

顔領域抽出機能を備えたビデオカメラの開発 Development of Video Camera with Face Area Extraction Function

山本 貴彦[†]
Takahiko Yamamoto

川村 尚生[‡]
Takao Kawamura

菅原 一孔[‡]
Kazunori Sugahara

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発に研究されている。顔画像処理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、それを追跡することが挙げられる。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で行うと、処理速度が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装置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。顔領域の抽出方法にはカメラからの入力画像に対して、あらかじめ用意した顔画像のテンプレートと比較し、類似した領域を抽出するテンプレートマッチングを用いている。また探索方法には大域的な探索と局所的な探索が可能な遺伝的アルゴリズムを使用しており、入力画像内の端から全領域を比較するよりも効率よく解を得ることができると考えられる。

開発するハードウェア装置の利用方法の一例として、図1に示すように、室内におけるTV会議での利用が考えられる。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が重要である。顔領域抽出機能を備えたハードウェア装置が実現すると、発話者の様子与会場全体の様子という2種類の映像をモニタに表示することが可能となり、会場の雰囲気のより詳細な情報を得ることが期待できる。

今まで我々は顔領域抽出処理のためのハードウェア開発について検討を重ねてきた[2][3]。本稿では顔領域抽出アルゴリズムのFPGA上への実装とその動作について報告する。

2. システム構成

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

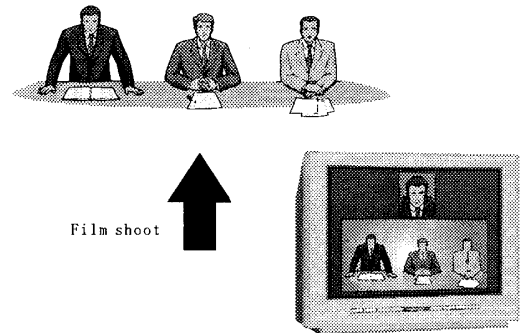


図1: 開発するハードウェア装置の利用例

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は次のとおりである。あらかじめ 120×150 [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、 640×480 [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[1]。

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。処理の手順を以下に示す。

1. 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。
2. 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。
3. 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。
4. 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。
5. 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す。

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

[†]鳥取大学大学院工学研究科知能情報工学専攻

[‡]鳥取大学工学部知能情報工学科

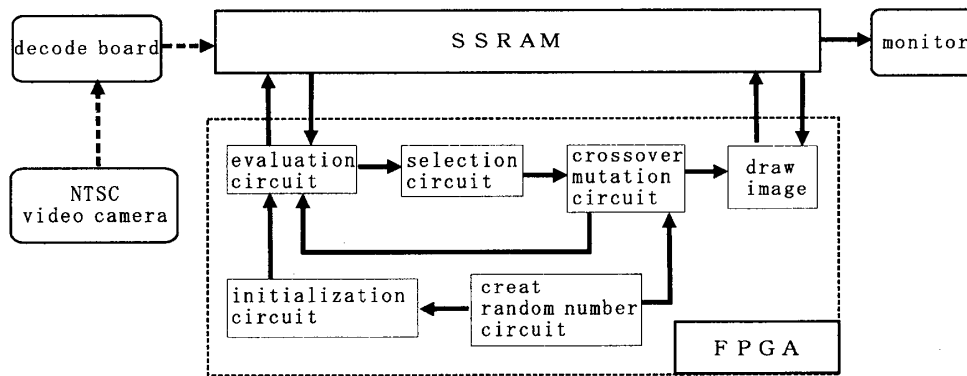


図 2: システム構成

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。

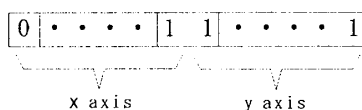


図 3: 遺伝子コード

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。

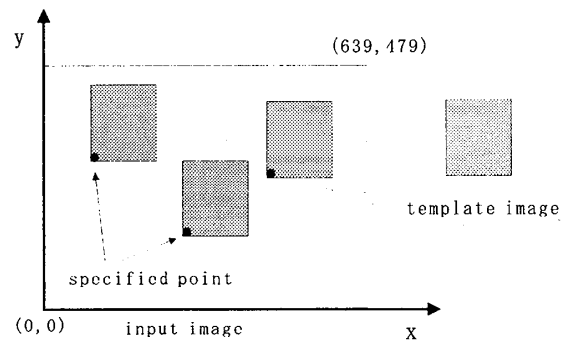


図 4: 比較する領域の指定

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

親個体選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。

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きする。

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

回路名	logic element 数
乱数発生回路	883
初期個体生成回路	262
適応度評価回路	890
選択回路	1244
交叉・突然変異回路	1812
合計	5091(61%)

4. 使用機器

本研究で使用している機器を図5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ので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 μ tec 社製の IVC-400 を使用しており、有効画素数 $768 \times 494[\text{pixel}]$ 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。

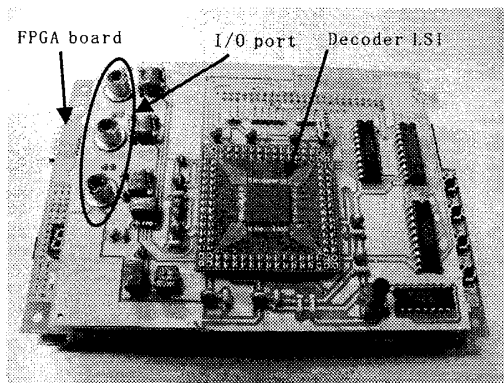


図 5: FPGA ボードとデコードボード

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抽出を行った。その手順としては、カメラの代替として 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 $640 \times 480[\text{pixel}]$ で、テンプレート画像のサイズは $120 \times 150[\text{pixel}]$ 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6に示す。テンプレート画像は20代前半の男性10人の顔の画像をもとにして作成した、平均顔の画像を使用した。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7～図9に示す。図7は探索の初期の状態である第1世代目の結果である。図8は途中経過にあたる第5世代目の結果である。図9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。また回路動作開始から終了までの時間を計測したところ、現段階では約2.90秒要することがわかった。



図 6: テンプレート画像

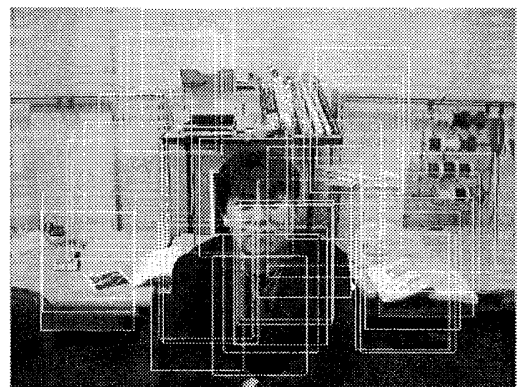


図 7: 初期状態

5.2 実験 2

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 of 数枚の画像の中から顔領域の探索を行った。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。また

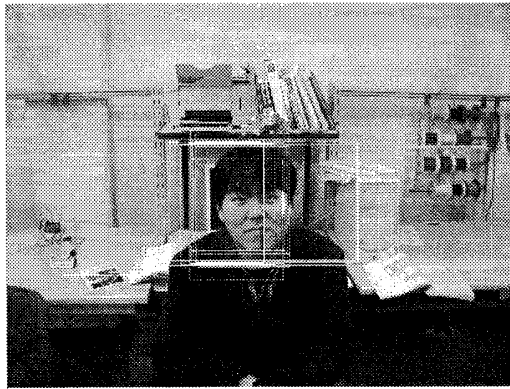


図 8: 第 5 世代

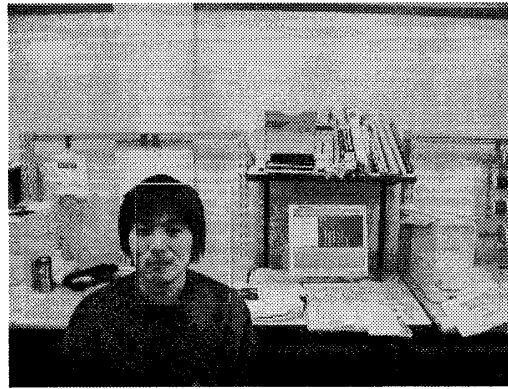


図 10: 追跡の様子 1

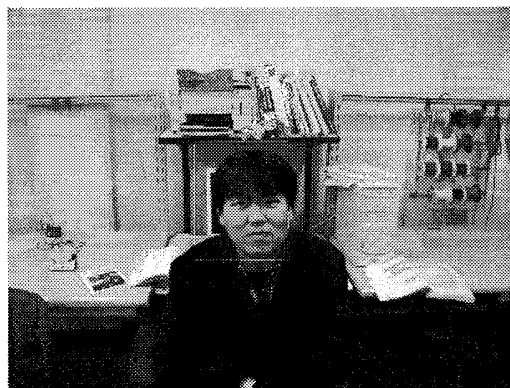


図 9: 最終状態



図 11: 追跡の様子 2

実験 2 のテンプレート画像には、実験 1 で用いたものと同じ平均顔のテンプレート画像を使用した。実験結果を図 10、図 11 に示す。

5.3 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 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考えられる。テンプレートとの比較回数は初期世代では 32 箇所、その後は一世代あたり 16 箇所行っており、15 世代で収束すると考えると 256 回程度比較していることになる。単純に入力画像の端からテンプレートマッチングを行うと 17 万回以上の比較が必要なのに比べて、本手法ではメモリアクセスを少なくし解を得ることが可能である。

実験 2 より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。動画は静止画像が連続して送られてくるものと考え、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。

6. おわりに

本稿では顔領域抽出アルゴリズムを FPGA 上でワンチップ実現し、その回路動作について検討した。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。

参考文献

- [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.