

濃淡・距離動画像を用いた顔認識システムの有効性の検証

Face recognition system based on the sequence of gray-scale and depth images

小山貴大 † 福井和広 ‡

Takahiro KOYAMA † Kazuhiro FUKUI ‡

概要

従来の顔認識システムにおいては、主に顔の濃淡動画像を用いて識別する手法が採られている。しかし、濃淡動画像を用いた認識は、顔向きや照明条件に大きく依存する。一方、距離動画像は照明条件に頑健である。そこで本研究は、顔認識精度を向上させるため、濃淡動画像に加え、距離動画像を用いた顔認識システムについて検証する。評価実験を行い、その結果、顔認識における距離動画像の有用性が示された。

1 まえがき

顔動画像を用いた顔検出、顔認識は個人同定、テレビ電話、セキュリティの分野など多くの場面で応用が期待できる。このような背景から、現在までに多くの顔検出手法や顔認識手法が提案されてきた [1]。

顔動画像のような動画像を入力として扱える識別器として、相互部分空間法 (MSM) [2]、制約相互部分空間法 (CMSM) [3]、直交相互部分空間法 (OMSM) [4] などが挙げられる。本研究では、さらにその拡張であるカーネル非線形直交相互部分空間法 (KOMSM) [5] を用いる。KOMSM は、画像パターンの変動を非線形部分空間で近似し、非線形な変動に対して高い識別性能を持つ識別器である。KOMSM の計算量を削減する手法 [6] を用い、濃淡動画像を用いたリアルタイム顔認識システムも提案され、高い性能を有することが示されている [7]。

従来手法に多く用いられている顔の濃淡画像は、3次元形状と顔面上のテクスチャーを反映しており、顔認識において有効である。しかし、濃淡画像は照明条件による変動が大きく、照明変動の大きな環境において顔認識に用いる際に性能低下を招く原因になる。

そこで本研究では、濃淡動画像に加え、距離動画像を用いた顔認識システムについて検証する。距離動画像は、距離センサから対象までの距離値を動画像として取得したものである。つまり、距離動画像は対象の3次元形状特徴を直接表しているため、照明条件に対し頑健で

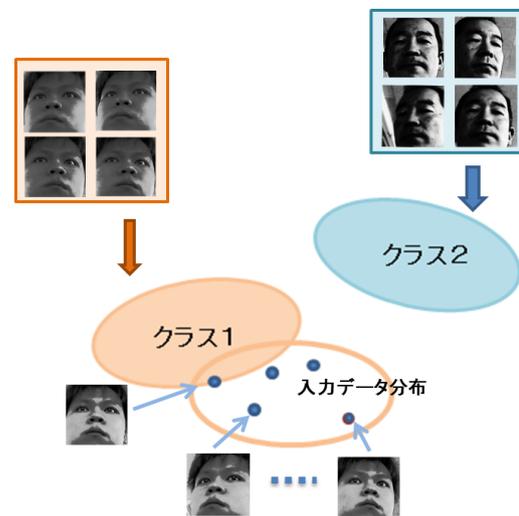


図1: データ分布同士の比較識別

ある。この点から、濃淡・距離動画像を顔認識に用いることで、識別性能の向上が期待できる。

本研究の検証では、連続的に取得した濃淡と距離動画像を用いて KOMSM により顔認識を行う。以下では本研究のベースとする濃淡動画像を用いる場合について説明する。まず、比較すべき2つの濃淡動画像に含まれる動画像セットの分布をそれぞれ非線形部分空間で近似する [8]。そして、得られた2つの非線形部分空間の成す角度を類似度として顔認識を行う。これは図1のように各非線形部分空間に含まれるデータ分布同士を比較する識別に相当する。距離動画像についても、同様の識別を行う。さらに、濃淡・距離の両動画像から得られた2つの類似度を同時に考慮する事で、濃淡と距離動画像を用いた顔認識の有効性を検証する。

† 筑波大学情報学群情報科学類

‡ 筑波大学大学院システム情報工学研究科

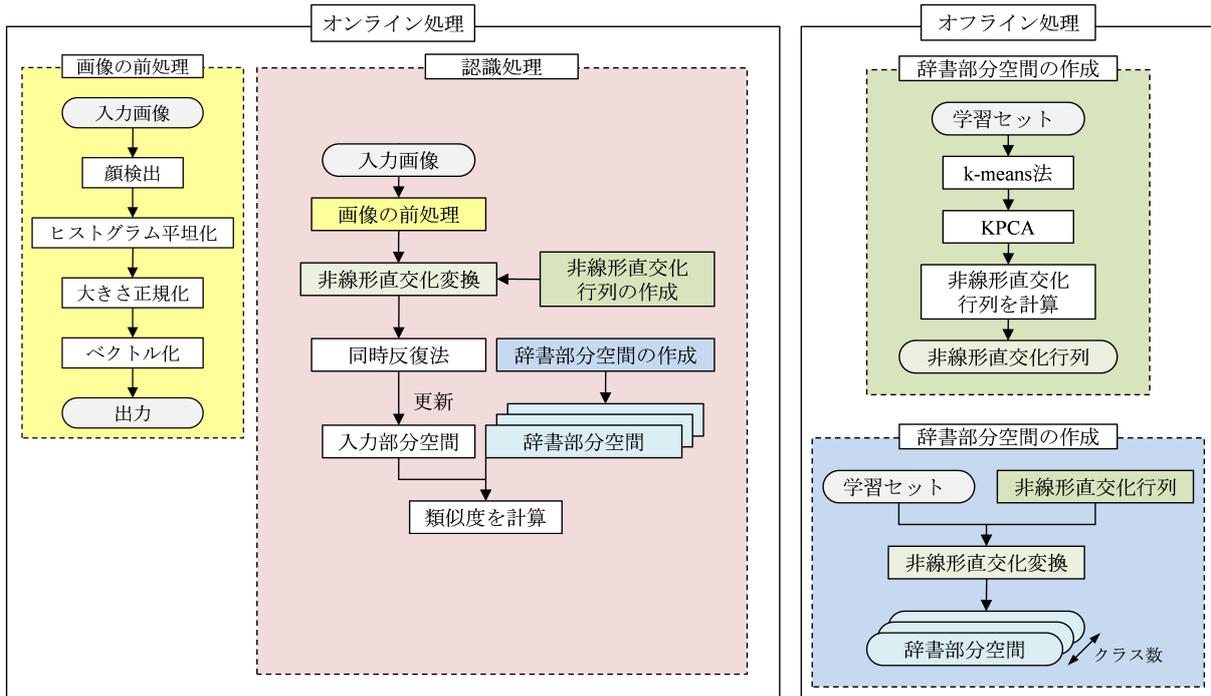


図2: 顔認識システムの処理の流れ

2 ベースとなる KOMSM を用いた顔認識システムの概要

図2に KOMSM を用いた顔認識の処理の大まかな流れを示す。処理は大別すると、画像の前処理、非線形直交変換行列の作成、辞書部分空間の作成、認識に分けられる。非線形直交変換行列の作成と辞書部分空間の作成は、オフラインで行われる。

2.1 入力画像からの特徴ベクトルの抽出

入力画像に対して、次の前処理を行う。

- (1) OpenCV[9] を用いて顔検出を行う。
- (2) 切り出した顔画像に対しヒストグラムの平坦化を行う。
- (3) 大きさ (30 × 30 ピクセル) に正規化する。
- (4) 上記の顔画像をベクトル化して 900 次元の特徴ベクトルを得る。

2.2 識別処理

入力画像に対して前処理を行い、非線形直交変換行列を用いて変換し、入力部分空間を得る。識別においては、

入力部分空間と辞書部分空間との間の正準角を計算し、各クラスの類似度を得る。

m 次元部分空間 P と n 次元部分空間 Q が与えられたとき、部分空間 P と部分空間 Q の間には m 個の正準角 ($0 \leq \theta_1, \dots, \theta_m \leq \frac{\pi}{2}$) を定義することができる。第1正準角 θ_1 は、 P 、 Q 間の最小角となる。第2正準角は θ_2 は θ_1 に直交する方向で測った最小角となる。第3正準角以降も、同様にして θ_i ($i = 1, \dots, m$) が決まる。

部分空間 P の第 i 正規直交基底を p_i 、部分空間 Q の第 j 正規直交基底を q_j とする。これらを列方向に並べた行列を $U = [p_1, p_2, \dots]$ 、 $V = [q_1, q_2, \dots]$ とする。 $U^T V$ における第 i 番目に大きい特異点が第 i 正準角に対応する $\cos^2 \theta_i$ となる。2つの部分空間の類似度は、第 n 正準角までを用いて、以下の式で定義される。

$$Sim = \sum_{i=1}^n \frac{\cos^2 \theta_i}{n} \quad (1)$$

2.3 辞書部分空間および直交変換行列の生成

学習データセットに対して、k-means 法 [6] を適用してデータ数を削減したうえで、KPCA[10] により各クラスの非線形辞書部分空間を生成する。次に、全クラスの非線形辞書部分空間からこれらを直交化するための直交変換行列を求める。各クラス部分空間はこの直交変換行

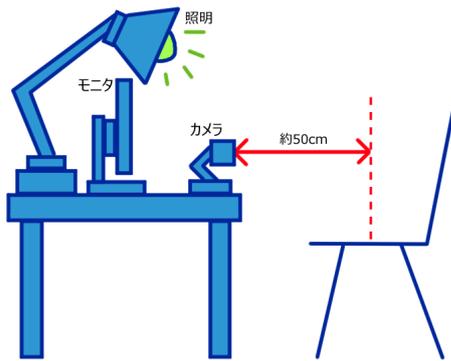


図 3: 実験環境

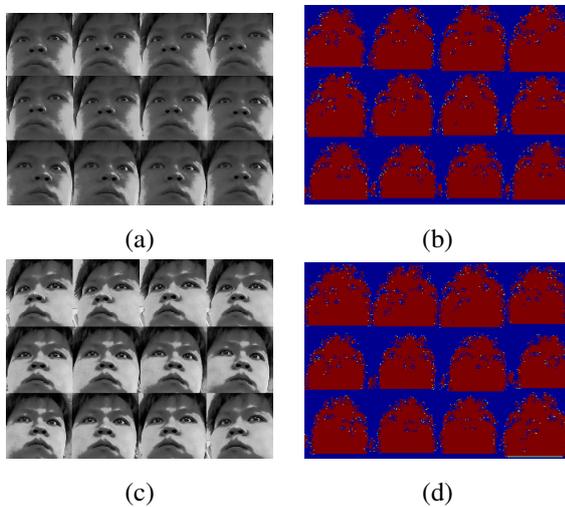


図 4: 顔画像連続 12 フレーム :

(a) 照明条件 1 の濃淡画像, (b) 照明条件 1 の距離画像,
(c) 照明を当てた照明条件 2 の濃淡画像, (d) 照明条件 2
の距離画像.

列を用いてお互いに直交化しておくことで、クラス部分空間の分離を高め、クラス識別を大幅に高める。

kOMSM では、カーネル関数により入力ベクトルを特徴ベクトル空間へ非線形写像したうえで、線形の OMSM を実行する。特徴ベクトル空間は超高次元、場合によっては無限次元となるので、通常は内積計算が困難となる。そこでカーネルトリックと呼ばれる計算技法により、特徴ベクトル空間における内積を可能としている。カーネル関数は次式を示されるガウシアンカーネル ($\sigma = *$) を用いる。 σ は最適な値を用いた。

$$k(x, y) = \exp\left(-\frac{\|x - y\|}{2\sigma^2}\right) \quad (2)$$

KOMSM の詳しい処理は文献 [6] を参照されたい。

3 濃淡 + 距離動画像の有効性の検証

濃淡動画像と距離動画像の組み合わせの有効性について実験を行った。本実験では、以下の条件で式 (1) を用いて識別結果を求める。

- ・条件 1 :
距離動画像から得られる正準角を単独で用いる。
- ・条件 2 :
濃淡動画像から得られる正準角を単独で用いる。
- ・条件 3 :
濃淡と距離動画像から得られる正準角の平均を用いる。

3.1 評価動画像列データの収集

4 人の被験者から一人あたり水平方向に顔向きを変化させた濃淡動画像 200 フレーム、垂直方向に顔向きを変化させた濃淡動画像 200 フレーム撮影した。実験環境を図 3 に示す。濃淡動画像 200 フレームのうち、100 フレーム (辞書データ) を撮影し、残りの 100 フレーム (入力データ) を照明条件を変化させ撮影した。同様の枚数を、距離動画像で撮影した。撮影した画像の一例を図 4 に示す。図 4 における (a) と (b), (c) と (d) はそれぞれ全フレーム同時時間に撮影され対応している。入力データ 100 フレームを連続した 5 フレームを 1 セットとして分割した。

3.2 実験および考察

識別を行う際、辞書部分空間、入力部分空間の次元は予備実験により求めた最適な値を用いた。

実験により得られた各実験条件における ER, EER の値を表を示す。図より、本実験では条件 3 における ER, EER が良いことが確認できる。条件 2 の識別結果と比べ、ER では 2.81% から 1.02%, EER では 2.04% から 1.79% へ改善されたことが確認できる。

また、図 5 に、KOMSM による各条件の識別結果に対する ROC 曲線を示す。ROC 曲線はクラス間の重なりが少ないほど、グラフが左上に近付いていく曲線である。濃淡と距離動画像から得られる正準角の平均を用いることで、識別性能が向上するということが確認できた。

上記の結果は、距離動画像の持つ照明条件に影響を受けないという特性が、濃淡動画像の欠点を補っていると考えられる。

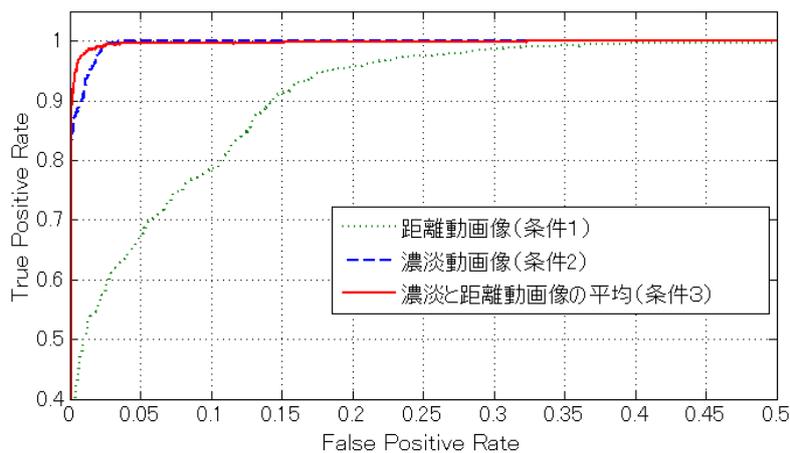


図 5: ROC 曲線

表 1: ER-EER

	ER(%)	EER (%)
距離動画像 (条件 1)	22.83	13.5
濃淡動画像 (条件 2)	2.81	2.04
濃淡と距離動画像の 平均 (条件 3)	1.02	1.79

4 おわりに

本研究では、濃淡・距離動画像を組み合わせた顔認識システムを検証し、有効性を確認した。濃淡動画像におけるリアルタイム認識は既に実現できているので、距離動画像を共に用いたリアルタイム認識を実装していくことを今後の課題とする。また、距離動画像の識別性能の向上がシステム全体の性能を決定するので、距離動画像単体での顔認識の識別性能の向上も今後の課題とする。

謝辞

本研究は科研費 (24300290) の助成を受けたものである。

参考文献

- [1] 岩井儀雄, 勞世, 山口修, 平山高嗣, “画像処理による顔検出と顔認識”, 情報処理学会研究報告, CVIM-149, pp.343-368, 2005.
- [2] 山口修, 福井和広, 前田賢一, “動画像を用いた顔認識システム”, 電子情報通信学会技術研究報告,

PRMU, Vol. 97, No. 113, pp. 17-24, 1997.

- [3] 福井和広, 山口修, 鈴木薫, 前田賢一, “制約相互部分空間法を用いた環境変動にロバストな顔画像認識: 照明変動の影響を抑える制約相互部分空間の学習”, 電子情報通信学会論文誌 D-II, Vol. J82-D-II, No. 4, pp. 613-620, 1999.
- [4] 河原智一, 西山正志, 山口修, “直交相互部分空間法を用いた顔認識”, 情報処理学会研究報告, CVIM-151, pp. 17-24, 2005.
- [5] Kazuhiro Fukui, Osamu Yamaguchi, “The Kernel Orthogonal Mutual Subspace Method and its Application to 3D Object Recognition”, ACCV2007, Part-II, pp.467-476, 2007.
- [6] Yasuhiro Ohkawa, Kazuhiro Fukui, “Hand Shape Recognition based on Kernel Orthogonal Mutual Subspace Method”, IAPR MVA2009, pp.222-225, 2009, JAPAN.
- [7] 五十嵐洋介, 大川泰弘, 北川景介, “高精度リアルタイム顔認識システムの構築”, ICT2009, 2009.
- [8] Martin Peris, Kazuhiro Fukui, “Both-hand Gesture Recognition Based on KOMSM with Volume Subspaces for Robot Teleoperation”, IEEE-Cyber2012, pp.191-196, 2012.
- [9] opencv.jp, <http://opencv.jp/>, 2014/6/27.
- [10] 坂野鋭, “パターン認識における主成分分析-顔画像認識として-”, 統計数理, 第 49 巻, 第 1 号, pp.23-42, 2001.