器を AdaBoost によって統合するという手法を提案した。矩形特徴は図 1 のような局所領域に含まれる画素の輝度の総和の組合せで表される。Viola らの手法は従来の手法に比べて非常に高速なため、オープンソースの画像認識関連のライブラリである OpenCV³⁾ にも利用されている。本研究では Web ニュース画像から顔を抽出するのに OpenCV を用いている。

本研究と同様の研究として、T.L.Berg らは画像特徴として PCA ではなく kPCA と LDA を用いた変形固有顔法を利用し、分類方法として k-means クラスタリングを用いて英文ニュースにおいて顔と人物名の対応付けを行った⁴⁾。Berg らの手法を使えば、英文ニュースにおいては約 95% の精度で顔と人物名が対応付けられるため、本研究では Berg らの手法を日本語 Web ニュースに応用して顔と人物名の対応付けを行うことにする。

3. 対応付けの方針

3.1 方針

本研究における顔と人物名の対応付けは抽出段階と対応付け段階からなる。抽出段階ではニュース画像から顔を切り出し、ニュース記事から人物名の抽出を行う。対応付け段階では、始めに候補ファイルと呼ばれるものを作成する。候補ファイルとは、各ニュースから抽出されたそれぞれの顔に、同じニュースから抽出された人物名を候補人物名として持たせたものをリスト化したものである。次に、候補ファイルをもとにし

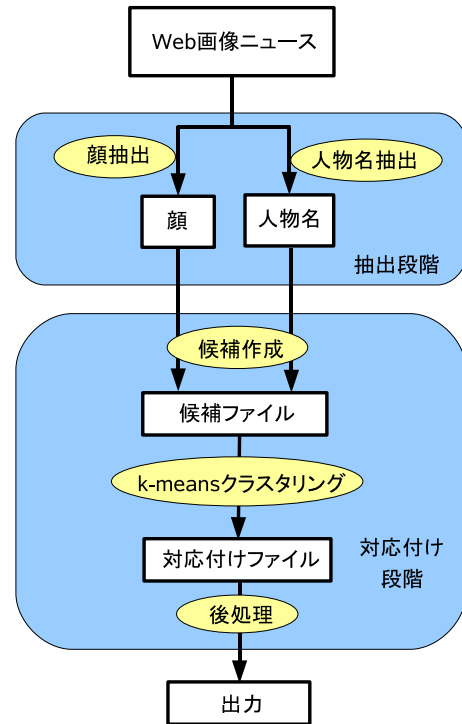


図 2 対応付け方法の流れ

て、各顔を候補の人物名の中で最適なものに分類する。画像特徴量として正面顔認識に有効な固有顔を用い、分類には k-means 法を利用する。最後に k-means クラスタリングの結果に対して後処理を加える。図 2 は対応付け方法の流れを図に表したものである。

3.2 固有顔

ここでは本研究で画像特徴量として利用する固有顔について述べる⁵⁾。固有顔とは、濃淡画像の画素値をベクトルの要素とみなして主成分分析を行い、その結果得られた圧縮されたベクトル (固有顔ベクトル) を画像特徴量とするものである。そして、複数の顔画像から固有顔ベクトルをもとに部分空間 (固有顔空間) を作成し、この固有顔空間の中で認識処理を行う。

固有顔空間の作成は、まず様々な人物の顔画像 N 枚を用意し A とする。($\vec{a}_i$ は各顔のベクトル)

$$A = \{\vec{a}_1, \dots, \vec{a}_i, \dots, \vec{a}_N\}$$

次に A の平均値 $\vec{u}$ 、共分散行列 C を求める。

$$\vec{u} = \frac{1}{N} \sum_{i=1}^N \vec{a}_i, \quad C = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N (\vec{a}_i - \vec{u})(\vec{a}_j - \vec{u})^T$$

これらの値を使って固有値問題

$$C\vec{v} = \lambda\vec{v}$$

を解き、固有値 λ_j 、固有ベクトル $\vec{v}_j$ を得る。ここで添字 j は固有値の大きな順に並べ、固有値の上位 d 次元を選択する。こうして得られた d 次元の固有ベクトル

ル $\{\vec{v}_1, \dots, \vec{v}_d\}$ が固有顔空間となる。本研究では固有値の寄与率が 90%以上となる上位 d 次元を選択した。

4. 対応付け方法

4.1 抽出段階

4.1.1 顔の切り出し

顔の切り出しには画像認識ライブラリ OpenCV に含まれる顔画像検出プログラムルーチンを利用する。ここで、切り出す顔は固有顔を利用できるように正面顔のみとし、正面顔の範囲は OpenCV が切り出せる正面顔の範囲とする。

4.1.2 人物名の抽出

人物名の抽出には日本語形態素解析ツール「茶筌」⁶⁾を利用する。ニュース記事を「茶筌」で解析した結果、品詞が以下のようにになっている名詞列を人物名とする。

抽出する人物名

- 苗字+名前 (例 小泉純一郎)
- 苗字+役職 (例 小泉首相)
- 苗字+国+役職 (例 ブレア英首相)
- 名前+・+苗字 (例 マイケル・ジャクソン)

ここで、役職名として苗字の次の単語がニュースによく登場する以下の 16 個を抽出することにする。

役職名

首相, 総理, 大統領, 大臣, ○○相, 皇太子, 皇后, 王子, 監督, 議長, ○○長, 代表, 総裁, 知事, 官房長官

ニュース記事内には、その記事を書いた記者の名前や写真を撮ったカメラマンの名前が含まれていることがある。これらの名前を抽出すると対応付けの精度を落とすことになるため、本研究では抽出しないことにする。実際には、「担当: “名前”」「【“名前”】」「“名前” 写す」「“名前” 撮影」のようにになっているものを抽出しないようにする。ただしこれは本研究で用いたニュースデータベースに対してのみ有効な処理である。

4.2 対応付け段階

4.2.1 候補ファイル作成

まず、顔・人物名共に 1 個以上抽出されたニュース(対応ありニュース)に対して、各人物名がそのニュース内で複数回出てくる場合はそれらの重複を排除する。次に、各ニュース内で「小泉純一郎」「小泉首相」のように同じ人物を指す名前が出てくる場合は、これらを統合して 1 つの人物名とする。この時、同じ人物を指すかどうかの判定は、「小泉純一郎」と役職名を取り除いた「小泉」のマッチングで判断する。マッチングが成功した場合は文字数の少ない方を捨てる。このようにして残った人物名に対して、すべての対応あり

ニュース中で 5 回以下しか現れない人物名を排除する。最後に、各対応ありニュースから抽出されたすべての顔に対して、そのニュース内で最終的に残った人物名をすべて候補に持つようなりリストを作成する。これが候補ファイルとなる。

4.2.2 k-means クラスタリングによる対応付け

候補ファイルで、各顔が候補のどの人物名に対応付けられるかについては、k-means 法によるクラスタリングを用いる。k の値は候補ファイル作成で最終的に残った人物名の種類数とし、各クラスは 1 つの人物名に対応するものとする。k-means クラスタリングは以下の手順で行う。

k-means クラスタリング

1. k 個のクラスタ中心決定

k 個のクラスタ中心を、そのクラスタの人物名を候補に持つ顔の中から 1 つ選んで決める。まず顔と人物名の 1 対 1 対応がある人物名クラスタについて決め、1 対 1 対応がなければ複数対応があるものの中からランダムに決める。ここで、すべての人物名クラスタの中心を決められるように、候補人物名の出現頻度が低いものからクラスタ中心を決定していく。

2. 各顔を適切な人物名クラスタに割り当て

各顔をその顔が候補に持つ人物名のクラスタ中心の中で最も近いクラスタに割り当てる。画像特徴量には固有顔を用い、距離はユークリッド距離を使用する。ただし距離が予め決めた閾値を超える場合は、正しい人物名が候補の中にないものとみなしてクラスタに割り当てることはしない。閾値は複数の値を実験し、最適なものを選ぶ。

3. クラスタ中心計算

各クラスタごとに中心を再計算する。

4. 判定

すべてのクラスタ中心が変化しなければ終了、そうでなければ 2 に戻る。

4.2.3 後 処 理

k-means クラスタリングの結果、あるニュースから切り出された複数の顔が、そのニュースから抽出された同一の人物名に対応づけられてしまうということがありえる。本来ならこれはあってはならないことである。そこで、同じニュースから切り出された複数の顔が同一人物名に対応付けられた場合は、その人物名クラスタの中心に最も近い顔のみを対応付け結果として許すことにする。この処理を行った結果を、最終的な顔・人物名対応付け結果とする。

5. 実 験

5.1 実験方法

実験データとして、当研究室で収集した 2005 年 3 月 6 日から 2006 年 8 月 31 日までの約 1 年半分の Yahoo!JAPAN ニュースを利用し、顔と人物名の対応付け実験を行った。抽出した顔画像は 50 × 50 のサイズに正規化し、これらの中から正しい顔画像を手動で 100 枚選択して固有顔空間を作った。k-means クラスタリング時における顔画像と各クラスタとのユークリッド距離に対して複数の閾値を設定し、適合率が最もよくなる閾値を調査した。ユークリッド距離は以下の式で表される (i :固有顔の要素, d :固有顔の次元数, F :顔画像, C :クラスタ中心)。

$$\text{ユークリッド距離} = \sqrt{\sum_{i=1}^d (F_i - C_i)^2}$$

実際にはユークリッド距離の二乗の値に対して、(1) なし、(2)6,000,000、(3)4,000,000、(4)2,000,000 の 4 種類の閾値を設定して実験を行った。

5.2 評価方法

実験結果の評価には適合率を用いることにする。適合率は、適合率=(正しい対応付け結果) / (すべての対応付け結果) と表される。Web 画像ニュースの顔と人物名の対応付けという本研究の目的のためには、適合率だけでなく再現率も重要になってくる。しかし対象とする Web 画像ニュースの数は膨大であり、再現率を求めることは困難である。そこで、まずは再現率をある程度考えずに適合率を高めることに重点を置く。

5.3 実験結果

5.3.1 顔の切り出し結果

ニュース画像から顔を切り出した結果を図 3、図 4 に示す。顔を切り出す際には、図 3 のように顔以外の領域を間違えて切り出してしまうこともあった。特に、ネクタイをした人物の首周り、制服姿の腰周辺、円の上半分などを顔と間違えて切り出してしまうことが多かった。切り出された顔画像は全部で 38,650 枚あった。これらすべての画像が顔として正しいかどうか調査することは困難なので、ランダムに選んだ 100 枚が顔として正しいかどうか調べたところ 78 枚が顔であった。つまり顔の切り出しに対する適合率は 78%と推定される。



図 3 顔画像の切り出し結果 (失敗例)

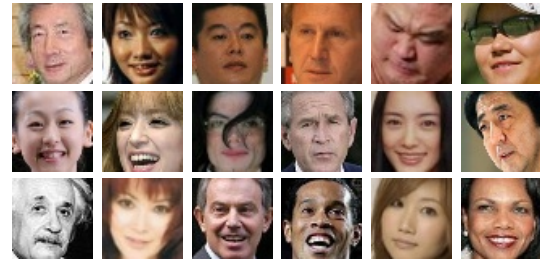


図 4 顔画像の切り出し結果 (成功例)

5.3.2 人物名の抽出結果

ニュース記事から人物名を抽出した結果を次に示す。
人物名の抽出結果

成功例

小泉純一郎, 松井秀喜, ブッシュ米大統領,
ボビー・バレンタイン など

失敗例

沢一郎, 戸田恵, 浩一・鈴木 など

失敗例は、逢沢一郎, 戸田恵梨香, 齋藤浩一・鈴木義純の誤抽出である。抽出された人物名の種類は全部で 13,579 種類であった。顔画像と同様にランダムに 100 種類選んで人物名として正しいかどうか調べたところ、78 種類が人物名として正しく抽出されていた。つまり人物名抽出の適合率は 78%と推定される。

5.4 対応付け結果

表 1 は、全対応付け結果とクラスタ画像が 3 枚以上の対応付け結果に対するランダムサンプリング 100 個の評価である。閾値 (1)~(4) は、実験方法で述べた値である。また、対応付け結果個数が多かった上位 5 人の人物名に対するランダムサンプリング 100 個の評価を表 2 に示す。図 5 に対応付け結果が最もよかった閾値 (4) における、対応付け個数上位 5 人の結果を載せる。図 5 のそれぞれの適合率は表 3 の通りである。図 5 の安倍晋三の対応付け結果は適合率 100%となっているが、これは同じようなパターンの顔写真が多いためと、表情が乏しいためにうまくいったものと思われる。また、小泉純一郎と小泉首相の 2 つの人物名クラスタが存在するのは、1 つのニュースから小泉純一郎しか抽出されない場合と小泉首相しか抽出されない場合があるからである。図 6 はクラスタ画像が 3 枚以上の結果の対応付け個数と適合率のグラフである。

6. 考 察

表 1 より、顔と人物名の対応付けの結果は最低で適合率 32%、閾値処理を施すと最高で 62%の適合率を得た。しかし、当然のことながら閾値を厳しくするにつれて対応付け結果画像枚数は少なくなってしまう

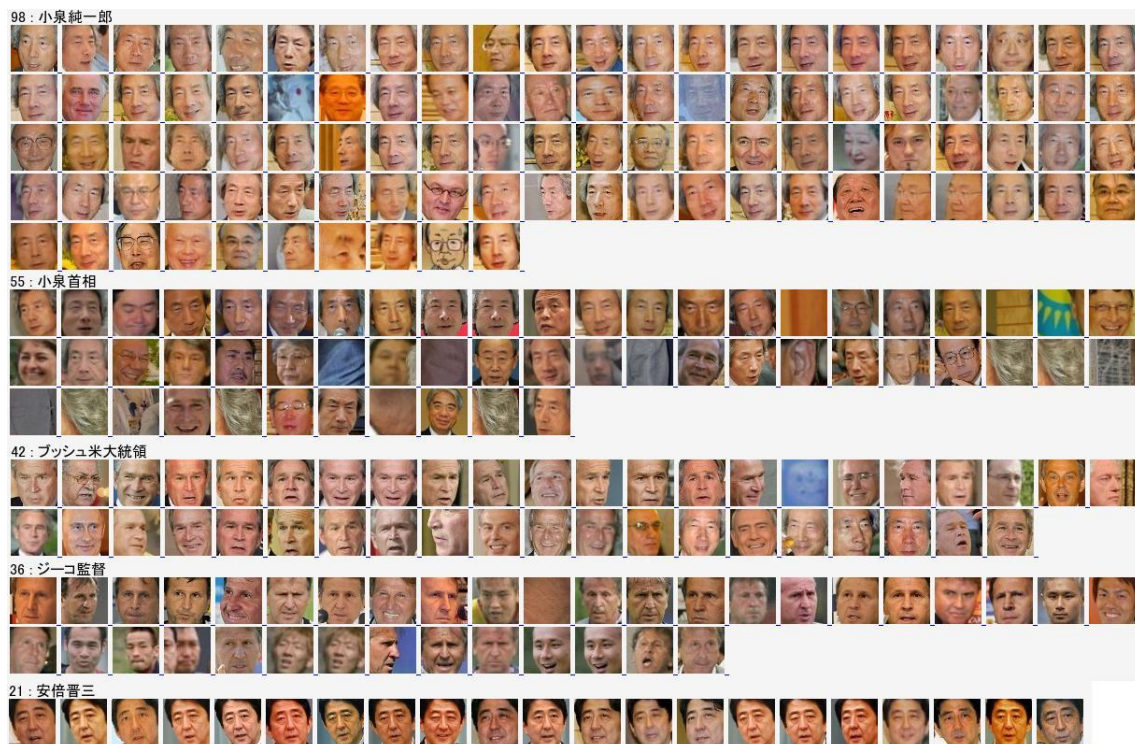


図 5 対応付けの様子

表 1 対応付け結果

閾値	全対応付け		3 枚以上	
	個数	適合率	個数	適合率
(1)	9329	34%	9108	32%
(2)	6480	42%	6027	46%
(3)	3853	40%	3125	44%
(4)	1281	54%	379	62%

表 2 閾値別上位 5 人の結果

閾値	対応付け個数	正解数	適合率
(1)	1420 個	729 個	51%
(2)	1195 個	673 個	56%
(3)	546 個	892 個	61%
(4)	252 個	165 個	65%

表 3 図 5 の適合率

人物名	適合率
小泉純一郎	72.4%
小泉首相	40.0%
ブッシュ大統領	64.3%
ジーコ監督	66.7%
安倍晋三	100%

た。最高適合率が得られた閾値 (4) では、対応付け枚数上位 21 位の人物名において、すでに 3 枚しか顔画像が存在しなかった。また全対応付け結果に対する評

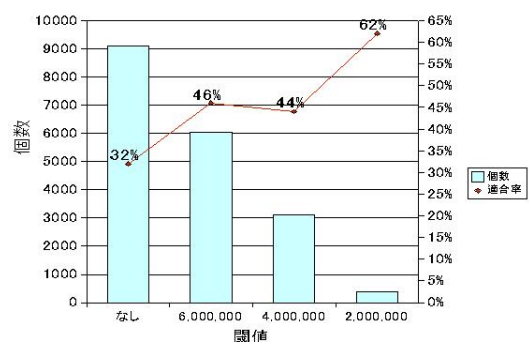


図 6 各閾値における適合率と個数の関係

価よりも 3 枚以上のクラスタに対する評価の方が全体的に上回っている。このことから、対応付け結果の少なかったクラスタでは対応付けの精度が悪くなってしまっていると考えられる。本研究の対応付け段階で用いた固有顔と k-means 法によるクラスタリングという手法は、T.L.Berg らによって英文ニュースではうまくいくことが示されている。対応付け段階は言語の違いは関係ないため、今回の実験結果が Berg らよりも悪かった原因は対応付け段階の前の抽出段階にあったと思われる。以下で抽出段階での問題点について考える。

6.1 顔抽出における問題

1. 存在する顔を切り出せない
ニュース画像には顔領域が存在するのに、切り出せないということがあった。本研究では正面顔のみを顔として抽出しているため、横顔などが抽出されずに残ってしまった。すべての顔と人物名を対応付けするためには、これらの顔も抽出しなければならない。OpenCV を利用して横顔を抽出することは可能であるが、本研究で用いた画像特徴量である固有顔は正面顔に対してしか有効であることが示されていない。そのため、横顔も含めてあらゆる方向に対応した画像特徴量、つまり方向に依存しない画像特徴量を固有顔の代わりに利用しなくてはならない。
2. 抽出した顔が傾いている
OpenCV で切り出してきた顔画像は、正面顔とはいってもある程度傾きをもったものも含まれており完全な横顔も少数あった。そのため、正面顔に対して有効な固有顔法では期待するほどの精度が得られなかったと思われる。よって、切り出された顔画像を正面顔の向きに修正する操作が必要になってくる。また、固有顔以外の顔の向きにとらわれない顔認識方法の適用も検討課題である。

6.2 人物名抽出における問題

本研究では苗字+名前、苗字+役職、苗字+国+役職、名前+・+苗字の4通りしか人物名を抽出していないため、抽出できない人物名があった。特に、以下の3つの場合が抽出できなかったものとして挙げられる。

1. 一般的でない苗字や名前
茶釜は辞書に登録された苗字や名前しか判別できないため、茶釜の辞書に載っていないような、あまり一般的ではない苗字や名前が抽出できなかった。また、これらの一般的でない苗字や名前の一部が、茶釜の辞書に登録されている別の苗字や名前だったりすると、誤抽出してしまった。他にも、外国人の人物名は茶釜に登録されている苗字と名前が少ないため、あまり抽出することができなかった。特に中国人や韓国人は漢字表記、カタカナ表記ともにほとんど抽出できず、誤抽出も多かった。
2. 役職名
本研究では役職名として16個を抽出することになっているが、実際にニュース記事に現れる役職名は多岐にわたり、役職名が抽出できないために「苗字+役職」の形でしか現れない人物名を抽出できなかったことがあった。また、本研究では苗字の次の単語を役職名としているため、「厚生労働大臣」や「文部科学大臣」のように複数の名詞からなる役職名を抽出することができなかった。今後、これ

ら複数名詞からなる役職名を含め、より多くの役職名を抽出できるようにしなければならない。

3. 苗字だけの人
本研究の抽出方法では、苗字だけしか記事に出てこない人は抽出できない。しかし、スポーツ選手などは苗字だけしか記事に現れないことが多いため、このような苗字だけの人物名を抽出することはできなかった。

これらの人物名が抽出できないために、対応付け段階において、顔は抽出されたにもかかわらず対応する人物名がなくて対応付けできなかったり、別の人物名に間違えて対応付けてしまったりするミスが発生した。こうした問題に対応するべく、今後は茶釜の辞書追加や、自然言語処理を行うことで人物名の抽出精度向上を図る必要がある。

7. おわりに

本研究では、約1年半分の写真画像付き Web ニュースに対して、画像内の顔と記事内の人物名を抽出し、それらの対応付けを行った。顔抽出には OpenCV を用い、38,650 枚の顔画像を適合率 78% で抽出した。人物抽出には茶釜を利用し、こちらも適合率 78% で 13,579 種類の人物名を抽出した。これらの抽出結果を使って、画像特徴として固有顔法を用いた k-means クラスタリングにより、顔と人物名の対応付けを行った。様々な閾値処理を加えた結果、62% の適合率を得た。英文ニュースにおける同様の研究の結果に比べると精度が悪い原因として、顔・人物名の各抽出の精度が悪いことが考えられる。今後の課題としてこれらの抽出精度向上が挙げられる。

参考文献

- 1) Turk, M. and Pentland, A.: Face recognition using eigenfaces, *Proc. of Computer Vision and Pattern Recognition*, pp.586–591 (1991).
- 2) Viola, P. and Jones, M.: Rapid object detection using a boosted cascade of simple features, *Proc. of Computer Vision and Pattern Recognition*, Vol.1, pp.511–518 (2001).
- 3) Open Source Computer Vision Library: <http://www.intel.com/technology/computing/opencv/>.
- 4) Berg, T., Berg, A., Edwards, J., Maire, M., White, R., Teh, Y., Learned-Miller, E. and Forsyth, D.: Names and Faces in the News, *Proc. of Computer Vision and Pattern Recognition*, pp.848–854 (2004).
- 5) 坂野鋭: パターン認識における主成分分析顔画像認識を例として, 統計数理, Vol.49, No.1, pp. 23–42 (2001).
- 6) 茶釜: <http://chasen.naist.jp/hiki/ChaSen/>.