

カメラ付きヘッドマウントディスプレイを用いたウェアラブル顔認識システム

加藤 祐也[†] 内海ゆづ子[†] 黄瀬 浩一[†]

[†]大阪府立大学大学院工学研究科 電気・情報系専攻 知能情報工学分野 〒599-8531 大阪府堺市中央区学園町1-1
E-mail: †katou@m.cs.osakafu-u.ac.jp, ††{yuzuko,kise}@cs.osakafu-u.ac.jp

あらまし 本稿では、カメラ付きヘッドマウントディスプレイを用いて、出会った人を個人識別し、識別した人物の認識結果を表示するウェアラブルな記憶補助システムを提案する。提案システムでは、人物認識の手法として顔認識を用いる。顔認識の手順は、まずカメラにより撮影された画像から顔を検出する。その後、検出した顔領域を追跡することで、顔認識の際に顔向きや表情に頑健な認識ができるよう複数枚の顔画像を得る。そして得られた顔画像セットを用いて、Affine hullを用いた顔向きの変化に頑健な顔画像集合に対する認識手法により、人物を認識し、人物の情報をヘッドマウントディスプレイに表示する。提案システムでは、顔検出、追跡、認識をリアルタイムで実現し、認識結果の表示に成功した。

キーワード 顔認識, 顔検出, 顔追跡, ヘッドマウントディスプレイ, ウェアラブル

1. はじめに

近年、人の生活・行為・体験などの思い出をデジタルデータとして記録するライフログが注目を集めており、人の記憶を記録し、表示する研究が行われている [1]。人の記憶には限界があり、日常の全てのことを覚えておくのは困難である。誰しも、一度会ったことのある人に、どこかでもう一度会った時に声を掛けられたが、名前が思い出せず困惑した経験があるだろう。人は日々新しい出会いを経験しており、頻繁に会わない人の名前は忘れてしまう。そこで、一度会ったことのある人に再会した際に、自動で名前を表示するサービスがあれば有用である。そこで、本稿では、出会った人を記録し、記録した人に再会した場合、リアルタイムで名前を表示する記憶補助システムを提案する。

記憶補助システムの結果を表示する装置として、図1のようなヘッドマウントディスプレイ (HMD) が考えられる。HMDは、メガネ感覚で自然に装着可能であり、ハンズフリーで使える利点がある。また、HMDにカメラを取り付けることで、相手と対面した状態から新たなデバイスを使用せずに人物認識することが可能である。HMDに認識結果を表示させることで、会話などのコミュニケーションを中断することなく、自然に情報を確認できる。そこで本稿では、システムの結果を表示する装置として、機器の操作をせず自然に結果を見ることができると提案する。

記憶補助システムを作成する場合、各個人を認識し、人物を見分ける必要がある。人物を認識する方法として、音声、服装、顔などを認識する手法が考えられるが、本稿では、短期的に変化しない同一特徴量が得られ、また機器への接触を必要とせず、離れていても相手の顔を見るだけで認識可能である顔認識を用いる。

提案システムで顔認識をする場合、会話中に認識対象



図1 HMD

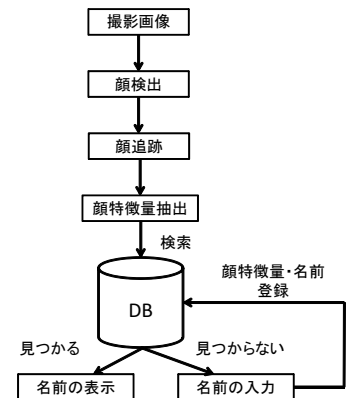


図2 処理の流れ

の人物が横を向いたり表情を変えたりすることがある。そのため、本システムに用いる顔認識手法は、認識対象人物の顔向きや表情にロバストなものが望ましい。顔向きや表情にロバストな顔認識を行う手法として、1人に対して複数枚の画像を用いて認識するものがあげられる。そこで、HMDに取り付けられたカメラで撮影された動画の中から顔を検出し、顔を追跡することで、同一人物の複数枚の画像を獲得し、得られた顔画像を複数枚に対応した顔認識手法に適用した。また、人に出会った場合には、瞬時に顔認識がなされ、名前が表示されることが求められる。そこで、顔検出、顔追跡に高速な手法を用いることで、リアルタイムに顔認識を実現した。

2. 提案システムの概要

提案システムの処理の流れを図2に示す。まず、出会った人をHMDに取り付けたカメラで撮影する。撮影された画像より顔領域を検出する。そして、顔追跡と顔特徴量の抽出を行い、人物を登録したデータベースから検索する。検索した結果、人物が見つければ名前を表示する。見つからなければ、初めて会った人とし、名前情報と共に

に顔特徴量をデータベースに登録する。

3. 顔検出・顔追跡手法

本稿では、顔検出手法として、高速な顔検出手法である Haar-Like 特徴量を用いた手法 [2] を使用した。Haar-Like 特徴量を用いた手法では、まず顔画像と顔の含まれていない画像を何枚かずつ予め用意して、Haar-Like 特徴量を抽出する。学習アルゴリズムとして AdaBoost を用いて作成した強識別器を、Attentional Cascade と呼ばれる手法により連結する。Attentional Cascade で連結することにより、顔でない画像を早い段階で除去することが可能であり、非常に高速な顔検出を実現している。

顔が検出された場合、顔追跡を行うことで、複数の顔画像を取得する。本稿では、会話中に認識対象人物が顔向きを変えることに対応するため、顔向きの変化にロバストな Mean-shift を用いた顔追跡法 [3] により、顔を追跡する。Mean-shift を用いた顔追跡法は、追跡対象のカラーヒストグラムから、ヒストグラム間の類似度関数に従い求められる各ピクセルの重み関数の最適化を行う。領域内のヒストグラムで比較するため、追跡対象の形状変化や部分的な隠蔽に対して頑健である。また、Mean-shift を用いた手法は、局所的な重みを計算するため、高速でリアルタイム処理が可能である。

4. 顔認識手法

顔検出、追跡で得られた複数の顔画像に対して顔認識を実行する。認識手法には、Affine Hull を用いた顔認識手法 [4] を使用した。この手法は、認識の際にクエリ・データベース共に、1枚の画像を用いるのではなく、複数の画像を用いて認識を行うため、顔向きや表情の変化に対して頑健に認識可能である。Affine Hull を用いた手法では、Affine Hull と呼ばれる凸包を作成し、Affine Hull 同士の距離を使用して、認識を行う。Affine Hull 同士の距離を求める手法の概要を説明する。顔画像サンプルの集合を $\mathbf{x}_{ci} \in \mathbb{R}^d$ とする。ここで、 C 人の顔画像がある場合、各個人の画像セットのインデックスを $c = 1, \dots, C$ とし、各画像セット c のそれぞれの画像のインデックスを $i = 1, \dots, n_c$ とする。これらの画像セットを含む最小の Affine 部分空間は、式 (1) となる。

$$H_c^{\text{aff}} = \left\{ \mathbf{x} = \sum_{k=1}^{n_c} \alpha_{ck} \mathbf{x}_{ck} \mid \sum_{k=1}^{n_c} \alpha_{ck} = 1 \right\} \quad (1)$$

式 (1) で表される凸包を Affine Hull とする。

この Affine Hull をパラメータで表現するため、適当な基準点を取り μ_c とする。 μ_c を用いて、式 (1) を書き換えると、式 (2) のように表すことができる。

$$H_c^{\text{aff}} = \{ \mathbf{x} = \mu_c + \mathbf{U}_c \mathbf{v}_c \mid \mathbf{v}_c \in \mathbb{R}^l \} \quad (2)$$

ここで、 \mathbf{U}_c は Affine 部分空間の正規直交基底、 \mathbf{v}_c は部分空間の次元を削減した時のパラメータベクトルである。



図 3 システムの表示結果

\mathbf{U}_c は $[\mathbf{x}_{c1} - \mu_c, \dots, \mathbf{x}_{cn_c} - \mu_c]$ に特異値分解 (SVD) を適用することで得られる。

2つの交差しない Affine Hull $\{\mathbf{U}_i \mathbf{v}_i + \mu_i\}$, $\{\mathbf{U}_j \mathbf{v}_j + \mu_j\}$ が与えられた時、2つの Affine Hull の最も近い点同士の距離は式 (3) から求められる。

$$\min_{\mathbf{v}_i, \mathbf{v}_j} \|(\mathbf{U}_i \mathbf{v}_i + \mu_i) - (\mathbf{U}_j \mathbf{v}_j + \mu_j)\|^2 \quad (3)$$

式 (3) より、2つの Affine Hull 同士の距離を求められる。クエリと登録されている全てのデータベースとの Affine Hull 同士の距離を求め、距離が最も近い人物を認識結果とする。図 3 は、提案システムにより、顔検出・顔追跡・顔認識を行った結果である。顔が横を向いても、認識可能であることがわかる。

5. おわりに

本稿では、カメラ付き HMD を用いて、リアルタイムに名前を表示させるウェアラブルな顔認識システムを提案した。Cascade 型顔検出器と Mean-shift を用いた顔検出手法を組み合わせることで、高速にかつ顔向きに頑健な顔検出、顔追跡を行った。また、認識に 1枚の画像だけを用いるのではなく、クエリ・データベースともに、複数枚の画像を使用する Affine Hull を用いた手法を使用することで、顔向きに頑健な顔認識システムを実現した。今後の課題として、認識精度の向上やさらなる高速化があげられる。また、文字認識と組み合わせ、名刺や名札などから自動的に名前を取得して登録するシステムを実装することも今後の課題である。

文 献

- [1] M. Lamming and M. Flynn, “Forget-me-not: Intimate computing in support of human memory”, Proc. FRIEND21, pp. 125–128, February, 1994.
- [2] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features”, Proc. Computer Society Conference on Computer Vision and Pattern Recognition, no. 1, pp. 511–518, December, 2001.
- [3] D. Comaniciu, V. Ramesh and P. Meer, “Real-time tracking of non-rigid objects using mean shift”, Proc. Computer Society Conference on Computer Vision and Pattern Recognition, no. 2, pp. 142–149, December, 2000.
- [4] H. Cevikalp and B. Triggs. “Face recognition based on image sets”, Proc. Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2567–2573, June, 2010.