

全方位画像列からの個人識別

小原 ゆ う[†] 八木 康 史[†]
横 山 太 郎[†] 谷内田 正彦[†]

全方位視覚センサは、広視野を同時に観測できるということからサーベイランスなどの目的に利用されつつある。しかし、入力画像の解像度は、撮像素子の素子数と密接に関係していることから、サーベイランスにおいても、人物の発見、追跡ならびに大きなジェスチャ認識といったおおまかな画像理解に関する研究にとどまっていた。本論文では、サーベイランスとしての機能をより充実する目的で全方位視覚センサにより撮像された顔画像からの個人識別という問題に取り組む。提案手法では、シーン中を歩行する人物画像から歩行者の顔画像列を抽出し、低解像度ながら長時間観測できる特徴を利用し、顔画像列から個人を識別する。

Face Recognition from an Omnidirectional Image Sequence

YUU OHARA,[†] YASUSHI YAGI,[†] TARO YOKOYAMA[†]
and MASAHICO YACHIDA[†]

Face is one of the most attractive information for personal identification. In this paper, we propose the personal identification method from an omnidirectional image sequence. Since an omnidirectional image sensor HyperOmni Vision observes a 360-degree view around the robot, it can observe a global azimuth information of the person (face). We track the human face while the person walks around the camera. Under an assumption of smooth human motion, we identify the corresponding person from facial database.

1. はじめに

人の行動を追跡、理解するうえで重要なことは環境を含めて人の情報をできるだけ多く取得することである。そのため一度に広範囲の画像を得ることができる全方位視覚センサの人物観測への利用が注目され、サーベイランス、ビデオインタラクションのための人物追跡と位置の検出やジェスチャ認識、頭部姿勢推定などの研究が報告されてきた^{1)~10)}。しかし、一般に全方位視覚センサは、通常のカメラ1台により撮像しているため、空間分解能が低く、詳細な人物の情報計測には適さない。しかし、人を観測するうえで、大まかな身体動作だけではなく、その人が誰であるかを知ることが重要である。特にセキュリティシステムへの応用を視野に入れた場合、この点は非常に重要である。そこで、本論文では、個人を識別する手段として人物の顔に着目し、人物を観測しながら同時に個人を識別できる手法の実現を目指す¹¹⁾。

ところで顔画像による個人識別は利用者の負担が少ないということから、近年、さかんに研究されている¹²⁾。代表的な顔研究としては、Turkらによる Eigen-Face の研究¹³⁾、Gabol Waveret を用いた方法^{14),15)} などがある。これらの研究の多くは、正面顔画像を対象としていた。また、姿勢や照明に対する自由度を増すために、複数の姿勢と照明条件で撮影された複数枚の画像を教示画像として用いる手法も数多く提案されている。これらの手法の多くは、1枚の入力画像とデータベース間での照合によりロバストな個人認識が実現されている。しかし、一般にこれらの手法は、通常のカメラで撮影された画像を対象にしているため、全方位画像のような空間分解能の低い画像に対しては、単一画像による認識手法では安定な認識が難しい。

連続顔画像列を対象とした認識手法としては、平均の認識率を計算する方法や忘却係数を設けて最新の認識結果を有効に利用する方法¹⁶⁾などが利用されるが、歩行時の画像の場合、現在の情報が必ずしも有用とは限らず、観測された奥行きや顔の向きを考慮した評価の仕方が重要となる。福井らによる相互部分空間法¹⁷⁾やさらにその発展である制約相互部分空間法を用いた

[†] 大阪大学大学院基礎工学研究科システム人間系専攻
Department of Systems and Human Science, Graduate
School of Engineering Science, Osaka University



図 1 全方位視覚センサ
Fig.1 HyperOmni Vision.

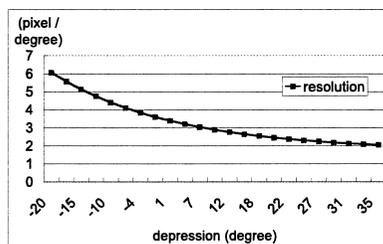


図 2 HyperOmni Vision の解像度特性
Fig.2 The angular resolution of HyperOmni Vision related to the dipression.

手法¹⁸⁾は、複数の視点画像を利用することで、単一の画像では不安定な認識を安定に行っている。この手法の場合、顔器官を特徴として、顔領域を正確に切り出し、大きさと向き正規化を行ってから、さまざまな方向や照明条件の正規化顔領域をセットにして部分空間を構成している。しかし、全方位画像の場合、画像中での顔画像の領域が小さいため、正確な顔器官の抽出が困難なため、あらかじめ正確な位置、姿勢、大きさの正規化を前提とした手法は適用しにくい。

また全方位画像を利用した個人識別手法が Huangらにより提案されているが¹⁹⁾、歩行時も含めてセンサに最も人物が近付いたときの単一画像を認識に使用している。しかし歩行時においては人物とセンサの距離が近いことが必ずしも個人識別に有用とは限らない。

それに対し、本手法では、差分処理により抽出された人物領域から大まかな頭部領域を抽出し、安定にできる顔幅を基準に大きさの正規化を行い、各フレームでの顔の向き推定後、探索区間を設けて各登録人物との照合度を計算することで、各フレームでの正確な位置合わせを前提としない方法をとる。そして、各フレームでの評価結果の合計により個人の識別を行う。なお、本手法ではテンプレート照合（実際には正規化相互相関を用いているが、他のパターン認識手法を用いてもよい）を行っているが、顔の向きが大きく異なる顔画像間、奥行き異なる画像間では、結果の信頼性や安定性が異なるため、そのまま結果を積算することはできない。そこで、あらかじめ顔の向きと人物の奥行き（顔幅）と相互相関の関係をモデル化することで、時系列間のデータの統合を行っている。

2. 個人識別システムの概要

2.1 全方位視覚センサ HyperOmni Vision

全方位視覚センサ HyperOmni Vision は図 1 に示すように双曲面ミラーを使用して 360 度の方位を観測することができる。しかし、双曲面ミラーの反射面の深さ（観測方向（俯角）の違い）により得られる画像の

解像度が異なる。本論文で用いた HyperOmni Vision の場合、解像度特性は図 2 に示すとおりである。通常、我々が用いる標準レンズ（垂直画角が 30 度程度）を用いたカメラの場合、垂直画素数を 512 ピクセルとすると、角度分解能は約 15 ピクセル/度になる。それに対し、HyperOmni Vision は、最大でも 6 ピクセル/度と、3 分の 1 程度の解像度しかない。さらに、角度分解能は俯角が増すともない低くなる。すなわち、HyperOmni Vision は、シームレスに観測できる利点を持つ反面、空間分解能が低く、単一画像から詳細な認識を行うことが難しいという性質を持つ。

それに対し、以下で示す提案手法では、この全方位視覚センサの解像度の低さを補うために、時系列顔画像から総合的に評価を行い個人識別を行う。

2.2 個人識別アルゴリズムの前提条件

本手法では、人物が前方を見ながら遠方より近付いてきてカメラの横を横切る場合のように顔の向きが連続的に変化する画像列を対象とする。そして歩行時に顔は前後に傾斜しないものとする。また、今回は、照明条件が一定の屋内環境を人物は歩行するものとし、シーン中には、識別対象人物が 1 名しかいないものとする。HyperOmni Vision は、床面から高さ 1.6 m の位置に鉛直に設置した。

2.3 処理の流れ

本論文で提案する全方位視覚センサを用いた個人識別システムでは、まず全方位画像中から背景差分を利用して人物とその頭部を発見する。次に発見された人物の頭部を透視投影変換し、あらかじめ登録しておいた顔の向き変化が連続したテンプレート列から入力画像の顔の向きに近いテンプレートを探査する。テンプレート探索の際には、歩行時の顔の向きは連続的に変化するという前提条件から、直前の顔の向きを参考に次フレームでの探索範囲の限定を行う。そして、推定された顔の向きで、正規化相互相関法を用い、すべての登録人物の照合度を計算し、人物の奥行きと顔の向きを考慮した評価値の高い人物を候補とする。そして、

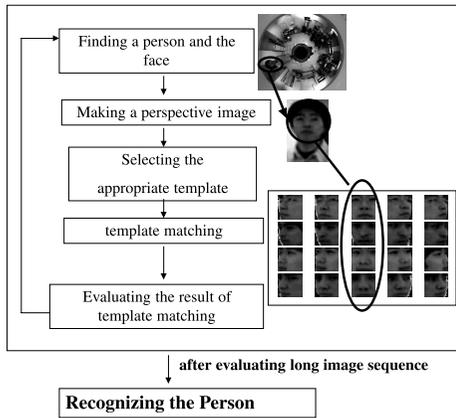


図 3 個人認識システム

Fig. 3 The recognition system.

解像度が低い場合、単一フレームで認識結果は不安定になりやすいことから、ある一定頻度以上、候補となり、かつ、評価値が高い人物を、個人識別結果とする。本システムの概要をまとめると、図 3 のようになる。

以下、全方位視覚センサを用いて得られた画像から人物を発見する方法、顔画像の切り出しを行う方法、さらに時系列顔画像から個人識別を行う方法について順に述べる。

3. 全方位画像からの人物と顔の発見

あらかじめ撮像しておいた全方位の背景画像との差分画像に対し、濃度ヒストグラムを作成し、動的しきい値により二値化する。変化領域が抽出された場合、歩行者として以後の処理を実施する。まず、全方位画像の中心から放射方向に変化領域の画素数を数数することで、人物方位を横軸とした 1 次元投影処理を行う。1 次元投影データに対し、平均値フィルタにより、スムージングを行ったうえで、頻度値最大の方位を人物方位 θ_p とする (図 4 (a))。ここで、HyperOmni Vision の設置高が 1.6 m の場合、標準的な日本人 (身長 170 cm 前後) が画像中で占める大きさは、接近した場合 (センサから約 30 cm) で、方位に換算して約 40 度程度になることから、人物方位角を中心に左右 20 度の角度範囲を次の処理範囲とする (図 4 (a))。図 4 (b) に示すように、上記で設定された処理範囲の差分二値化画像に対し、方位方向 θ に計数した後、平均値フィルタによりスムージングを行い、1 次元投影データを作成する。そして、本 HyperOmni Vision では、人物の全身が写る場合、足元での頻度値は最大頻度値の 10% 程度以上になることから、最大投影値から画像中心方向に最大投影値 y_b の 10% となる位置を探し、人物の足元とする。また頭頂位置 r_{top} も、頭部の幅が安

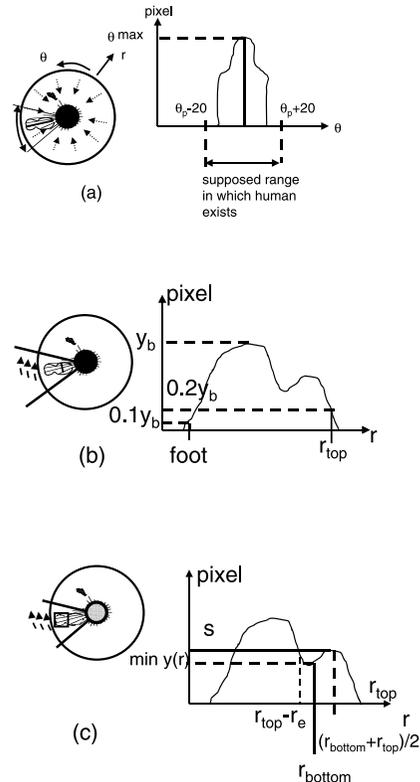


図 4 全方位画像からの人物の発見

Fig. 4 Human and face region detection.

定する平らな部分が頭頂となるように、経験的に最大頻度値 y_b の 20% となる位置を画像外周から探すことで求める。さらに、一般に首は、頭部や胸部に比べ細いことから、次式に示すように、顔の下端 (首位置) は、頭頂位置 r_{top} から、一定の範囲内で、頻度値が最小となる位置として抽出する。

$$y(r_{bottom}) = \min y(r) \quad (1)$$

$$(r_{top} - r_e \leq r \leq r_{top})$$

なお、 $y(r)$ は、画像中心から放射方向の位置 r での、1 次元投影データの値 (身体の幅) で、 r_e は、頭頂位置から首位置を探す範囲である (図 4 (c))。

ここで幅 s は、次式で定義することができる。

$$s = y((r_{top} + r_{bottom})/2) \quad (2)$$

以上の人物と頭部領域抽出結果の 1 例を図 5 に示す。また時系列での人物の観測方位角 (図 4 (a) 中 θ_p) と検出された顔の幅 (図 4 (c) 中 s) の時系列での推移結果の 1 例を図 6 に示す。また切り出した顔画像系列を図 7 に示す。

全被験者 15 名について 2 m 離れた地点から HyperOmni Vision の脇を通り過ぎるように歩行してもらった結果、図 6 に示すように人物の方位角および顔の幅は

滑らかに変化し、この結果より距離 2 m 以内での自然な歩行動作に対して安定に顔位置の切り出しができた。

以後述べる個人識別処理には、抽出された頭部を透視投影変換した画像を用いる。抽出された頭部の大きさは個人差、距離の違いなどで大きさが異なる。そこで、すべての抽出された顔領域に対して、顔の幅を基準に透視投影画像上で正規化を行う。透視投影の投影

中心は、頭部中心とする。

4. 全方位時系列画像を用いた個人の識別手法

全方位視覚センサは、周囲 360 度が常時観測できるため遠距離からの人物観測に加え、広範囲の顔の向きを追跡することができる。そして低解像度ながら長時間の多様な人物観測が可能である。この特性を生かし、歩行中に得られた連続的に顔の向きが変化する顔画像列から、個人識別を行う。具体的には、前述の方法により発見された顔領域に対し、顔の向き推定のためのテンプレート照合を行う。時系列間での顔の向き推定では、姿勢の連続性から探索範囲を制限し時系列間での顔の向き変化を推定する。次に各フレームでの顔の向き推定結果をもとに、すべての登録者との個人照合を行う。そして、時系列間で最も認識の良かった人物が識別結果として出力される。ただし、認識結果が悪い場合は、未認識としている。以下詳細について述べる。

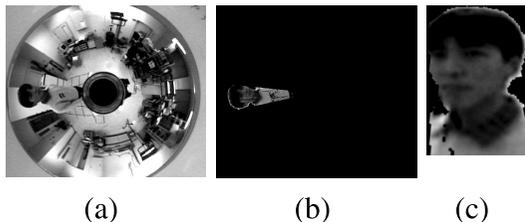


図 5 人物領域抽出結果例

Fig. 5 Extracting a person and the head; a) The image of HyperOmni Vision, b) Extracting a person from a), c) The perspective transformed face image.

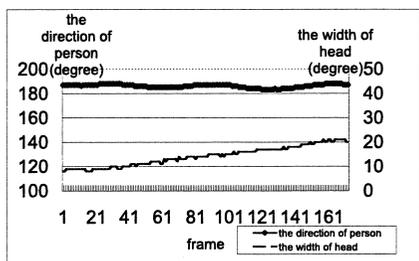


図 6 人物方位角と顔の幅

Fig. 6 A result of the direction and the width of head.

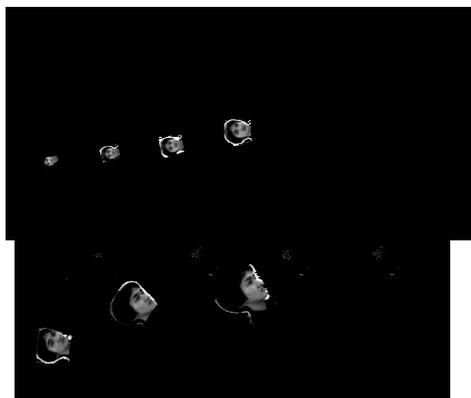


図 7 顔画像列抽出結果例

Fig. 7 A result of face region extraction.



図 8 個人用テンプレート画像

Fig. 8 An example of template images.

4.1 個人識別用テンプレート

学習データとして、連続して変化させた画像列を用いてテンプレート集合をつくる。今回の実験では、顔の向きを -60 度から 60 度 (正面顔を 0 度とする) まで約 6 度ずつ変化させた 20 枚の顔画像を用いた。図 8 に作成されたテンプレートの一例を示す。

テンプレートの作成は、上述の正規化された透視変換画像から手作業で行う。具体的には、正規化透視変換画像中で顔中央を指示し、この指示点を中心にして、図 8 に示すような 45×45 の矩形領域をテンプレート領域とした。

4.2 顔の向きに対する重み

4.4 節で述べる顔の向き推定手法では、テンプレート画像と入力画像との正規化相互相関値を計算し、相関値の高いテンプレート画像の顔の向きを入力画像の顔の向きと考えて、各フレームの顔の向きを推定している。

一般に顔の姿勢ずれがない場合には、高い相関値が得られるが、特に、顔が横向きときは、前提条件のように首を振らずに前方を見ながら歩いてもらっても、顔の上下の変化の影響を受けやすくなり、相関値が上がりにくくなる。そこで、本手法では、テンプレートの向きごとの微小変化への安定性を考慮し、各テンプレートに顔の向きに応じた重み付けを行う。

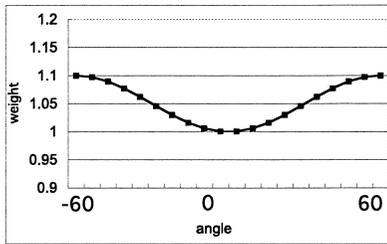


図 9 顔の向きを重み

Fig. 9 The weight for the angle of face of template images.

具体的には 15 名の人物について、各人物ごとに顔の向きが隣接するテンプレート間の相関を計算する。そして、この相関値の平均の逆数をテンプレートの姿勢に応じた重みとして用いる（図 9）。この結果から、正面顔ほど隣接する向きの顔との相関が高く、姿勢変化の影響を受けにくく、高い相関値が期待できることから重みが低くなり、横顔の方が姿勢変化により相関値が下がりやすいことから重みを上げていることが見てとれる。

4.3 顔のスケールの正規化と相関計算

前述のように本手法では、顔の幅を基準に顔の大きさの正規化を行っている。

具体的には、顔の幅が 90 ピクセルとなるように透視変換画像上で正規化している。なお、透視変換は、全方位画像より得られた人物の頭頂から首付近までの頭部領域を変換している。そして、この頭部領域内で、正規化相互相関法を用い、テンプレート照合を行っている。テンプレート領域の大きさは、 45×45 である。

4.4 顔の向き推定

フレーム t での入力画像と人物 p のテンプレート r^p との相関値を $K(r^p, t)$ とし、以下に初期フレームとそれ以降の顔の向きを推定法を示す。

- 初期フレーム ($t = 0$)

登録人物中の 10 名の代表人物 p に対してそれぞれ最も相関値が高いテンプレートを顔の向き候補として選ぶ。すべての代表人物の顔候補を顔の向きごとに投票し、顔の向きが 20° 以上離れた投票結果が上位 2 つのテンプレート（顔の向き）を初期フレーム $t = 0$ でのテンプレート r_t^* , r_t^{**} として選択する。

- 初期フレーム以降 ($t > 0$)

フレーム $t-1$ で対応づいたテンプレートを r とするとテンプレート $r-1, r, r+1$ を各々中心にその前後のテンプレートからなるテンプレートグループ G_{r-1}, G_r, G_{r+1} を作る（図 10）。ここでテンプレートの姿勢に応じた重み w_r （図 9）

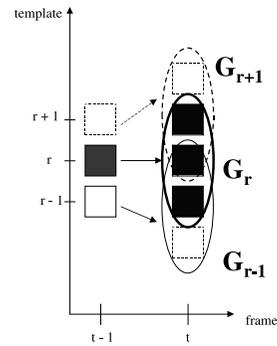


図 10 テンプレート選択方法

Fig. 10 Selecting the template image at the next frame considering around template at the prior frame.

を考慮し、人物 p のグループ G_r について $V(G_r^p)$ を次式で定義する。

$$V(G_r^p) = 1/3\{K((r-1)^p, t) \times w(r-1) + K(r^p, t) \times w(r) + K((r+1)^p, t) \times w(r+1)\} \quad (3)$$

そして同様に $V(G_{r-1}^p), V(G_{r+1}^p)$ を計算し、 V が最大のグループ中心のテンプレート（顔の向き）を人物 p のフレーム t での照合テンプレート r_{\max}^p とし、評価値を $Y(p, t)$ とする。

$$Y(p, t) = \max\{V(G_{r-1}^p), V(G_r^p), V(G_{r+1}^p)\} \quad (4)$$

登録人物のうち代表的な人物に対し、以上の手順でテンプレートを選択し、最も多く選択されたテンプレートをフレーム t での選択テンプレート（顔の向き） r_t^* とし決定する。次のフレーム $t+1$ では、フレーム t での選択テンプレート r_t^* をもとに同様の処理を行い、時系列の顔の向きを推定する。 r_t^{**} についても、上記の r_t^* と同様の処理を行い、顔の向きを推定する。

以上の処理では、 r_0^*, r_0^{**} から始まる 2 種類の顔の向きを推定することで、動的計画法やビームサーチと比べ信頼性はやや劣る可能性はあるものの、処理コストが小さい顔追跡が実現できる。

4.5 単一フレームでの識別処理

顔の向き候補 r_t^*, r_t^{**} のそれぞれについて以下の識別処理を行い、個人を識別する。

しきい値 D 以上で評価値が最上位の人物を探す。しきい値 D 以上の最上位人物が見つからなかった場合には認識不可能なフレームと見なす。このとき、各フレームでの個人識別値 $Q(p, t)$ は、以下のように定義する。

- 評価値が最上位の人物の場合

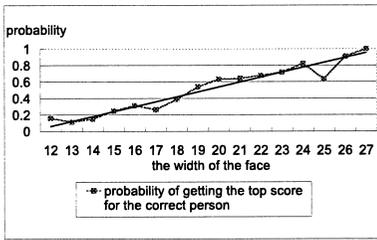


図 11 正しい人が最上位評価値になる確率

Fig. 11 Probability of getting the top score by correct person at each face size.

$$Q(p, t) = d(t)Y(p, t) \quad (5)$$

- 評価値が最上位以外の人物の場合

$$Q(p, t) = 0 \quad (6)$$

ここで、 $d(t)$ は、人物とカメラ間の距離の違いによる重み係数である。一般に、人物が遠方にいる場合、画像中での顔領域は小さくなり、他者との区別は難しくなる。すなわち、顔領域の入力画像中での大きさは、顔認識の信頼性に大きくかかわる。そこで、本手法では、人物とカメラ間の距離の違いを考慮した重み係数を顔認識の際に利用している。具体的には、あらかじめ、全方位画像中での顔幅（人物の奥行き）と認識率の関係を調べ、モデル化することで、重み係数 $d(t)$ として用いている。

フレーム t での全方位画像中での顔の幅を $s(t)$ とすると、重み係数 $d(t)$ は、15名の被験者に前方を見ながら直進してもらったときの画像列を用いて、正面から観察したときの顔幅 $s(t)$ と正規化相互相関法により個人識別を行ったときに正しい人物が選ばれる割合との関係を線形近似することで作成する。図 11 は、線形近似を行った結果である。図 11 から分かるように、カメラに近づくほどに認識率が上がり、信頼性が増している。距離の重み係数 $d(t)$ は

$$d(t) = b \cdot s(t) \quad (7)$$

とし、 b は図 11 の傾きとした。なお本実験では $b = 0.05$ であった。

4.6 時系列関係を考慮した個人識別処理

さまざまなノイズ要因ならびに顔の向きにより一見似かよって見える顔立ちの人の存在などにより、正しい人物が、他の人物より高い相関値をだすとは限らない。しかし、多くの顔画像があり、かつそれらが異なる方向から観察した画像列である場合、部分的に他者の相関が上位にくることはあっても、全体では、正しい人物が上位を占めることが期待できる。本手法では、この考えに基づき個人識別を行う。実際には、単一フレームの処理により個人識別が可能と判断されたフレームの個人評価値の平均値を、人物 p の個人識別度

$R(p)$ とし、以下の式で定義する。

$$R(p) = 1/F \sum_{i=0}^t Q(p, i) \quad (8)$$

ここで、 F は、評価値が最上位となったフレーム数である。

そして、この評価を r_0^* , r_0^{**} から始まる 2 種類の系列に適用し、いずれかの系列で、 F が一定以上で、個人識別度が所定のしきい値を先に超えた人物を最終的な個人識別結果とする。

ただし以下の場合には、識別不能として $R(p) = 0$ とする。

- 識別できないフレームが一定フレーム続いたとき
- 向きが -60 度または 60 度のテンプレートが一定時間選択され続けたとき

5. 評価実験

5.1 テンプレートの生成と方法

15名の被験者が全方位視覚センサの前 30 cm の位置で首を水平方向に振った系列を撮影する。この際、被験者には首をかきげないように、遠方に設置した目標物を見ながら首を振ってもらった。この首振り画像系列のうち、顔の傾きが -60 度から 6 度刻みで 60 度（正面顔を 0 度とする）までの画像列 20 枚をテンプレートとして使用する。

顔の向きの推定そして識別は顔を確認できる距離から行う。今回の実験で使用した HyperOmni Vision の場合には、全方位画像中で顔のテンプレート領域に相当する部分が 10×10 ピクセル（顔の幅が約 12 度）以上となったフレーム以降を識別処理の初期フレームとした。なお、平均的日本人（身長 170 cm）の場合、顔幅約 12 度は、センサから約 1.0 m の距離に被験者がたっている状況に相当する。

表 1 に示す画像系列 1 から画像系列 5 の歩行系列で評価実験を行った。歩行系列 1 は 13 名の被験者が 1 名 1 回の歩行をし、歩行系列 2 から歩行系列 5 については 4 名の同一被験者が 1 名 2 回の歩行を行った。顔の向きの変化が、全被験者同一になるように、被験者には遠方の目標物を見ながら歩行してもらった。画像系列 1 から画像系列 5 は、図 12 中の歩行経路と目標物の次の組合せである。

- 画像系列 1: A を見ながら I を歩行
- 画像系列 2: B を見ながら II を歩行
- 画像系列 3: C を見ながら II を歩行
- 画像系列 4: B を見ながら III を歩行
- 画像系列 5: C を見ながら III を歩行

表 1 実験画像系列パターン

Table 1 Walking patterns for the experiment.

	センサからの距離 (m)	顔の向き変化 (度)
画像系列 1	1.0~0.3	0~-60
画像系列 2	1.0~0.7	10~60
画像系列 3	1.0~0.7	30~60
画像系列 4	1.0~0.5	10~60
画像系列 5	1.0~0.5	30~60

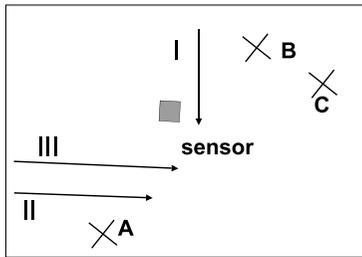


図 12 歩行方法

Fig.12 Walking passes and the targets.

全方位画像中の最も距離の変化の大きい画像系列 1 の場合には、顔の大きさは歩行の間に全方位画像中の顔でテンプレート領域に相当する部分が約 10 × 10 ピクセルから約 60 × 60 ピクセルまで変化した。

なお

- 顔の向きが -60 度以上かつ 60 度以内 (正面を 0 度とする)
- 顔の向きの変化は横方向に限定

とする。テンプレート画像、入力画像ともに 30 枚/秒の速さでサンプリングした。

5.2 顔の向き推定結果

5.2.1 顔の向き为重みの有効性の検証

画像系列 1 の顔の向き推定の結果を図 13 に示す。

図 13 は顔の向きに対する重み (図 9) をかけなかった場合と、重みをかけた場合の顔の向き推定例である。なお各テンプレートの顔の向きは手入力で与えた。重みをかけなかった場合にはテンプレートが正面顔付近からほとんど推移しなかったが、顔の向きに対して重みをかけた本手法の場合には真値と同様の顔の向きのテンプレートが選択されていることが分かる。よって顔の向きに対する重みは有効であるといえる。

同様に 13 名の歩行者に対し画像系列 1 で顔の向き推定について実験した結果を表 2 に示す。表 2 より重みをかけた場合には 12 名について顔の向きを正しく追跡することができた。

一方、重みなしの場合、正面顔から最大でも 20 度程度の変化の顔の向きまでしか追跡することができなかった。これは実際の歩行時には、顔の向きが横方

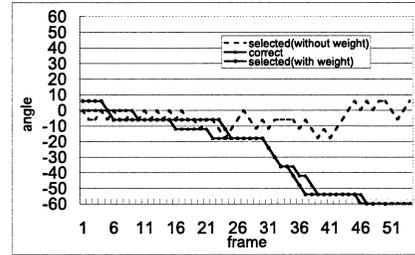


図 13 画像系列 1 顔の向き推定例 (人物 A)

Fig.13 An example of face pose estimation.

表 2 顔の向き推定へのテンプレートの重みの影響

Table 2 The result of face pose estimation;

left: with weight for the angle of face,

right: without weight for the angle of face.

顔の向き推定の成功確率	
重みあり	重みなし
92.3%	7.7%
(12/13)	(1/13)

表 3 顔の向き推定, 個人識別結果

Table 3 The result of a pose estimation and the recognition rate.

r_t^+, r_t^{**} のどちらかに正しい推定を含む確率		88.9% (40/45)	
	認識率	誤認識率	未認識率
r_t^+ で認識	35.0% (14/40)	2.5% (1/40)	7.5% (3/40)
r_t^{**} で認識	50.0% (20/40)	5.0% (2/40)	
合計	85.0% (34/40)	7.5% (3/40)	

きになると正面顔に比べて顔の前後の姿勢変化に対する影響が大きく単純に隣接するテンプレートの相関和をとただけでは正面の顔に比べて相関値が上がらなかったためと思われる。

図 14 は画像系列 1 の場合の歩行画像の 1 例である。左の画像列が全方位視覚センサの画像, 中央の画像列が全方位画像より人物の顔を発見し透視投影変換した画像, 右の画像列が透視投影変換画像との相関の結果より選択されたテンプレート画像である。

5.2.2 顔の向き推定

顔の向き推定は r_0^+, r_0^{**} の 2 つの推定を行うが, 2 つの顔の向きを推定することが必要かどうかを調べる。顔の向き推定結果と個人識別結果を表 3 に示す。

表 3 に示すように, 約 89% の割合で正しい顔の向き系列がいずれかの顔の向き系列に含まれていた。そして, 2 本目の顔の向き系列 r_t^{**} が有効に働いたケースが 50% もあった。

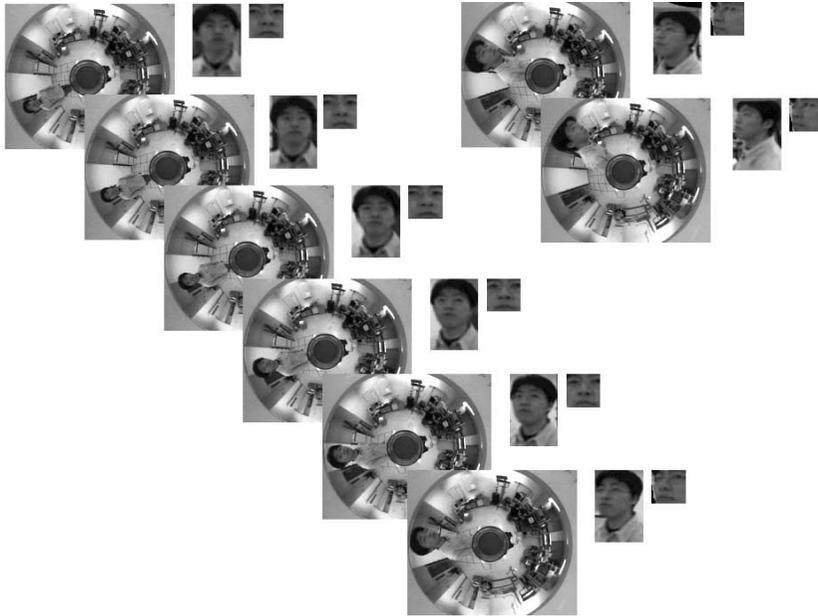


図 14 画像系列 1 顔の向き推定結果画像例

Fig. 14 Examples of input images and the pose estimation at each frame; left: the HyperOmni Vision images used as input images, center: the Perspective images of detected person in the input images, right: the Selected template images.

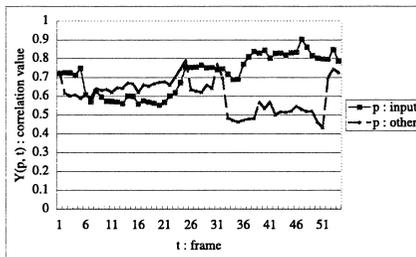


図 15 画像系列 1 評価値推移の例(人物 A)

Fig. 15 An example of the evaluate value (subject A).

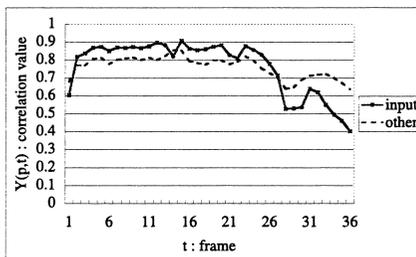


図 16 画像系列 1 評価値推移の例(人物 B)

Fig. 16 An example of the evaluate value (subject B).

5.3 個人識別結果

最初に画像系列 1 に関して、評価値と識別値の時系列の推移についてまとめる。図 15、図 16 は各々、人物 A、B が歩行したときの評価値の推移結果である。

評価値 $Y(p, t)$ の計算の際のしきい値 D は、予備実験から、経験的に $D = 0.75$ 以上とした。また最終的に特定される人物は最も早くに $F \geq 15$ で $R(p) \geq 0.80$ となる人物とする。input は、正しい人物のテンプレート画像での評価値、other は、正しい人物以外で最上位となる人物の評価値である。図 15 では、当初、遠方を歩行時は、他者との評価値が反転している場合があったが、センサに近づくに従い、正しい人物(人物 A)の評価値が上昇している。一方、図 16 では、遠方では、正しい人物の(人物 B)の評価値が高く、近くでは、歩行時の首の揺れによる影響がでたために評価値が下がるという結果がでた。このように単一フレームでは、評価値が安定しないが、連続で変化する時系列画像列から向きや距離を考慮し各フレームの評価を積算することでいずれの場合も個人識別ができた。

図 17 は、図 15 の場合の識別結果、図 18 は図 16 の場合の識別結果である。いずれの場合の正しい人物の識別値 $R(p)$ が、徐々に積算され、他者よりも安定に高い値を示し、最終的に正しく人物を特定することができた。

表 3 の個人識別結果よりさまざまな歩行パターンで、かつ複数の被験者により複数の実験を行った結果、85%の認識率が得られた。実験より、入力画像中の人物の距離により顔画像の解像度が変化し、テンプレ

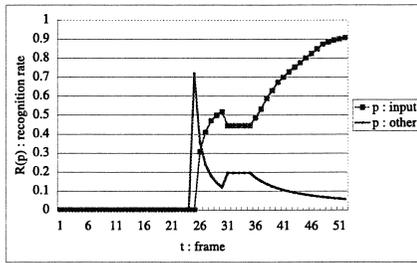


図 17 画像系列 1 識別値の例 (人物 A)

Fig. 17 An example of the recognition rate (subject A).

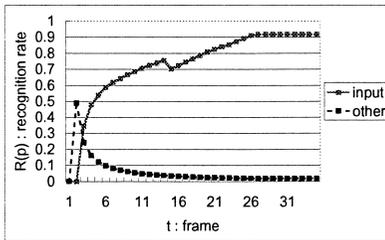


図 18 画像系列 1 識別値の例 (人物 B)

Fig. 18 An example of the recognition rate (subject B).

ト作成時と異なる解像度の顔画像であったが、個人識別が可能であることが示された。また、全方位視覚センサにより得られた顔画像の解像度は低かったが、一般の正規化相互相関法による認識率を考えると、十分な結果が得られたものと考えられる。

6. まとめ

本論文では、全方位視覚センサを用いた顔画像による個人識別手法を提案した。一般に全方位視覚センサを用いて個人識別を行う場合、入力画像の空間分解能が低いことから、単一フレームでは個人識別が困難であると考えられてきたが、長時間、顔の向きが変化している画像列を観測し、各フレームの評価結果を時系列で総合評価することで、個人識別が可能であることを示すことができた。また、歩行時の画像を扱う場合、顔の向きと顔の大きさの変化が大きくなるという新たな問題が生まれたが、顔の向きに関するモデル(重み係数)と顔の大きさの変化に関するモデル(重み係数)を設けることで、安定な認識を可能にした。

実験の結果、顔の向き推定に関しては、向きの連続した画像をテンプレート列として持つことにより顔の向き追跡率 88.9%という結果が得られた。認識に関しては、顔の向きが正しく追跡できている場合には識別率 85.0%、誤認率 7.5%、未認識率 7.5%という結果が得られた。よって顔の向き追跡を含めた本システムの識別率は 77.3%である。

また今回は認識手法として正規化相互相関法を用いたが、個人認識研究で高い認識率が報告されている PCA, ICA, FFT などの特徴量を用いることでより高い認識率が得られることが期待できる。今後、他の特徴量を用いた場合についても検討したいと考える。

今回は、識別対象人物は 1 名としたが、全方位視覚センサの魅力の 1 つは、同時に複数の人物を観察できることもある。複数の人物追跡方法に関しては、さまざまな手法がすでに提案されているので、それらの手法を参考に、多人数に対応できるシステムに拡張することは十分に可能であると考えられる。

また、今後は登録人物を増やし、実用的なシステムとするために大規模なデータに対して評価を進めていく。さらに今回は横方向の顔の向き変化のテンプレートを用いたが、実際の歩行では顔の上下変化も考えられ、自然な歩行画像から個人識別ができるべく、今後は上下の変化も考慮したいと考える。

参考文献

- 1) Nayer, S.K. and Boulton, T.E.: Omnidirectional VSAM System, *Proc. DARPA Image Understanding Workshop*, pp.55-61 (1997).
- 2) Onoue, Y., Yokoya, N., Yamazawa, K. and Takemura, H.: Visual Surveillance and Monitoring System Using an Omnidirectional Video Camera, *Proc. 14th IAPR Int. Conf. on Pattern Recognition*, pp.588-592 (1998).
- 3) Miki, R., Yokoya, N. and Takemura, H.K.: A Real-time Surveillance and Monitoring System using Multiple Omnidirectional Video Cameras, *Proc. ACCV*, pp.528-534 (2000).
- 4) Ng, K.C., Ishiguro, H., Trivedi, M. and Sogo, T.: Monitoring Dynamically Changing Environments by Ubiquitous Vision System, *Proc. Workshop on Visual Surveillance*, pp.67-73 (1999).
- 5) Sogo, T. and Ishiguro, H.: Real-time Target Localization and Tracking by N-Ocular Stereo, *Workshop on Omnidirectional Vision*, pp.153-160 (2000).
- 6) 十河卓司, 石黒 浩, モーハン M. トリペティ: 複数の全方位視覚センサによる実時間人間探索システム, 電子情報通信学会, Vol.J83-D-II, No.12, pp.2567-2577 (2000).
- 7) 寺沢征彦, 山澤一誠, 竹村治雄, 横矢直和: 複数の全方位画像センサを用いた遠隔監視システムにおける複数移動物体の存在領域推定, *PRMU*, p.195 (2001).
- 8) Kopilovic, I., Vagvolgyi, B. and Sziranyi, T.: Application of Panoramic Annular Lens for Motion Analysis Tasks, Surveillance and

- Smoke Detection, *Proc. ICPR*, pp.714-717 (2000).
- 9) Nishimura, T., Yabe, H. and Oka, R.: Indexing of Human motion at Meeting Room by Analyzing Time-varying Images of Omni directional Camera, *Proc. 4th Asian Conference on Computer Vision*, pp.1-5 (2000).
- 10) Stiefelhagen, R., Yang, J. and Weibel, A.: Simultaneous Tracking of Head Poses in a Panoramic View, *Proc. ICPR*, pp.726-729 (2000).
- 11) 横山太郎, 小原ゆう, 八木康史, 谷内田正彦: 全方位画像列からの個人識別, *CVIM-125*, pp.119-124 (2001).
- 12) 赤松 茂: コンピュータによる顔の認識—サーベイ, 電子情報通信学会, Vol.80, No.8, pp.2031-2046 (1997).
- 13) Turk, M. and Pentland, A.: Eigenfaces for recognition, *Cognitive Neuroscience*, Vol.3, No.1, pp.71-86 (1991).
- 14) Maurer, T. and von der Malsburg, C.: Single-view based recognition of faces rotated in depth, *Proc. International Workshop on Automatic Face and Gesture recognition*, pp.24-253 (1995).
- 15) ダデット プラマディハント, 呉 海元, 谷内田正彦: 多様な姿勢を有する入力画像からの個人識別, 電子情報通信学会, Vol.80, No.8, pp.2232-2238 (1997).
- 16) Weng, J., Evans, C.H. and Hwang, W.S.: An Incremental learning method for face recognition under continuous video stream, *Proc. IEEE Conf. on Automatic Face and Gesture Recognition*, pp.251-256 (2000).
- 17) 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認識, 電子情報通信学会, Vol.82, No.4, pp.613-620 (1999).
- 18) Yamaguchi, O., Fukui, K. and Maeda, K.: Face recognition using temporal image sequence, *Proc. International Conf. on Automatic Face and Gesture Recognition*, pp.318-323 (1998).
- 19) Huang, K.S. and Trivedi, M.M.: Networked omnivision arrays for intelligent environment, *Proc. Applications and Science of Soft Computing IV* (2001).

(平成 13 年 9 月 10 日受付)

(平成 14 年 3 月 8 日採録)

(担当編集委員 長尾 健司)



小原 ゆう

2001 年大阪大学基礎工学部システム科学科卒業。現在, 同大学大学院基礎工学研究科修士課程在学中。主としてコンピュータビジョンの研究に従事。



八木 康史 (正会員)

1983 年大阪大学基礎工学部制御工学科卒業。1985 年同大学大学院修士課程修了。三菱電機(株)産業システム研究所を経て, 1990 年大阪大学基礎工学部助手。現在, 同大学大学院基礎工学研究科助教授。1995~1996 年オックスフォード大学客員研究員。全方位センシング, 画像理解, 知能ロボットに関する研究に従事。工学博士。



横山 太郎 (正会員)

1996 年大阪大学基礎工学部システム工学科卒業。1998 年同大学大学院基礎工学研究科修士課程修了。主としてコンピュータビジョンの研究に従事。工学博士。



谷内田正彦 (正会員)

1971 年大阪大学大学院工学研究科修士課程修了。同大学基礎工学部制御工学科助手, 助教授を経て, 現在, 同大学大学院基礎工学研究科システム科学分野教授。1967~1968 年デนมーク原子力研究所留学。1972~1973 年米イリノイ大学にて Research Associate。1980~1981 年西独ハンブルグ大学 Research Fellow。1982 年米ミネソタ大学 CDC Professor。コンピュータビジョン, 人工知能, 移動ロボットの研究に従事。工学博士。