

擬似マハラノビス相互距離法を用いた動画像顔照合 Face Recognition using Pseudo Mahalanobis Mutual Distance

井上 晃†
Akira Inoue

佐藤 敦†
Atsushi Sato

1. まえがき

動画像を用いた顔照合は、Query 側データとして静止画ではなく、連続した入力画像シーケンスを利用出来るのが特徴である。山口[1]らは、入力シーケンスの部分空間と、登録画像が張る部分空間との角度を利用した相互部分空間法[2]を動画像顔照合に適用した。動画像顔照合において、このように入力画像シーケンスを有効に利用する例は少なく、実際には各フレームに対して静止画照合を連続して行うことが多い。

動画像照合は、入力画像シーケンスによって形成される入力特徴分布と、登録された画像群の形成する辞書特徴分布との分布間距離を評価することであると考えられる。このようなクラス間距離としては、Kullback 情報量、ダイバージェンス、パタチャリア距離などが提案されている[5]。しかしながら、サンプルデータが少ない場合、計算に必要な特徴分布関数の正確な推定は一般に困難であり、有効な距離値を計算することはできなかった。

そこで著者は、少ないデータからでも実用的な分布間距離を算出する方法として、擬似マハラノビス距離を用いた相互距離法を提案する。本手法を用いて実際の動画像による顔照合実験を行った結果、静止画ベースの照合に比べて大幅な性能向上が見られた。

2. 擬似マハラノビス相互距離法

2.1 相互距離

2つの分布間の距離値 D を以下のように定義する。

特徴ベクトル x と、クラス C_1, C_2 との距離が、あらかじめ $d_1(x), d_2(x)$ として定義されているものとする。双方の分布の中心ベクトル (平均ベクトル) を m_1, m_2 とすると、 m_2 からクラス C_1 への距離は $d_1(m_2)$ 、 m_1 からクラス C_2 への距離は $d_2(m_1)$ となる。図 1 に $d_1(x)$ と $d_2(x)$ の分布の一例を示す。求める分布間距離 D は、 m_1 と m_2 を結ぶ直線上の関数 $d_1(x) + d_2(x)$ 上に存在すると考える。

$d_1(x), d_2(x)$ は $x = m_1, m_2$ のとき最小値 0 をとり、そこ

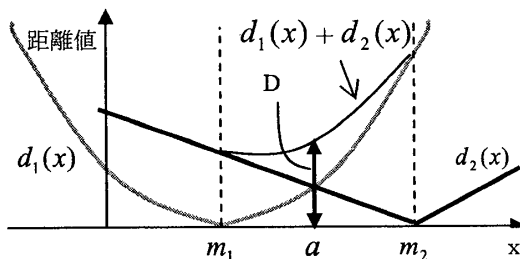


図 1 距離値の分布と相互距離 D

から離れるにつれて単調増加する関数とする。このとき m_1, m_2 を結ぶ直線上の $[m_1, m_2]$ の範囲で、 $d_1(x)$ と $d_2(x)$ が等しくなる点 a が一つ存在する。この点は 2 つの分布から等距離にある点と考えられる。そこで、点 a における、 $d_1(x)$ と $d_2(x)$ の和として、相互距離 D を定義する。

$$D = d_1(a) + d_2(a) \quad (1)$$

2.2 特徴ベクトルと分布との距離

次に距離 $d_1(x), d_2(x)$ を定義するため、マハラノビス距離を基にした測度を定義する。マハラノビス距離は、正規分布を仮定した分布と特徴ベクトルとの間の標準化距離であり、広範囲なパターン認識分野に利用されている。次元数 n の特徴ベクトルを x 、分布の平均ベクトルを m 、分布の分散共分散行列を Σ とすると、マハラノビス距離は式(2)によって定義される。

$$d_m^2 = (x - m)^T \Sigma^{-1} (x - m) \quad (2)$$

しかしながら、認識対象の画像サンプル数が小さい場合には、サンプルから計算した分散共分散行列が特異になり、算出される距離値が不安定になる問題がある。そこで従来から、この不安定要因を回避する距離算出法が提案され認識に適用されている[3][4]。本手法においては文献[3]を参照し、次のように分散共分散行列を推定する。

$$\Sigma = \hat{\Sigma} + \sigma^2 I \quad (3)$$

$\hat{\Sigma}$ は標本サンプルから算出される分散共分散行列で、第二項は分散共分散行列の初期推定値を表す (σ^2 は定数)。分布の標本サンプルから求められる固有値を λ_i 、固有ベクトルを Φ_i とするとマハラノビス距離は次式で得られる。

$$d_m^2 = \sum_{i=1}^n \frac{1}{\lambda_i + \sigma^2} \{\Phi_i^T (x - m)\}^2 \quad (4)$$

固有値の小さい成分はノイズが支配的となる。そこで次元 $i > k$ のとき $\lambda_i \ll \sigma^2$ と仮定して変形し式(5)を得る。本手法ではこの平方根を、擬似マハラノビス距離として定義し、分布との距離 $d_1(x), d_2(x)$ に用いる。

$$d_p^2(x) = \|x - m\|^2 - \sum_{i=1}^k \frac{\lambda_i}{\lambda_i + \sigma^2} \{\Phi_i^T (x - m)\}^2 \quad (5)$$

式(5)より $d_p^2(x)$ は、 $x - m$ の長さの 2 乗に比例するため、擬似マハラノビス距離は、 $x - m$ の長さに比例することがわかる。図 2 に擬似マハラノビス距離を用いたときの距離分布を示す。 $d_1(x), d_2(x)$ は、 m_1, m_2 を結ぶ直線上において次の直線で表される。(但し $L = |m_2 - m_1|$)

$$d_1(x) = \frac{d_1(m_2)}{L} x, \quad d_2(x) = \frac{d_2(m_1)}{L} (L - x) \quad (6)$$

$d_1(x) = d_2(x)$ を解くことで、等距離点 a が得られる。

† NEC マルチメディア研究所

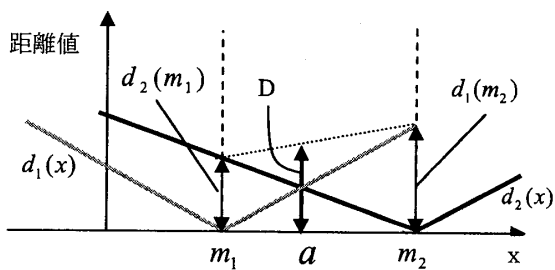


図2 擬似マハラノビス距離による相互距離D

$$a = \frac{d_2(m_1)}{d_1(m_2) + d_2(m_1)} L \quad (7)$$

式(1)より相互距離Dは次のように得られる。

$$D = 2 \cdot \frac{d_1(m_2) \cdot d_2(m_1)}{d_1(m_2) + d_2(m_1)} \quad (8)$$

なお、ダイバージェンス[5]は、分布に正規分布を仮定し、定数項を無視すると、マハラノビス距離の相加平均となる。そこで本手法と同様、擬似マハラノビス距離を用いた、擬似ダイバージェンス(式(9))を定義し比較を行った。

$$D_D = \frac{1}{2} (d_1(m_2) + d_2(m_1)) \quad (9)$$

3. 照合実験

3.1 テスト画像シーケンス

本手法は主に動画を対象とした照合方法であり、登録辞書、及びテスト画像が複数存在することを前提としている為、動画シーケンスを用いた顔照合実験を行った。実験に用いた画像は、DVカメラを用いて室内の異なる12方向の光源下で撮影された動画シーケンスから、市販のキャプチャカードを使ってフレームごとに取得した正面顔画像群であり、顔画像は輝度値の正規化後、手動で与えた両目位置を基準として12x18画素の大きさに正規化する。また登録辞書は12方向の光源シーンを含んでいる。表1に実験に用いたデータセットを、図3に正規化後の画像例を示す。



図3 実験画像例

表1 実験に用いたデータセット

撮影環境	登録枚数	テスト枚数
室内	200枚/人	4420枚
光源12方向	8人分登録	16人

3.2 照合実験の流れ

照合実験は次のように行った。

- 1) あらかじめ、人物毎に登録画像を用いて平均値、固有値、固有ベクトルを求めておく(辞書統計量)。
- 2) テスト用の画像シーケンスから、N枚の連続画像(フレーム番号f=1からNまで)を取得する。
- 3) 入力N枚のデータから、平均値、固有値、固有ベクトルを求める(入力統計量)。
- 4) 人物毎に、辞書統計量と入力統計量を用いて、擬似マハラノビス相互距離Dを計算し、最も距離値に近い人物を照合結果として出力する(またはリジェクトする)。
- 5) 次のN枚(フレームf=2からN+1まで)のシーケンスを用い、3) 4) の処理を繰り返す。以降フレーム番号を増

加しつつテストシーケンスが終了するまで繰り返す。

3.3 実験結果

図4に顔照合実験結果をROC曲線によって示す。投影次元数kは入力16,登録16とし、入力フレーム数はN=32とした。また本手法の有効性を検証するため、同じデータセットを用いて部分空間法(CLAFIC、投影距離法)、相互部分空間法、擬似ダイバージェンスとの性能を比較した。なお部分空間法では、入力Nフレームのスコア平均を用いて照合を行った。表2に平均処理時間を示す。

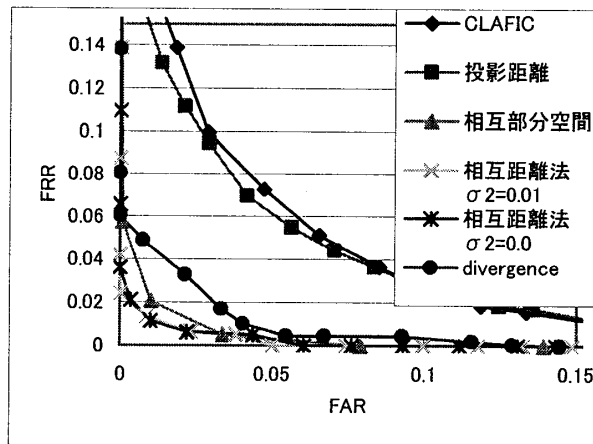


図4 照合実験結果

表2 処理時間比較 (Queryあたり、N=32)

手法	平均処理時間 msec
部分空間法 N回	32
相互部分空間法	23
擬似マハラノビス相互距離法	12

4. おわりに

本稿では、動画画像の照合に有効な複数入力画像対複数登録画像の照合方式、擬似マハラノビス相互距離法を提案した。本手法は分布の平均を結ぶベクトルをそれぞれの分布に投影して2つの擬似マハラノビス距離を算出し、それらを統合して算出する。擬似マハラノビス距離は、分布の平均と固有値の大きなk個の固有ベクトルのみを用いて算出するため、処理コストが低く、サンプル数が少ない場合でも安定的に距離を推定することができる。

正面顔画像シーケンスを用いた照合実験を行った結果、部分空間法に比べ大幅に性能が改善され、擬似ダイバージェンスを上回り相互部分空間法に匹敵する照合性能を持つことを示した。また計算時間に関しては相互部分空間法の約半分の処理速度で実行可能であり、実用性も極めて高いことを示した。

参考文献

- [1] 山口, 福井, “動画画像を用いた顔認識システム”, PRMU97-50, 1997
- [2] 前田, 渡辺, “局所構造を導入したパターンマッチング法”, 信学論(D), Vol.J68-D, No.3, pp.345-352(1985)
- [3] F.Kimura et.al, “Handwritten Numerical Recognition based on Multiple Algorithms”, Pattern Recognition, vol.24, no.10, 1991
- [4] B.Moghaddam, A.Pentland, “Probabilistic Visual Learning for Object Detection”, Proceedings of the ICCV, 1995
- [5] 高木, 下田, “画像解析ハンドブック”, 東大出版会, pp.656-658, 1991