

# 顔領域追跡アルゴリズムのハードウェア上への実装について

山本貴彦\*, 川村尚生, 菅原一孔 (鳥取大学)

Hardware realization of face area tracking algorithm

Takahiko Yamamoto, Takao Kawamura, Kazunori Sugahara (Tottori University)

Face image processing are the topics, which are actively being researched to develop intelligent systems such as novel human machine interfaces. Face area tracking of arbitrary image sequences is the one of basic techniques of the face image processings. Variety of the face area tracking methods have been proposed, however, no hardware system has been developed. In order to decelop human machine interfaces, it is effective that camera itself has the face area tracking function. In this paper, the hardware realization of face area tracking function is proposed and the feasibility of the proposed system is confirmed by experiments.

キーワード：顔領域特定，FPGA，遺伝的アルゴリズム

**Keywords:** face area extraction , FPGA , genetic algorithm

## 1. はじめに

現在，顔画像処理技術は様々な分野で利用されつつあり，ヒューマンマシンインターフェースや監視システムなどへの応用を目指し，さらなる高機能化のため活潑に研究されている．顔画像処理を行うにあたり，その最も基本的な処理として，任意の画像中から顔領域を抽出したり，それを追跡するものがある．通常，この顔領域抽出や追跡などの顔画像処理は，すべてソフトウェアにより実現されている．しかし，顔領域特定をソフトウェア処理で行うと，処理速度や精度がPCの計算能力に依存してしまう．またコンピュータが必須となりシステムの構成が複雑になることが予想される．

実際にヒューマンインターフェースなどでの利用を考えると，カメラなどのハードウェア側に顔領域抽出や追跡の機能を持たせるということも一つの有効な手段として考えられる．そこで本研究ではビデオカメラからの入力画像から顔領域抽出の処理を行うことができるハードウェアの開発を目指している．ハードウェア化することによって，顔画像処理のためのシステムを小型化することが可能になり，システム全体としての処理速度の向上も期待できる．今まで我々は顔領域抽出処理の基本的なハードウェア開発について検討してきた<sup>(2)(3)</sup>．本稿では顔領域抽出アルゴリズムのハードウェア上への実装について述べ，その動作実験結果について報告する．

## 2. 想定している利用環境とシステム構成

**2.1 想定している利用環境** 本研究で開発するハードウェア装置の利用方法は，室内におけるTV会議での利用を想定している．遠隔地で会場を撮影した映像を見ている側からすれば，発話者の様子だけでなく会議全体の雰囲気を知ることは重要である．しかし，カメラ1台で会場全体を撮影しようとするとき，画面に表示される出席者が小さくなってしまい，発言者の表情を読み取ることができなくなる．逆に出席者がある程度の大きさで撮影しようとするとき，カメラの撮影画角に限度があるため会場全体を撮影することができない．この2つの状況を改善するためにはカメラを複数台用意して，1台で会場全体を広く撮影し，もう一方で発話者を近くから撮影する必要がある．この場合，発話者が変わった場合は手動でカメラを操作しなければならない．

この点を改善するため，本研究ではNTSCビデオカメラを用いて撮影した画像の中から顔領域を特定する，ハードウェア装置の開発を目指す．その動作はNTSCビデオカメラで会場全体を撮影し，発話者の顔領域の特定を行い，その様子を拡大して表示する．このハードウェア装置によると，2種類の映像をモニタに表示することが可能となり，会場の雰囲気の詳細な情報を得ることが期待できる．本稿では顔領域の特定処理のハードウェア化について述べる．

**2.2 システム構成** 本研究におけるシステムの構成を図2に示す．本システムはNTSCビデオカメラ，デコー

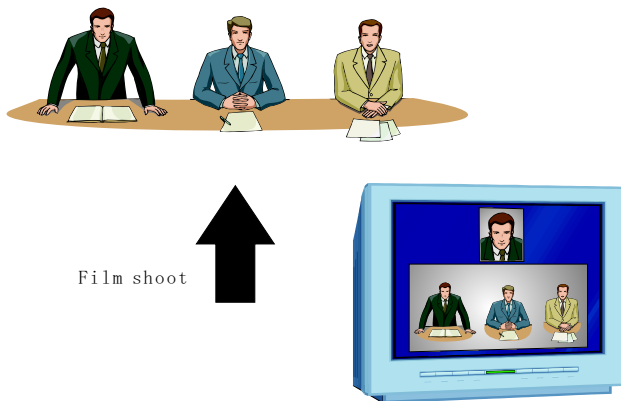


図 1 想定している利用環境  
Fig. 1. Assumed use situation.

ドボード, SSRAM, FPGA ボードから構成されている. NTSC ビデオカメラから送られてきたアナログ映像信号をデコードボードを介することによってデジタルの RGB 各 8bit のデジタル信号に変換し FPGA ボード上の SSRAM に蓄積する. そして入力された画像データ中から顔領域特定の処理を行い, FPGA 上の各回路において特定した画像を PC へ転送しディスプレイ上に表示する.

### 3. 顔領域の特定方法と回路構成

本研究における顔領域の特定方法としては, あらかじめ  $120 \times 150$ [pixel] の顔の画像をテンプレート画像として用意しておく. そして,  $640 \times 480$ [pixel] の入力された画像内からテンプレート画像に類似した領域を比較し特定するテンプレートマッチングを遺伝的アルゴリズムを用いて探索を行った<sup>(1)</sup>.

FPGA 上には遺伝的アルゴリズムを実現するために乱数発生回路, 初期個体生成回路, 適応度評価回路, 選択回路, 交叉・突然変異回路をそれぞれ作成した. 処理の手順とし

ては初期個体生成回路内に含まれる乱数発生回路で作成した乱数を個体の遺伝子コードとして設定する. 次に適応度評価回路で設定された個体情報をもとに入力画像とテンプレート画像を比較して評価値を各個体に付加する. そして親個体選択回路では交叉に用いる個体の選出を行う. 送られてきた個体を利用して交叉・突然変異回路で交叉および突然変異の遺伝的操作を実行し, 次世代の個体を作成する. その後, 適応度評価回路において新たに作成された個体に対して評価を付加するという一連の処理を繰り返し行う. この時に用いた遺伝的アルゴリズムのオペレーションとしては選択方法はランク戦略, 交叉方法は一様交叉, 突然変異は一点突然変異である. 以下にそれぞれの回路について詳しく説明する.

**3.1 乱数発生回路** M 系列を用いて乱数を作成している. その方法としては 10bit の "0", "1" で構成されている数列から 10bit 目と 6bit の排他的論理和をとり, その後全体を 1bit 左シフトする. そして排他的論理和の結果を 1bit 目に代入する. この操作をくり返すことで乱数を作成しており, 10bit 全体で 2 進数で表現した一つの乱数を構成している.

**3.2 初期個体生成回路** 入力画像中で探索に用いる個体数は 32 とし, 作成した乱数を個体に設定する. 各個体の遺伝子コードは入力画像中で探索に指定する領域の左下端の  $x, y$  座標を表しており, 前件部が  $x$  座標, 後件部が  $y$  座標で構成されている. また入力画像の左下を  $(0, 0)$  として横方向を  $x$  座標, 縦方向を  $y$  座標として定義しており, このとき指定する領域はテンプレート画像と同じサイズである. また指定領域が入力画像からはみ出した場合や入力画像外を探索するような座標値が得られた場合は除外し, 再設定している. 遺伝子コードの構成を図 3 に示す.

**3.3 適応度評価回路** それぞれの個体で指定した領域において RGB 値を読み取りテンプレート画像と一画

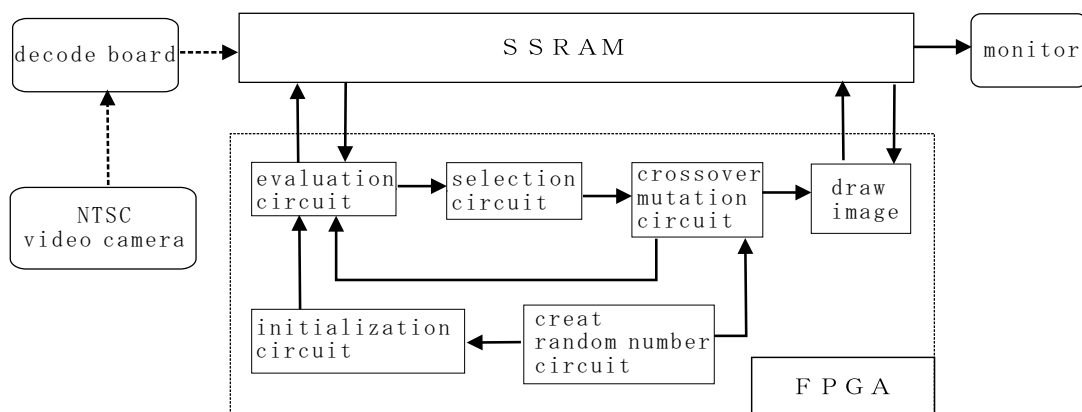


図 2 システム構成  
Fig. 2. System architecture.

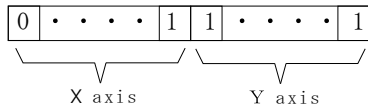


図 3 遺伝子コード

Fig.3. Genetic code.

素ごとに比較を行う．比較方法としては RGB 成分それぞれで差分をとり，その領域における差の絶対値和を評価値として各個体に付加する．このとき評価値の値が小さい方がテンプレート画像により近い領域であると考え，設定された評価値が小さいものほど適応度が高い個体とした．また評価値にしきい値を設定し，差分の加算結果がそのしきい値を越えると指定した領域はテンプレート画像とは似ていない領域とみなすこととする．そしてその後の比較処理を中断し，評価値には最大の値を設定するようにしている．比較する領域の指定を図 4 に示す．

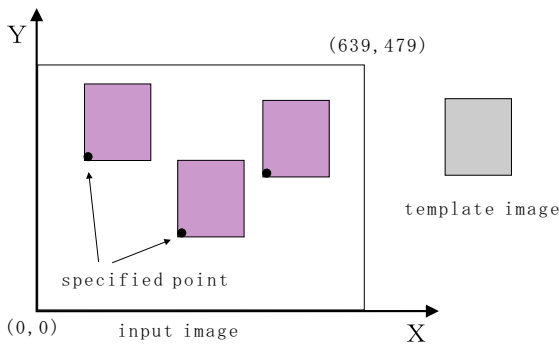


図 4 比較する領域の指定

Fig.4. Comparison of input and template image.

**3.4 親個体選択回路** 適応度評価回路において設定された評価値をもとにして各個体を昇順に並び換える．そして 1 番目から 16 番目までの上位 16 個体の中から，乱数で交叉に用いる親個体となる 2 個体を選び出す．このとき親個体として選択される確率はランクづけされた順位に依存している．

**3.5 交叉・突然変異回路** 交叉の様子を図 5 に示す．親個体撰択回路から選ばれた 2 個体と乱数の bit 列を用いて一様交叉を行う．乱数はマスクの役割をしており，各 bit をチェックし”1”となっている箇所のみ親個体の値の入れ換えを行い子個体を作成する．

次に突然変異の様子を図 6 に示す．交叉を終えた個体に突然変異を行うかどうかを決定するために各個体に対して乱数を与える．そしてその乱数が突然変異率を満たしていれば，bit 反転を行う遺伝子座をさらに乱数を用いて決定する．その後 32 個体のうち後半の 17 番目以降の個体に新たな個体として上書する．

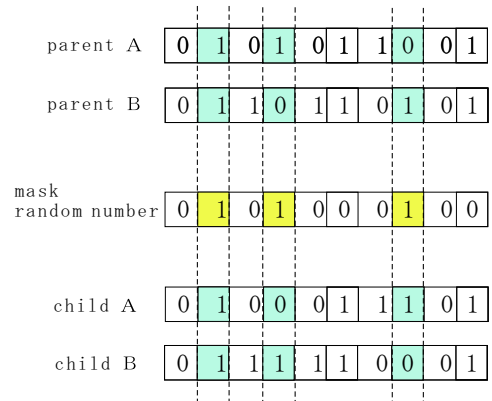


図 5 交叉部

Fig.5. Cross over block.

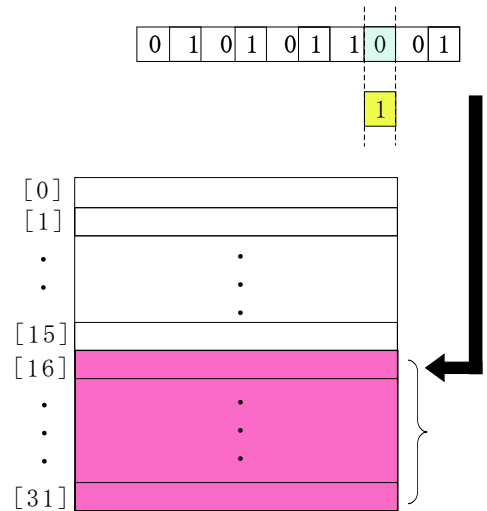


図 6 突然変異部

Fig.6. Mutation block.

以上の作業をくり返すことによって次世代の個体群の作成を行う．適応度が最も高かった個体を特定した顔領域としている．

**3.6 回路規模** これまでに述べた遺伝的アルゴリズムを構成するのに要した各回路の logic element 数を表 1 に示す．

表 1 回路の logic element 数  
Table 1. Logic elements of circuit.

回路名	logic element 数
乱数発生回路	883
初期個体生成回路	293
適応度評価回路	905
選択回路	3603
交叉・突然変異回路	1989
合計	7643(92%)

#### 4. 使用機器

本研究で使用している機器を図7に示す。FPGAボードに搭載しているFPGAはAltera社製のAPEX20KCである。このFPGAにおける使用可能なlogic element数は8320であり、これは回路規模20万ゲートに相当する。またFPGAボードには36bit幅の18MbitSSRAMが2つ搭載されており、それぞれバスが独立しているため2つのメモリに同時アクセスが可能である。NTSCビデオカメラはμtec社製のIVC-400を使用しており、有効画素数768×494[pixel]である。またデジタルのRGB信号に変換するデコードボードをFPGAボードの上に実装している。

論理回路設計はInnoveda社のVisual Eliteを用いてVHDLで記述し、論理合成はSynplicity社のSynplify Pro 7.3.1を使用した。配置配線、検証、FPGAへの書き込みにはAltera社のQuartusII2.2を用いた。

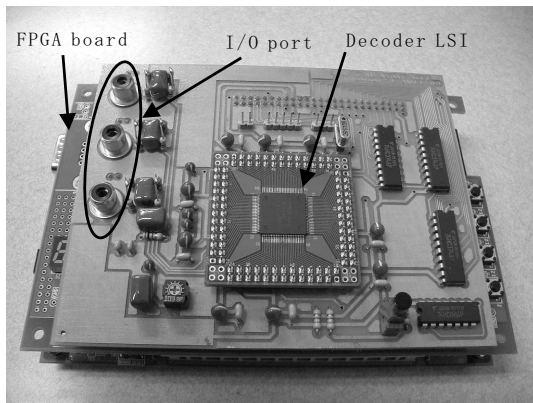


図7 FPGAボードとデコードボード  
Fig. 7. FPGA board and decode board.

#### 5. 動作実験

FPGA上に作成した回路の動作実験として数種類の画像の中から顔領域の探索を行った。その手順としては、まずPCから入力画像とテンプレート画像のRGB情報をFPGAボード上のSSRAMに転送し、それぞれ決められた領域に格納する。次にそれらの情報をもとにしてFPGA上の回路を動作させることによって顔領域の探索を行う。そして最後に顔領域が正しく特定できているかを検証するために、画像に特定した領域を線で囲みPCへと転送し、ディスプレイ上に表示する。このとき実験に使用した入力画像のサイズは640×480[pixel]で、テンプレート画像のサイズは120×150[pixel]である。

5.1 実験1 実験1で用いたテンプレート画像を図8に示す。テンプレート画像Aは入力画像中で被写体

として写っている人物と同じ人物の顔の画像をテンプレート画像として用いている。一般にテンプレートマッチングでは検出対象とテンプレートが異なると認識処理に影響が生じるという問題があるので、前処理としてテンプレート画像にぼかす処理を行っている。このような処理を行った画像を用いることによって、多少の異なりがあったとしても特定に大きく影響しないため、テンプレートマッチングでは有効な手法として用いられている。

このときの実験結果を図9～図11に示す。図9は探索の初期の状態である第1世代目の結果である。図10は途中経過にあたる第5世代目の結果である。図11は探索がほぼ収束した状態である第10世代目の結果である。世代数が進むにつれて収束してゆく様子が確認できる。



図8 テンプレート画像A  
Fig. 8. Template image A.

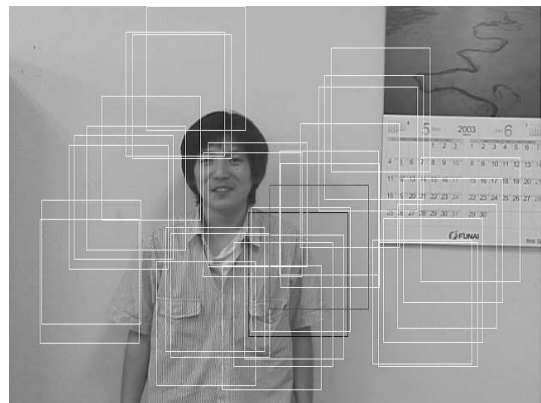


図9 初期状態  
Fig. 9. Initial state.

5.2 実験2 実験1で行ったようにテンプレート画像に特定の人物の顔の画像を使用すると、カメラでその人物を撮影した場合は精度よく顔領域を探索できるが、テンプレート画像とは異なる人物を撮影すると、その人の顔領域を特定することが困難になるのではないかと危惧される。そこで実験2ではテンプレート画像として20代前半の男性10人の顔の画像をもとにして、平均顔の画像を作成

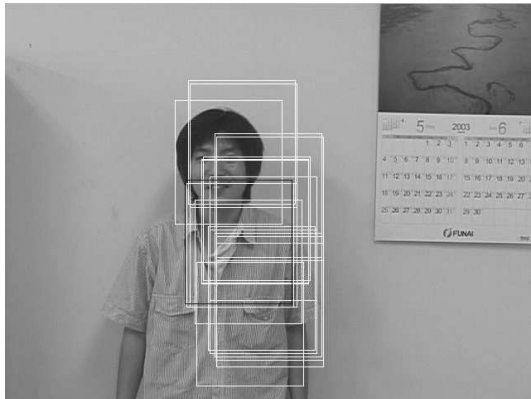


図 10 第 5 世代  
Fig. 10. The 5th generation

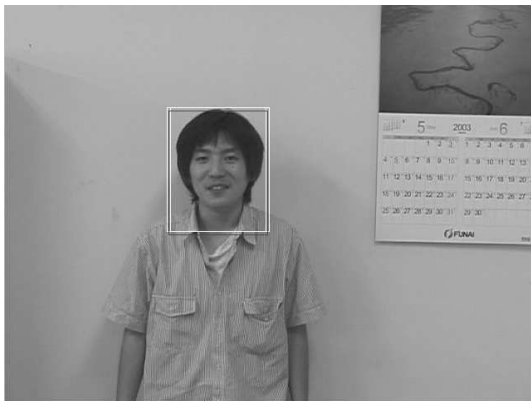


図 11 最終結果 (第 10 世代)  
Fig. 11. Final result.(The 10th generation)

しテンプレート画像 A と同様にぼかす処理を行った顔画像を使用した．実験 2 で用いたテンプレート画像 B を図 12 に示す．

入力画像は実験 1 で被写体として写っていた人物とは別の人物を被写体としている画像を用いた．このときの実験結果を図 13 に示す．ここでも人物の顔付近に収束していることが確認できる．

次にカメラで撮影されている人物が画像中で左から右へと移動している様子の数枚の画像の中から顔領域の探索を行った．このときの被写体は実験 1 の入力画像と同じ人物である．実験結果を図 14～図 16 に示す．人物の動きに合わせて収束した領域も顔を追跡して動いている様子が確認できる．

5・3 実験 3 実験 1, 実験 2 では入力画像中に人物は 1 人しかいなかったが, 実験 3 では入力画像中に人物が 2 人いる場合の画像を使用した．このときの実験結果を図 17 に示す．また実験 3 のテンプレート画像には, 実験 2 で用いたものと同じ平均顔のテンプレート画像 B を使用し

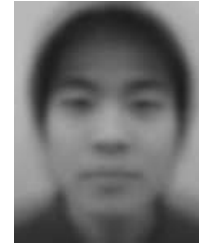


図 12 テンプレート画像 B  
Fig. 12. Template image B.



図 13 最終結果  
Fig. 13. Final result.



図 14 追跡の様子 1  
Fig. 14. Face tracking1.

た．入力画像内の 2 人のそれぞれの顔付近に収束していることがわかる．

5・4 考 察 実験 1 より回路が正常に動作していることが確認できた．またこのときテンプレート画像と同じ人物の顔を入力画像内から探索したが, 10～15 世代で収束するという結果が得られた．このことから画像一枚に対して 15 世代以内で顔領域を抽出することができると思われる．さらに回路動作開始から終了までの時間を計測した．計測作業を 20 回行いその平均をとることによって

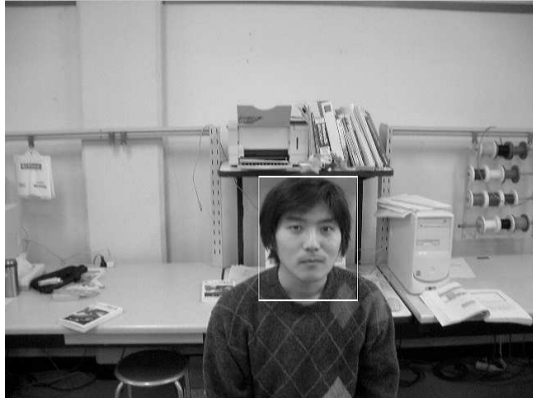


図 15 追跡の様子 2  
Fig. 15. Face tracking2.

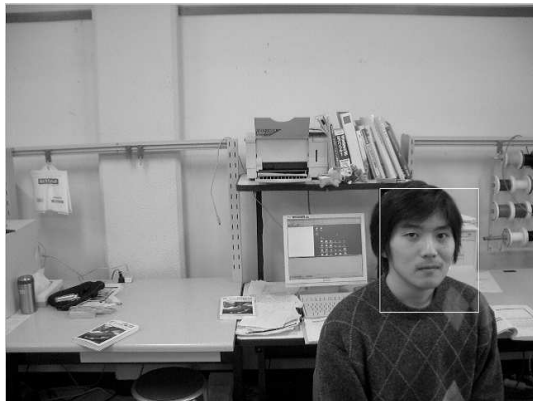


図 16 追跡の様子 3  
Fig. 16. Face tracking3.

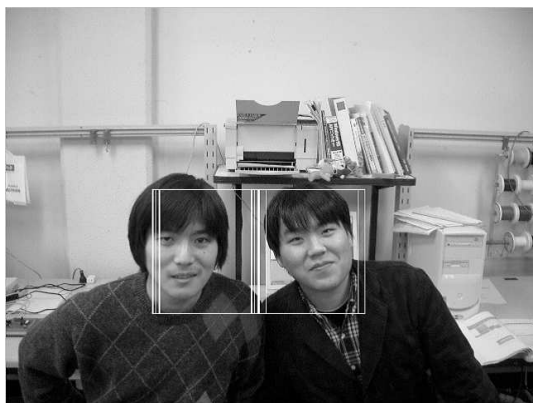


図 17 2人の場合  
Fig. 17. Face extraction result of two persons.

処理にかかる時間とした。その結果、現段階では約 2.90 秒要することがわかった。この原因として評価回路における比較処理に時間を要していると考えられるので、余分な処理を削減し速度を向上する必要がある。

実験 2 では多少のずれはあるものの人物の顔付近への収束が見られた。平均顔のテンプレート画像を用いることで特定の人物だけでなくとも対応する事が可能であり、平均顔の使用が有効な手段であるといえる。また動画は静止画像が連続して送られてくると考えると、フレームごとに顔領域特定の処理を行えば静止画と同様に扱うことができる。そのためビデオカメラで撮影している際に人物が移動した場合でも顔領域の追跡が可能であると考えている。

現在は適応度の最も高い個体を特定した顔領域としているが、実験 3 のように画像内で複数の箇所に収束が見られるような場合も想定される。このとき、適応度が最良の個体が示す領域からある程度離れた位置であれば、適応度が多少劣った個体でも顔領域として認識させることで、複数の顔領域の特定が可能になると考えられる。

## 6. おわりに

本稿では顔領域追跡アルゴリズムのハードウェア実装における回路動作について数種類の画像を用いて検証した。実験によって提案手法が顔領域の特定に有効に働いていることが確認できた。今後の課題として、まず回路規模の縮小ならびに処理速度の向上を目指し、開発した回路の最適化を進める。次に、本稿で用いた顔領域の特定方法では、顔を横に向けた場合や俯いた場合に顔領域の追跡が困難になることが予想される。そこで、正面顔画像以外でも精度のよい顔領域特定が行えるようシステムを改善してゆくことも今後の課題として残されている。

## 文 献

- (1) 長尾智晴, “進化的画像処理,” 昭晃堂, 2002.
- (2) 山本貴彦, 佐々木悠介, 川村尚生, 菅原一孔” 顔領域追跡アルゴリズムのハードウェア実現に関する研究”, 電気・情報関連学会中国支部第 54 回連合大会講演論文集
- (3) 山本貴彦, 佐々木悠介, 川村尚生, 菅原一孔” ハードウェア実現された顔領域特定アルゴリズムについて”, 第 5 回 IEEE 広島支部学生シンポジウム, 2003, pp.213-214.