

顔認証における前処理の検討

A study of pre-processing for face recognition

栗田 進
Susumu Kurita富川 武彦
Takehiko Tomikawa

1. まえがき

顔の認証における重要性は、その特徴をどう捉えるかにある。その一つに appearance-based methods が提案されているが、顔の向き、表情の変化、照明の違い等に抽出される特徴は大きく影響され、全てに対応できる「特徴の捉え方」が見つからないのが現状であろう[1]。本論文は顔認証において、顔画像の輝度情報を前処理することで照明の差による問題を低減し、認識率を向上させる方法を提案するものである。

照明を顔の左側から当てた場合、影が出来る、左側の輝度が高くなるといった問題が発生する。その結果、顔の特徴データが変化し、認識においては大きな問題となる。これを避ける方法として、画像の平均値、ヒストグラムを同じにする、或いは高域フィルターを用いる等の手法が提案されている。本提案はこれらと異なり、低域フィルターにて照明のシェーディング信号を抽出し、それにある係数を掛け算して元画像から差し引く事で顔の特徴量を残しながら照明の差を吸収しようとするものである。実画像でシミュレーションを行い、認識率の向上に有用であることを確認した。評価用に今回用いた認識アルゴリズムは顔画像を2次元フーリエ変換し、係数の内、有用なものを選択して線形判別分析を行う方法である。

2. 提案方法

顔画像は3次元情報が2次元に投影され、そこに輝度情報が加わっている。この2次元の情報の中で、輝度情報は意味があるのかと考えてみる。この情報は照明の輝度、角度、点光源か、面光源かといった条件に影響され、顔の輝度に変化し、特徴として抽出したフーリエ係数が変わってしまう。その変化が顔の違いか、照明の条件によるものかは区別できず、結果として誤認識の原因となる。丸顔、四角い顔、或いは目が大きいといった顔の形状こそが本来抽出すべきものであり、この輝度情報は認識の上では誤差要因となる(撮影条件を限定してしまえば、この輝度情報も重要な特徴量となる)。そこで我々はどのようにして顔の特徴量を変えずに輝度の影響を抑えるかに注目した。

$N \times M$ 画像の2次元のフーリエ変換は(1)式で表される。ここで f_{xy} は各画素の輝度値であり、照明条件が変われば変換係数である $F(u, v)$ は明らかに影響を受ける。

$$F(u, v) = \frac{1}{\sqrt{MN}} \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} f_{xy} e^{2\pi j \left(\frac{xu}{N} + \frac{yv}{M} \right)} \quad (1)$$

この照明条件の影響を減らすために以下の前処理を行う。顔画像に適当な低域フィルターをかけ、1以下のある係数をかけて元画像から差し引く。こうすることで、顔画像の

中の照明による影、照明のシェーディングを補正することができる。更に輝度情報の白黒の関係(例えば額は白、目は黒)を保存しながら不要な輝度変化を抑えるために、顔の中心部の輝度平均値に、ある係数を掛けたものを先の画像から差し引き、ある閾値を設けてそれを境に輝度を抑圧する(ゲインを変える)。これを式で表すと、提案する画像信号 I_{xy} は以下のようになる。

$$I_{xy} = (f_{xy} - f'_{xy} \times a - m) \times b + OS \quad (2)$$

ここで f_{xy} は元の信号、 f'_{xy} はシェーディング信号(元の信号に低域フィルターをかけたもの)、 a はシェーディングを補正するための1以下の係数、 m は画像の中心部の平均値にある係数を掛けたもの、 b は輝度を抑圧する係数、 OS は固定値のオフセットである。結果として図-1に示す入出力特性を実現し、照明の影響を軽減する。輝度情報に適当な係数を掛け算し、シェーディングを生じさせた画像、及び本提案法により処理された画像を図-2に、図2の矢印部の輝度信号を図-3に示す。

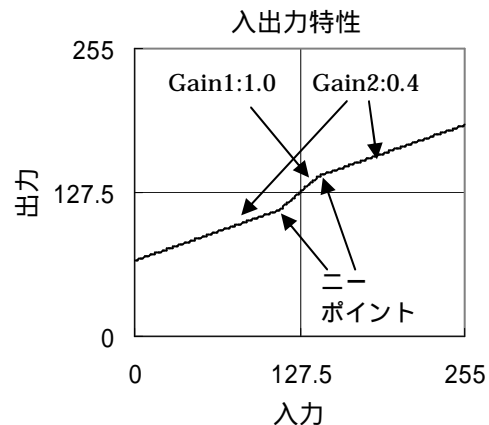


図 1 輝度情報を抑圧するための非線形入出力特性 (2) 式の b がこの入出力特性にあたる



図 2 輝度情報に適当な係数をかけることでシェーディングを生じさせた画像、及び本提案法による処理画像例。上列左から、元画像、1.6倍輝度画像、左右それぞれから照明された画像、部分的に照明された画像。元画像上の四角の枠は(2)式中の m を求めている部分を示す。下列はそれぞれの本提案の処理画像。

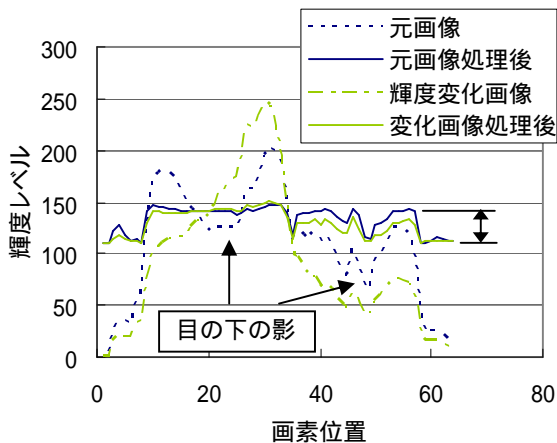


図 3 元画像、それに輝度変化を生じさせた画像及びそれら二つの処理後の信号例。輝度変化画像において 250 程の輝度を持つ部分が、まで圧縮され、処理後の二つの信号はほぼ同じとなっている。又元画像にある、目の下の不要な影がほぼ削除されている事が分かる。

以上の前処理により(2)式の OS 値に固定され、輝度レベルが抑圧された画像、図 2 が得られる。この画像を 2 次元フーリエ変換し、特徴量とした。

ここで(2)式の a を 1 とすると単なるハイパスフィルターの処理と等価になるが、本提案のポイントは a を 1 以下とすることで、顔の形、目、鼻、口を残し且つ、シェーディングを補正できるという点である。今回用いた各係数は $a=0.7$ 、 $m=0.15 \times$ 平均値、ゲインを変える二ポイントは ± 5 、ゲインはそれぞれ 1、0.4 である。これらは認識正解率が高くなるようシミュレーションにより決定された。

顔認識のアルゴリズムは様々な方法が提案されている。顔画像をそのまま主成分分析を行う Eigenface [2]、線形判別分析を用いる Fisherface [3]、ウェーブレット或いはフーリエ変換により特徴量を抽出し、多変量解析手段により認識を行おうとするものなどがある [4]。我々が今回用いたのはフーリエ係数のパワースペクトラムを線形判別分析する方法である。パワースペクトラムは原理的に顔の切り出し位置に依存せず特徴量を求めることが出来る利点があり、今回採用した [5, 6]。

3. シミュレーション方法

本提案の効果を確認するために 3 種類の画像を用いた。一つ目は一枚の画像を操作し、照明が変化したと等価になるように輝度を変えたシェーディング画像、二つ目は様々な表情を含む画像である。三つ目は顔認識における照明の影響を評価するため、照明位置を変えて撮影した画像である。一つ目は表情の変化等の影響を避けるため、二つ目は一般的に顔認識に求められる状況での問題を評価するため、三つ目は実際のシェーディング画像(幾分表情変化を含む)での効果確認のためである。一つ目の輝度操作画像は以下の 4 種類である。

- 輝度に一律 0.6、1.2、1.6 を掛けたものの 4 種
- 左から照明が当てられたと等価になるよう、中央に対し左側 +40%、右側 40%としたもの及びそれぞれを +80% 80%としたものの 2 種
- b) の逆で、右側から照明が当てられたと等価な 2 種
- 部分的に照明が 40, 80%強くなったのを想定した 2 種

* 各係数の最大値が使われた場合、輝度レベルは部分的に飽和或いはゼロとなり、画像情報が失われる大きさである。我々の目的は照明の影響排除の方法であり、それ以上の係数は意味がないと判断し、係数を決定した。

これら a)~d) に元画像を追加し、10 枚を一組として 40 組分、全部で 400 枚の画像を使用して認識正解率の差を評価した。400 枚の内の 1 枚を未知画像とし、それがどの人物のグループに属するかを判定し、認識正解率を求めている。二つ目の画像は、各人物 10 種類の表情の違うものを 20 人分使用し、200 枚の画像より認識正解率を求めた。提案法の効果を確認するために比較対象としてシミュレーションを行ったのは、輝度を変化させた画像をそのまま使用する方法(以後未処理法と呼ぶ)、顔の中心部(図 2 の黒枠部)の平均値が輝度レベル 127 となるように階調に適切な係数を掛け算した平均値法、及び本提案の方法の三つである。

使用した画像は AT&T より提供されている ORL database 及び照明の影響評価用の Yale Face Database B の画像を 64×64 のサイズに顔を切り出し使用した。

4. シミュレーション結果

4.1. 特徴ベクトルの次元数

特徴ベクトルとしてフーリエ変換係数を用いているが、どの周波数領域の、幾つの次元とするかは判別分析の結果に大きく影響する。これは正解率が最も高くなるよう図 2 の画像 200 枚を用い、シミュレーションにより決定した。図 4 に特徴ベクトル次元数と不正解率を二ポイントの位置をパラメーターとして示す。この結果から gain2 の影響も考慮し、不正解率が最も低くなるよう、特徴ベクトル次元数を 5、二ポイントを 5 とした。判別分析の手順は、全体を二つのグループに分け、未知画像がどちらに属するかを判別する。その後、未知画像が分類されたグループを更に二つに分け、同様の計算を行っている。計算の負荷を無視すればそれぞれの判別段階で最適な二ポイント、gain2 を選択できるが、今回は二ポイント、gain2 の値は最終的に不正解率が最も低くなるものを選択し、特徴ベクトル次元数はそれぞれの判別段階で最適の値を選択した。図 4, 5 は最終段階での判別の結果を示している。特徴ベクトル次元数の最適値はサンプル数に依存し、一般的にはサンプル数が増えるに従い、最適次元数は増える傾向にあるといえる。

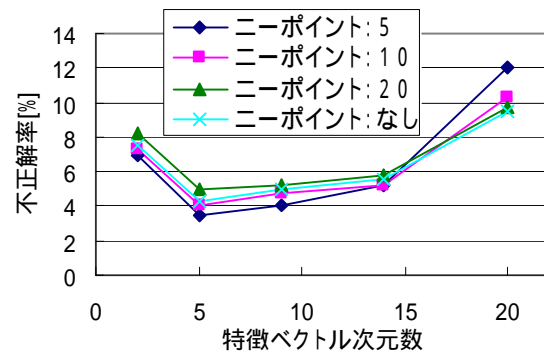


図 4 特徴ベクトル次元数と不正解率の関係。二ポイントは 5 で最も不正解率が低くなっている。因みに二ポイントなしとは非線形性がなくなることを意味する。

4.2. 非線形部の検討

図 1 における非線形の係数である Gain2 を決定するため、図 2 の画像を用い、シミュレーションを行った。特徴ベクトル次元数をパラメータとし、gain2 と不正解率の関係を結果を図 5 に示す。結果は gain2 : 0.4 が最も不正解率が低くなった。

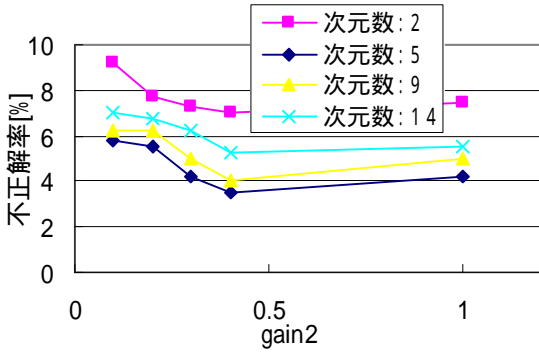


図 5 gain2 と不正解率との関係。gain2 : 1.0 とは非線形性がなくなることを意味する。

正解率が係数依存する理由として、係数が小さくなると細かな皺、髭などは消されてしまう。又信号が矩形波に近くなり、高周波分が発生すること。一方係数が大きくなると照明による影響を受けてしまうことが考えられ、最適値が存在すると推測できる。

4.3. 照明の影響排除の効果

表情の違いの影響を排除するため、図 2 に示した 200 枚の画像を用いて認識正解率を評価した。結果を表 1 に示す。単純な平均値法でも表情に差が無い画像であると、かなりの改善効果があることが分かる。結果として、照明の影響を何も補正しないと 94% の正解率となり、平均値を同じにするだけでも 97% に改善する。更に提案法により、照明の影響を削減し、98.6% まで改善できた。

表 1 認識正解率への照明の影響

	未処理	平均値法	提案法
正解率 [%]	94	97	98.6

4.4. 実撮影条件での提案法の効果

実際の照明条件の違いによる画像での本提案方式の効果を見るため、図 6 の 10 人分、100 枚と 10 種類の表情、顔の向きが異なる図 7 の画像、20 人分、200 枚を使つての認識正解率を評価した。結果を表 2 に示す。

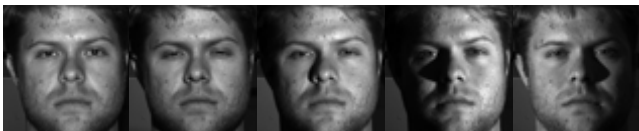


図 6 照明の位置を上下、左右に変化させた画像例



図 7 表情、顔の向きが異なる画像例

表 2 実撮影条件での正解率

	正解率 [%]	
	未処理	提案法
照明差の画像 (図 6)	95	97
表情差の画像 (図 7)	91	90.5

照明差の画像では提案法で若干の認識正解率の改善がみられるが、表情差の画像では優位差があるとは認められない。一方、表 1 に示した結果からは明らかに改善効果が認められる。これらを考え合わせると、表 2 の結果は、表情の違いがもたらすものと考えられる。図 6 の画像では照明位置により、鼻の横に強い影が発生しており、これは一種の表情変化ともいえる。実際に誤認識された画像は図 6 の右二つと同等の画像である。傍証として図 2 に見られる軽度の影は改善できるが、図 6 に見られる強い影は改善できていないのを確認している。又、図 7 の画像結果で両者に差が無いという事は提案法にデメリットが無い事を意味している。

5. まとめ

顔認識における前処理として、ローパスフィルターにより抽出した照明によるシェーディングを、ある係数を掛けて元画像から差し引き、顔の中心部の平均値に、ある係数を掛けたものを先の画像から差し引き、ある閾値を設けてそれを境にゲインを変えることで階調を圧縮し、照明の影響を軽減する方法を提案した。この画像をフーリエ変換し、得られたパワースペクトラム係数を線形判別する顔認識法において認識正解率が高くなることをシミュレーションにより実証した。結果は未処理の場合 94% であるものが、98.6% まで改善できた。

今回の提案法はシェーディングを抽出し、差分をとることで補正を行った。照明の光量と画像出力とは線形関係であり、補正は画像のゲイン調整で本来すべきである。その後非線形処理することで更に正解率向上の可能性が考えられ、その点の実証が今後の課題である。

参考文献

- [1] 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 (2001) 第 49 巻第一号 23-42
- [2] M.Turk and A.Pentland, “Eigenfaces for recognition” Journal of Cognitive Neuroscience Volume 3, No 1
- [3] Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J., “Eigenfaces versus Fisherfaces: recognition using class specific linear projection” IEEE Trans. Pattern Anal. March Intell. 19(7), 711-720
- [4] W.Zhao, R.Chellappa, P.J.Philips and Rosenfeld “Face Recognition: A literature Survey” ACM computing surveys, Vol.35 No.4, December 2003, pp.399-458
- [5] Akamatsu S., Sasaki T., Fukumachi H., and Suenaga Y. “A robust face identification scheme - KL expansion of an invariant feature space” Proc. SPIE, 1607, pp71-84, 1991-11
- [6] Kurita S., Tomikawa T. “A Study of The Evaluation of Face Parts in Facial Imaging” Proceedings of the IEEEJ IEVC2007