

## 写真ニュースにおける「話題顔」の検出

杉山 一成† 山田 剛一† 与那嶺 靖典† 中川 裕志‡

†横浜国立大学 工学部

‡東京大学 情報基盤センター

### 1 はじめに

画像、映像データベースなどマルチメディア情報の普及に伴い、その大部分を占めると考えられる人物(顔)を検索、編集などに役立てていく技術は、重要性を増してきている。この前処理として必要になるのが、シーン中から顔パターンを検出することである。従来のシーン中からの顔検出に関する研究は、[1],[2],[3],[4]など数多く行なわれているが、これらは画像中の顔をもれなく検出することを目的としている。また、ある程度の大きさの、主に正面向きの顔を扱っており、あらゆる場面に対応できるとは言い難い。例えば、新聞記事などの写真は、その場の様子をありのままに写したものであるため、顔の大きさや向きは様々である。また、記事の内容とは無関係な人物が写っていることもしばしばである。写真中の顔をクリックした時に、その人物と対応付けられた記事を検索するようなシステムができれば、ユーザはその人物に関して多くの情報を得られ、記事データベースの有効利用ができると考えられる。そこで、図1のように記事と写真から構成される毎日新聞写真ニュース[5]を対象とし、記事中に名前が出ていそうな顔の特徴をとらえ、検出する手法を提案する。本研究では、そのような顔を「話題顔」と呼んでいる。

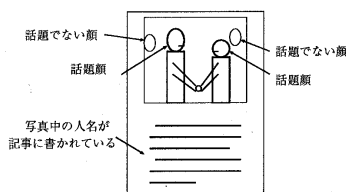


図1: 対象とする新聞記事

### 2 顔候補領域の抽出

本研究では、顔の検出において向きや遮蔽に比較的強いとされる色情報を利用する。サンプルとして85

“Topical Face Detection in Photograph Newspaper”  
Kazunari Sugiyama, Koichi Yamada, Yasunori Yonamine, Hiroshi Nakagawa: Faculty of Engineering, Yokohama National University

人分の顔画像から頬の肌色を抽出する。肌色は  $HSV$  空間の正規分布に従うと仮定し、これらのサンプルから濃度平均値  $M(H, S, V)^T$ 、分散共分散行列  $C$  を求め、入力画像の濃度値を  $I(H, S, V)^T$  として、

$$d^2 > (I - M)^T C^{-1} (I - M)$$

を満たす画素を未知画像から抽出する。  $d$  は実験によって定めた。

### 3 「話題顔」の検出

#### 3.1 ヒューリスティックなルールの適用

情報を伝えるという新聞記事の役割上、写真中における「話題顔」は読者に注目される存在である。すなわち、位置と大きさに関してなんらかの特徴をもつと考えられる。そこで、学習データから「話題顔」の位置と面積に関するヒストグラムを作成し、この結果から次のようなルールを適用した。なお、対象としたデータの写真の大きさが一定でないので、面積、位置ともに正規化してある。

ヒューリスティック 写真全体の面積の0.5%より大きく、重心の  $y$  座標が  $2/3$  以上の領域を抽出

このルールを適用したところ、再現率、適合率は、それぞれ93.9%、42.4%であった。すなわち、手や背景など余計な領域を検出していることが確認された。そこで、抽出領域に関して4節で述べる特徴パラメータを計算し、「話題顔」の特徴を学習によりとらえることとした。

### 4 各候補領域の属性

2節で抽出した各領域について、色、位置、形状、角度、対称性、画像の構図、面積からなる28個の特徴パラメータを学習の属性として用いた。

- 色、形  
 $R, G, B$  と  $Y$  の平均値
- 位置  
重心の  $x, y$  座標、画像全体を9分割したときの領域の位置

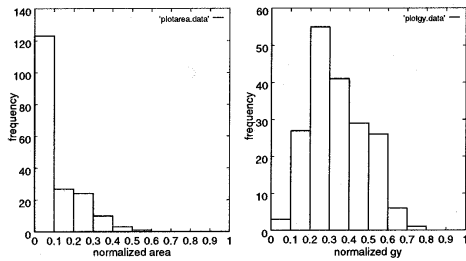


図 2: 面積、位置のヒストグラム

- 形状  
縦横比、矩形度、楕円度、円形度
- 角度  
抽出領域の上下左右の端点を頂点とする四角形の  
内角
- 対称性  
抽出領域の上下と左右の対称性を表すパラメータ
- 画像全体の構図、面積  
画像全体の縦横比、画像の中心、各頂点および各  
辺の中心から抽出領域までの距離、画像全体に対  
する面積比

#### 4.1 各学習法による識別性能

再現率、適合率を以下のように定義し、決定木、パーセプトロン、ニューラルネットを用いて識別性能を比較した。しかし、4節で導入した特徴には、相関の高い特徴の組が含まれている可能性がある。そこで、主成分分析によって特徴選択を行なった結果、16個の特徴で累積寄与率が99%に達することが確認された。特徴選択前後の結果を、それぞれ表1、2に、また、原画像と検出例を図3に示す。

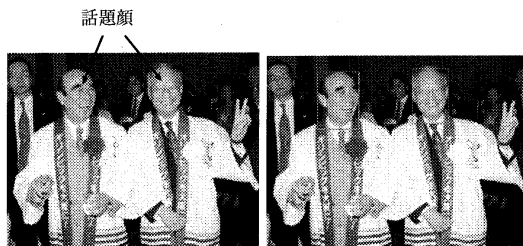


図 3: 原画像(左)と話題顔検出例(右)

## 5 まとめ

写真付き新聞記事からの「話題顔」検出法を提案した。この写真ニュースを用いて、記事から写真中に顔

表 1: 「話題顔」検出精度の比較 (特徴選択前)

	再現率	適合率
C4.5	93.0%	72.5%
C5.0	87.4%	84.9%
パーセプトロン	63.8%	63.8%
ニューラルネット (3層 3個)	86.7%	95.0%
ニューラルネット (3層 4個)	88.5%	95.1%
ニューラルネット (3層 5個)	89.3%	95.1%

表 2: 「話題顔」検出精度の比較 (特徴選択後)

	再現率	適合率
C4.5	84.0%	80.8%
C5.0	84.0%	87.5%
パーセプトロン	61.4%	61.7%
ニューラルネット (3層 3個)	87.8%	95.0%
ニューラルネット (3層 4個)	88.5%	95.1%
ニューラルネット (3層 5個)	89.3%	95.1%

の写っていきそうな人名の候補を抽出し、本システムによる「話題顔」候補との対応付けを行ない、メディア間の意味内容を解析した人物検索システムを構築中である。これについては、文献[6]を参照されたい。また、本システムを拡張することで、動画像において、数ある人物シーンの中から主要なシーンを選び出すことも可能であると考えられる。

## 謝辞

充実した写真ニュースを Web 上で発信している毎日新聞社に敬意を表するとともに感謝いたします。

## 参考文献

- [1] Venu Govindaraju, Sargur N. Sruhari, and David Sher. A computational model for face location based on cognitive principles. *In Proceedings of the National Conference on Artificial Intelligence(AAAI-92)*, Vol. 1, pp. 350-355, 1992.
- [2] K. Sung and T. Poggio. Example-based learning for view-based human face detection. *A.I. Memo 1521, MIT A.I. Lab*, 1994.
- [3] H. Rowley, S. Baluja, and T. Kanade. Human face detection in visual scenes. *Technical Report CMU-CS-95-158R, School of Computer Science, Carnegie Mellon University*, 1995.
- [4] Edgar Osuna, Robert Freund, and Federico Girosi. Training support vector machines: an application to face detection. *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 130-136, 1997.
- [5] 毎日新聞社. Aulos 写真ニュース (現在、毎日インタラクティブ <http://www.mainichi.co.jp/>), 1997.
- [6] 山田剛一, 杉山一成, 与那嶺靖典. 写真ニュースにおける顔領域と人名との自動対応づけ. *人工知能学会全国大会論文誌*, pp. 201-204, 1999.