

低解像度画像からの歩行者顔方向推定

平田 純子 森本 雅和[†] 藤井 健作

兵庫県立大学大学院工学研究科 〒671-2201 兵庫県姫路市書写 2146

E-mail: [†] morimoto@eng.u-hyogo.ac.jp

あらまし 顔パーツの特定が困難な程度の、低解像度の顔画像から顔方向を推定するシステムを提案する。顔方向推定はサポートベクターマシンを用いる。顔データベースを用いた実験結果より、6×6ピクセルの入力顔画像に対して15度刻み13方向の顔方向を90%以上の推定率で推定することが確認できた。実際に撮影した動画から切り出した顔画像に対する実験結果より、顔切り出し領域は首から上で頭部輪郭が画像内に収まる大きさが、低解像度画像では適切であることを確認した。また、そのときの平均推定誤差は34.3度、サンプル画像の50%以上が誤差15度以下で推定できることを確認した。
キーワード 顔方向推定、低解像度画像、人物行動認識、歩行者検出、SVM

ESTIMATING FACIAL DIRECTION FROM LOW RESOLUTION IMAGES

Junko HIRATA Masakazu MORIMOTO[†] and Kensaku FUJII

Graduate School of Engineering, University of Hyogo 2167, Shosya, Himeji-shi, Hyogo, 671-2201 Japan

E-mail: [†] morimoto@eng.u-hyogo.ac.jp

Abstract In this paper, we propose a facial direction estimating system from low resolution facial images captured by surveillance camera. The facial direction is estimated by using support vector machine (SVM). Experimental results show that, when we use 6x6 pixels facial database images as input, we can estimate 13 facial directions in increments of 15 degrees with accuracy more than 90%. Experimental results of surveillance camera images shows that, the appropriate size for low resolution images is from neck up and head outline fit into image. In that case, average estimate error is 34.3 degree and more than half images can be estimated within 15 degree errors.

Keyword Facial direction estimation, Low resolution images, Human behavior recognition, Pedestrian detection, SVM

1. はじめに

監視カメラは銀行の出入口や商店のレジ付近に設置され、不審者の監視、侵入禁止区域への侵入者の検出、犯罪抑制、犯人検挙などに利用されている。この監視カメラ映像を計算機により処理し、人物の行動を自動的に監視する研究や製品開発が広く行われている。これまでの研究開発では、侵入者の自動検出が主な目的とされており、遮蔽物や交差に対する人物追跡の頑健化や人物以外の移動物体（自動車等）の識別等、人物追跡の精度向上に関する研究が広く行われている。

また、監視カメラ映像を用いた人物行動認識に関する研究も行われており、過去の監視映像の統計的情報を用いた異常行動の検出や、マーケティング分野への

応用として商店における買物客の陳列棚付近の滞留時間の調査等の報告がある。

現在、人物行動監視の分野における挑戦的課題として「善人の集団に紛れた、善人の外見をした不審人物の検出」つまり「不審者が犯罪行為を起こす前にそれを阻止する」ことが求められている。これは事後検出ではなく事前予知であり、熟練した監視員によって行われる業務となっている。広域監視カメラ画像から不審人物を割り出すために、不審者をPTZカメラで追跡し詳細情報を取得する研究が行われている。しかし、この事前予知の自動化を実現するためには従来の人物追跡の解析以外に、人物の振る舞い（目や頭の動き、移動軌跡）を数値化し、それらの情報を総合的に判断

する必要がある。そこで、本報告では事前予知のために必要となる要素の一つである人物の顔方向推定に注目する。

従来の顔方向推定手法として、高解像度顔画像における目・鼻・口等の部品の配置から顔の 3D モデルを仮定し、正確な顔向きを推定する手法が研究されている[1]-[4]。しかし、一般的な監視カメラは常に広い範囲を監視するために広視野角の映像を撮影することが多く、顔部品検出を利用する従来法を適用するためには複数の監視カメラが必要となる。また、顔部品の検出が可能となる解像度の顔画像では、個人の特定が容易であり、マーケティング等の応用ではプライバシー保護の問題が発生する。

そこで本報告では、低解像度の人物画像から顔方向を推定することを目的とする。低解像度の人物画像から顔方向を推定することができれば、一台のカメラでより広い範囲の監視を行うことができるため、既存の監視カメラをそのまま利用することができる。さらに、顔部品検出が困難な解像度の顔画像からは個人の特定も難しいため、プライバシーを保護する必要がある場合にも利用できる。ここでは、低解像度の顔画像から顔方向を推定するために学習機械を利用する。提案手法では、サポートベクターマシン (support vector machine : SVM) [5]を用いて顔方向の判別を行う。SVM は最も優れたパターン識別手法の一つであり、画像認識、文字認識、顔検出等に広く用いられる[6]-[8]。

2 章では、低解像度画像からの顔方向推定に用いる SVM について述べる。3 章では、提案する顔方向推定の有効性を確認するための実験を行う。3.1 節で顔データベースを用いた実験結果、3.2 節で撮影した動画から切り出した画像を用いた実験結果、3.3 節で撮影した動画から異なる大きさで切り出した顔画像による推定誤差比較実験結果について述べる。最後に 4 章でまとめと今後の課題について述べる。

2. 低解像度画像からの顔方向推定

本報告では、サポートベクターマシンを用いて顔方向の推定を行う。頭部の検出として、動画から背景差分による移動物体の検出と、平均顔テンプレートを用いた頭部切り出しを行う。切り出した画像は背景差分を取る際にマスクをかける。本報告では、これらの処

理が正確に行えたという前提の下で顔方向推定を行う。

2.1. SVM を用いた顔方向推定

本報告では SVM を用いて、顔方向推定を行う。SVM は Vapnik によって提案されたマージンの最大化に基づく 2 値分類器であり、学習モデルの中で最も認識性の優れた手法の一つとして知られている。“カーネルトリック”と呼ばれる方法を用いて、線形分離不可能なデータ集合に対して、特徴ベクトルを非線形変換して高次の特徴空間に写像し、その空間で線形識別を行うことで非線形関数による分離が可能となる。

SVM では、超平面とその超平面に最も近いデータ間の距離 (マージン) を最大化する識別平面を求める。分離超平面とは K 係数一次多項式の解空間として定義される K^n の $n-1$ 次の部分線形空間、つまりデータを分離する部分空間のことである。このようにして SVM はカーネル関数を用いて特徴ベクトルを高次元にマップし、この空間における最適な分離超平面を形成する。学習データを $(\mathbf{x}_i, y_i), i=1, \dots, l$ where $\mathbf{x}_i \in R^n, y_i \in \{1, -1\}^l$ とすると、SVM は以下の最適化問題を解く必要がある。

$$\min_{\mathbf{w}, b, \xi} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^l \xi_i$$

$$\text{subject to } y_i (\mathbf{w}^T \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i, \quad (3) \\ \xi_i \geq 0.$$

学習ベクトル \mathbf{x}_i は関数 ϕ を用いて高次元空間にマップし、2 クラス分類器の場合に両クラスの分離超平面に最も近い特徴ベクトルのマージンが最大であれば最適となる (図 1)。超平面に最も近い特徴ベクトルは「サポートベクター」と呼ばれ、その他のベクトルの位置が超平面に影響を及ぼさないという事を意味している。多クラス分類に使用する場合は、2 クラス分類の SVM を組み合わせることによって実行できる。C はエラー時のペナルティパラメータ、 $K(\mathbf{x}_i, \mathbf{x}_j) \equiv \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$ はカーネル関数であり、代表的なものにガウス関数を用いた動径基底関数カーネル (Radial Basis Function : RBF)

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2) \quad (6)$$

がある。ここで、 γ はカーネルパラメータである。カーネル関数は超空間における 2 点間距離を計算する。

ここで、SVM 学習の実装は OpenCV[9] と LIBSVM[10]で行う。顔方向ごとにクラス分けするため、

SVMを多クラス分類に使用する。多クラス分類は以下のような原理となる。

1. 複数ある顔方向クラスから任意の2クラスを選び、推定したい画像がどちらのクラスに近いかによりポジティブクラスを選択する。他方はネガティブクラスとラベル付けする。
2. ポジティブクラスと1.でネガティブクラスのラベル付けがなされていないクラスを1.と同様に比較し、全てのクラスに対してポジティブ・ネガティブクラスのラベル付けを行う。最終的にポジティブクラスとなったクラスが推定結果となる。

また、SVMの入力ベクトルとして、本報告では画像の画素値を用いる。事前実験として、色空間変換(HSV, YCrCb, CIE XYZ, CIE L*a*b*)した画像の画素値、離散フーリエ変換した画像の画素値、グレースケール画像の画素値を入力ベクトルとして顔方向推定を行った。推定結果を比較して、最も結果が良かった HSV 色空間変換後の画素値を本報告の実験における入力ベクトルとする。

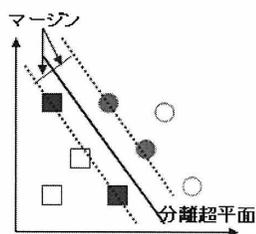


図1 マージン最大化

3. 推定実験

3.1. 顔データベースによる実験結果

2.1 節の顔方向推定手法の有効性を確かめるために財団法人ソフピアジャパン[12]提供の顔画像データベース*で実験を行う(図2)。顔データセットは男女150人ずつ計300人のカラー画像で、図3に示すように各人-90から90度までを15度刻みに13の顔方向を

もつ。学習は150人(男女75人ずつ)、推定は学習に用いていない150人で行う。画像は図4に示すように6種類(20×22, 14×15, 10×11, 8×9, 6×6, 4×4pixels)のサイズに縮小し、擬似的に低解像度画像とし、学習・推定画像の解像度は等しくする。推定する顔方向数は30度刻み7方向, 15度刻み13方向の2種類で行う。SVMの入力ベクトルはHSV色空間変換後の画素値, カーネル関数パラメータは, $C=10$, $\gamma=0.09$ で行う。あらかじめ画像に定められている方向と推定した方向が一致した場合に, 正しく推定でき誤差が0度だとする。

図5に7, 13方向推定の学習・推定画像解像度を等しくした場合のみの結果を示す。20×22から6×6の範囲では推定率90%以上と高い推定率を得ることができた。しかし, 4×4においては, 7, 13方向共に推定率70%以下であり, あまり高い結果ではない。ここで, 図4(e)の解像度6×6の顔画像を参照すればわかるとおり, 目や鼻等の顔のパーツの位置を特定することが難しい。それゆえ, 解像度6×6の顔画像に対して90%以上の顔方向推定率を得ることができれば, 低解像度という点において十分であると言える。



図2 顔データベース画像(60×65)

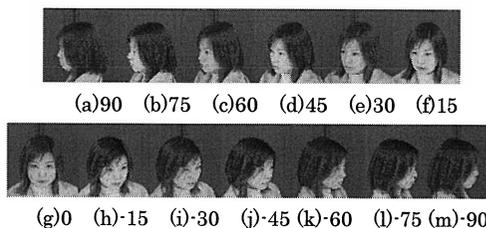
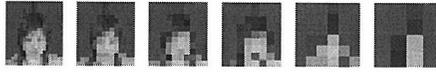


図3 15度刻み13方向の顔画像

* 本報告に使用した顔画像データは, 財団法人ソフピアジャパンから使用許諾を受けたものである。権利者に無断で複写, 利用, 配布等を行うことは禁じられている。



(a)20×22 (b)14×15 (c)10×11 (d)8×9 (e)6×6 (f)4×4

図4 低解像度顔画像

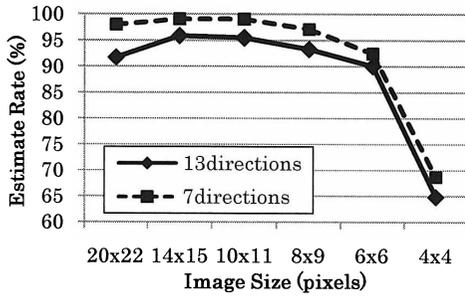


図5 7,13 方向推定結果

3.2. 撮影動画から切り出した顔画像に対する実験結果

実際に撮影した動画から切り出した顔画像に対して実験を行う。撮影カメラは、解像度は 640×480pixels(VGA)である。顔画像は 266 人のカラー画像でサイズは 20×20pixels。学習には 3.1 節で用いた顔データセット 300 人を用い、学習・推定画像共に図 7 に示すように 4 つの解像度(14×14, 10×10, 5×5, 4×4pixels)に縮小して顔方向を推定する。推定する顔方向数は 13 とする。

解像度を変えて実験した場合の推定誤差を図 8 に示す。20×20 から 5×5 にかけて解像度が低くなるにつれ推定誤差が小さくなっているが、4×4 では誤差が大きくなった。低解像度化することで推定誤差が小さくなることに関して、入力次元数に対して推定するサンプル数が少ないことが原因だと考えられる。

また、図 9 にそれぞれの解像度における、誤差分布を示す。凡例は画像の誤差である。正しく推定できた正解画像は全体の 20%程度である。また、正解角度の 1 つ隣、つまり誤差 15 度の画像と推定した画像と、正解画像の割合を足すことで全体の 50%程度になるが、3.1 のデータベースを用いた実験結果と比較して、誤差が大きい。

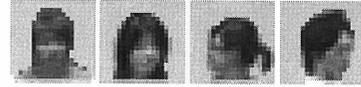
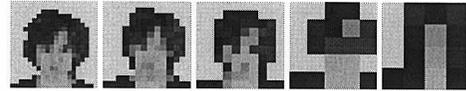


図6 撮影した動画から切り出した顔画像



(a)20×20 (b)14×14 (c)10×10 (d)5×5 (e)4×4

図7 推定する解像度

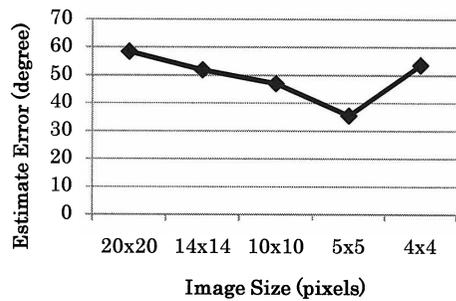


図8 実画像顔方向推定結果

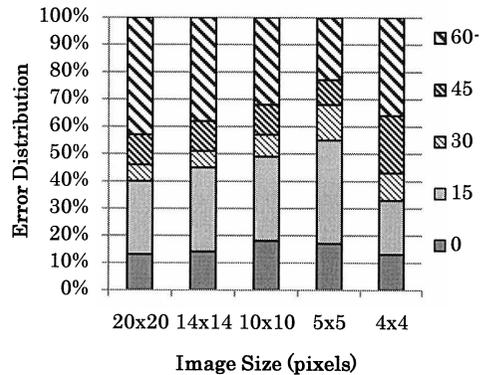


図9 解像度毎の誤差分布

3.3. 切り出した顔画像の絞り込みによる実験結果

3.2 節で使用した実画像は、首から下の服の部分を含む。それゆえ、服の色による影響が考えられるので、首から上の顔領域のみを切り出し修正することで顔の絞り込みを行う。図 7(a)の 20×20pixels の画像に対し

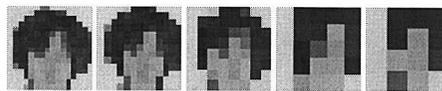
て図 10(a)のように 16×16pixels で顔の絞り込みを行う。さらに解像度を図 10(b)-(d)に示すように縮小して低解像度化し実験を行う。同様に、顔の絞り込みを 14×14pixels でも行う。

実験結果を図 12 に示す。凡例は、元画像を縮小した場合の結果を 20×20 に、絞り込み画像 16×16 を縮小した場合の結果を 16×16 に、絞り込み画像 14×14 を縮小した場合の結果を 14×14 に示してある。元画像 20×20 での推定誤差は 58.4 度であり、顔を絞り込んだ画像で縮小する前の推定誤差がそれぞれ、51.2、52.1 度であることから顔を絞り込むことで 6~7 度改善できることが確認できた。また、絞り込み画像 16×16 の場合、解像度を 5×5 まで落とすことで推定誤差が更に改善されたが、14×14 では解像度を落とすことで推定誤差が大きくなった。図 6(a)、図 10(a)、図 11(a)を比較すると、絞り込み画像は首から下の服の部分を含まないため推定誤差が改善されたと考えられる。また、16×16 の顔画像は髪型を含めて顔の輪郭が画像内に収まるが、14×14 の顔画像は髪の毛が部分的に画像内に収まらないため、14×14 の顔画像は低解像度化しても推定誤差が改善されなかったと考えられる。

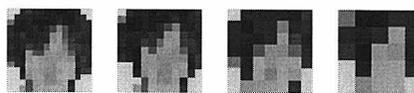
図 13 に絞り込み画像 16×16 を低解像度化した場合の誤差分布を示す。図 8 の絞り込む前の元画像を低解像度化した場合の誤差分布と比較して、誤差 15 度の画像と推定した画像と、正解画像の割合を足した割合は、5%程度の改善が確認できた。また、図 13 において 5×5 では誤差 0 度の正解画像の割合が大幅に増えている。よって推定誤差が最も小さくなったと考えられる。

図 14 に絞り込み画像 16×16 を 8×8 に縮小した場合の正解角度毎の認識率を示す。凡例は正解角度と推定角度の誤差を表す。誤差 0 度、誤差 15 度の画像を正しく顔方向推定ができたと仮定すると、正解角度が±90 度に近づくほど顔方向が推定できていることが確認できる。つまり、真正面よりも横顔のほうが正しく推定できる。

図 15 に絞り込み画像 16×16 を 8×8 に縮小して実験し、推定誤差が 45 度以上の画像を示す。髪型、照明条件、帽子の着用等の影響で正しく推定できなかったと考えられる。



(a)16×16 (b)13×13 (c)8×8 (d)5×5 (e)4×4
図 10 顔絞り込み画像と縮小画像(16×16)



(a)14×14 (b)11×11 (c)7×7 (d)5×5
図 11 顔絞り込み画像と縮小画像(14×14)

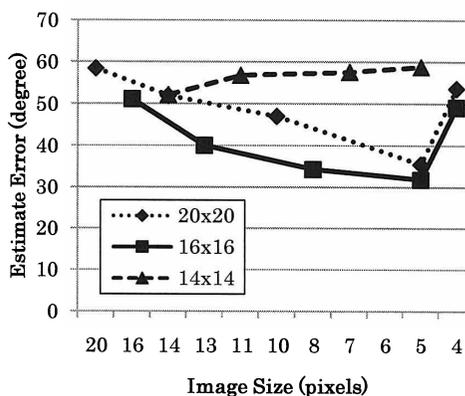


図 12 絞り込みによる推定誤差の変化

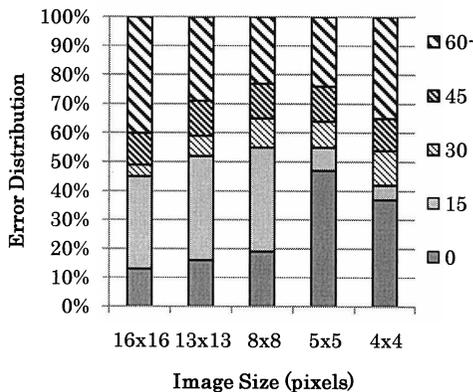


図 13 絞り込み画像 16×16 を低解像度化した場合の誤差分布

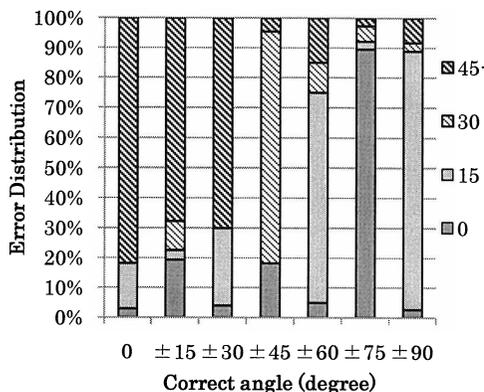


図 14 切り出し画像における正解角度毎の認識率

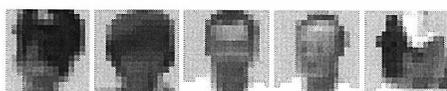


図 15 推定誤差の大きい画像(16×16)

4. まとめと今後の課題

SVM を用いた低解像度映像から歩行者の顔方向を推定する手法を提案した。

顔画像データベースを用いた顔方向推定実験により、解像度 6×6pixels の学習・推定画像で 15 度刻み 13 方向の推定において 90%の推定率を得ることができ、SVM を用いた顔方向推定手法の有効性を確認できた。

また、実際に撮影した動画から 20×20pixels で切り出し、10×10pixels に縮小して低解像度化した顔画像を用いた実験により、推定誤差は 47.0 度となった。顔領域の切り出しを頭部形状が画像内に収まるように 16×16pixels で切り出した場合、縮小して 8×8pixels に低解像度化した画像での推定誤差が 34.3 度となった。しかし、誤差 0 度で正しく推定できたのは画像数に対して 20%程度であり、顔データベースを用いた実験結果の 90%の推定率と比較して、大幅に推定率が下がる。

今後の課題として、推定精度の改善、髪型や帽子的着用等データベースにない条件の画像で正しく顔方向を推定できない画像を、「判別不能」と判定すること、動画からの連続画像取得を生かした顔方向推定として、前後フレームとの比較による推定等が挙げられる。

文 献

- [1] 荒木祐一, 島田伸敬, 白井良明, “背景と顔の方向に依存しない顔の検出と顔方向の推定,” 信学技報, PRMU2001-217, pp. 87-94, 2002.
- [2] 林健太郎, 橋本学, 鷺見和彦, “頑健性と精緻性を備えた顔特徴点追跡による顔方向推定,” 信学論 D-II, Vol. J84-D-II, No. 8, pp1762-1771, 2001.
- [3] Shigenori ISHII, Daisaku ARITA and Rin-ichiro TANIGUCHI, “Real-time Head Pose Estimation with Stereo Vision,” Proceedings of the 9th Korea-Japan Joint Workshop on Frontiers of Computer Vision, pp. 79-83, 2003.
- [4] Yuxing Mao, Suen C.Y, Caixin Sun and Chunhua Feng, “Pose Estimation Based on Two Images from Different Views,” Proceedings of the Eighth IEEE Workshop on Applications of Computer Vision, 2007.
- [5] Christopher J. C. Burges, “A Tutorial on Support Vector Machines for Pattern Recognition,” Data Mining and Knowledge Discovery 2:121-167, 1998.
- [6] Takeshi Matsuyama and Hiroshi Nakagawa, “Two Step POS Selection for SVM Based Text Categorization,” IEICE TRANSACTIONS on Information and Systems, Vol.E87-D, No.2, pp.373-379, 2004.
- [7] Wonjun HWANG and Hanseok KO, “Image Feature Extraction Algorithm for Support Vector Machines Using Multi-Layer Block Model,” IEICE TRANSACTIONS on Information and Systems, Vol.E86-D, No.3, pp.623-632, 2003.
- [8] Tae-Kyun KIM and Josef KITTLER, “Composite Support Vector Machines with Extended Discriminative Features for Accurate Face Detection,” IEICE TRANSACTIONS on Information and Systems Vol.E88-D No.10 pp.2373-2379, 2005.
- [9] “OpenCV”, Software available at <http://sourceforge.net/projects/opencvlibrary/>.
- [10] Chang C.-C., C.-J. Lin, “LIBSVM: a library for support vector machines”, Software available at <http://www.csie.edu.tw/~cjlin/libsvm>.
- [11] 財団法人ソフトピアジャパン, <http://www.softopia.or.jp/>.