

## hw/sw複合体を用いた Coarse to Fine テンプレートマッチングシステム

横川 昌俊<sup>†</sup> 佐藤季花<sup>†</sup> 須藤一郎<sup>†</sup> 湯野智己<sup>†</sup> 工藤健慈<sup>†</sup>

関根優年<sup>†</sup>

<sup>†</sup> 東京農工大学 工学府

〒184-8588 東京都小金井市中町 2-24-16

E-mail: †{masa,rika,ichiro,yuno,k2x}@sekine-lab.ei.tuat.ac.jp, ††sekinem@cc.tuat.ac.jp

あらまし 大域的な処理を coarse 画像（低解像度画像）で、局所的な処理には fine 画像（高解像度画像）を用いる手法を Coarse to Fine テンプレートマッチングと呼ぶ。本研究では coarse 画像の作成に、Wavelet 変換を使用した。Wavelet 変換によって、入力である fine 画像（640\*480 サイズ）の圧縮を行い、coarse 画像（80\*60 サイズ）を作成する。coarse 画像全域に対して、顔のテンプレート画像を用いてマッチングを行い、顔領域を探索する。次に coarse 画像におけるマッチングで一致した領域を、fine 画像に反映する。そして、その領域を中心に、fine 画像から 80\*60 サイズの矩形領域を切り出す。この切り出した fine 画像で再びマッチングを行うことで、顔領域の詳細な位置決めを行う。Wavelet 変換、テンプレートマッチング、画像切り出しを、FPGA 上にダウンロードした回路を用いることにより、高速処理を実現した。

キーワード 画像認識、テンプレートマッチング、多重解像度解析、FPGA

## Coarse to fine template matching system with the union of hardware and software

Masatoshi YOKOKAWA<sup>†</sup>, Rika SATO<sup>†</sup>, Ichiro SUDO<sup>†</sup>, Tomomi YUNO<sup>†</sup>, Kenji KUDO<sup>†</sup>, and

Masatoshi SEKINE<sup>†</sup>

<sup>†</sup> The Faculty of Technology, Tokyo University of Agriculture and Technology

2-24-16 Naka-chou, Koganei-shi, Tokyo, 184-8588 Japan

E-mail: †{masa,rika,ichiro,yuno,k2x}@sekine-lab.ei.tuat.ac.jp, ††sekinem@cc.tuat.ac.jp

**Abstract** Coarse to Fine template matching is the algorithm which uses the coarse image for global processing and the fine image for local processing. We use the Wavelet transform to make the coarse image in this research. When the fine image that is the input is compressed by the Wavelet transform, and the coarse image is made. We use the coarse image by searching for the face area. This area is the main area of the fine image. We extract the fine image which region size are 80\*60 at the point that is a center of the area. This extracted image is used for the decision of a detailed positioning the face area. We got realtime process using the circuit that downloaded the Wavelet transform, the template matching, and the image extractor on FPGA.

**Key words** image recognition, template match, multiple resolution analysis, FPGA

### 1. はじめに

#### 1.1 研究背景

近年、カメラを入力センサとして利用したシステムの開発が盛んに行われている。これらのシステムに搭載される画像認識手法は、単一レベルの解像度画像を利用して処理を行うものが主流である。精度を優先するようなシステムでは高解像度画像

を使用し、速度が優先されるようなシステムでは低解像度画像が利用される。Coarse to Fine テンプレートマッチングは、探索領域を fine 画像になるに連れて小さくすることで、高解像度化による処理コストの増加は少なくて済む。精度面に関しても、その有効性は関根らによる研究で示されている [1]。カメラ画像の高精細化が進む現在、単一の解像度レベルに対してのみの処理では、精度・速度両面で不十分である。Coarse to Fine テン

プレートマッチングと hw/sw 協調設計により、頑強な認識システムが構築できると考えた。

### 1.2 研究目的

fine 画像は、処理コストがかかるが精度が高いという特徴を持つ。coarse 画像は、精度は落ちるが高速処理が可能という特徴を持つ。fine 画像、coarse 画像双方の利点を活かせる Coarse to Fine テンプレートマッチングを hw/sw 複合体を用いて構築する。coarse 画像を得る為の圧縮処理に伴う負担増加を、ハードウェアへの処理分散により解消し、リアルタイム処理が可能となる。

更に、coarse 画像では個人の固有情報が薄れるため、テンプレート画像の汎用性が得やすいと考えられる。これによる認識精度の向上に関しても検証を行った。

## 2. Wavelet 変換

本研究では、Haar 関数を用いた Wavelet 変換を使用して多重解像度解析を行う。Haar の wavelet 関数を式 1 に、scaling 関数を 2 に示す。[2]

$$\psi_H(x) = \begin{cases} 1 & (0 \leq x < \frac{1}{2}) \\ -1 & (\frac{1}{2} \leq x < 1) \\ 0 & (\text{otherwise}) \end{cases} \quad (1)$$

$$\phi_H(x) = \begin{cases} 1 & (0 \leq x < 1) \\ 0 & (\text{otherwise}) \end{cases} \quad (2)$$

wavelet 関数  $\phi_H(x)$  を High pass Filter, scaling 関数  $\psi_H(x)$  を Low pass Filter として入力画像（原画像）の水平方向・垂直方向の順に 1 次元変換を行うことで  $\frac{1}{4}$  に圧縮された低周波成分 (scaling image) と高周波成分 (wavelet image) が得られる。図 1 には、Wavelet 変換による多重解像度表現を示す。これは、変換後の低周波成分に対して再度変換することにより、得られる信号の表現である。

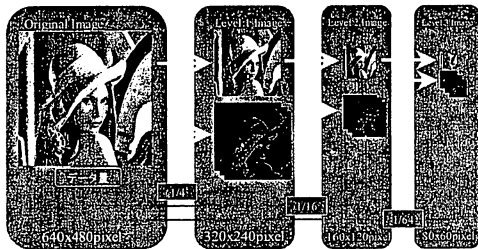


図 1 多重解像度

## 3. Coarse to Fine テンプレートマッチング

### 3.1 マッチング演算

本研究では、差分絶対値と演算を用いてパターンマッチングを行う。式 3 に、 $A(x,y)$  をテンプレート、 $B(x,y)$  をサーチ

ウィンドウとした場合の、差分絶対値と演算の一般式を示す。この場合、テンプレートのサイズは  $m * n$  である。

$$M = \sum_{i=0}^m \sum_{j=0}^n |A_{(i,j)} - B_{(x+i,y+j)}| \quad (3)$$

これを YUV カラー形式へ応用すると、式 4~7 のようになる。ここで、 $O_{\beta}^j$  がテンプレート画像、 $I_{\alpha 1}^j$  がサーチウィンドウ、 $K_c$  が混合係数である。混合係数とは、 $dY$  と  $dU$ 、 $dV$  の寄与度を同等にするために導入した係数で、[1] では  $K_c = 4$  に設定している。

$$dY = |Y_c(O_{\beta}^j) - Y_I(I_{\alpha 1}^j)| \quad (4)$$

$$dU = |U_c(O_{\beta}^j) - U_I(I_{\alpha 1}^j)| \quad (5)$$

$$dV = |V_c(O_{\beta}^j) - V_I(I_{\alpha 1}^j)| \quad (6)$$

$$M = \sum_{i=0}^m \sum_{j=0}^n |dY + K_c * (dU + dV)| \quad (7)$$

### 3.2 Coarse to Fine

多重解像度解析を用いたテンプレートマッチングのことを、Coarse to Fine テンプレートマッチングという。粗い画像でのマッチング結果を、高解像度画像におけるマッチングの対象領域に反映させ、高解像度マッチングでは、局所領域に対してのみマッチングを行うというものである。この処理のメリットは以下の通りである。

- 個人の特徴が薄れるため、テンプレートの汎用性が得やすい
- 演算処理が軽減される
- ノイズの影響を受けにくい

図 2 に、Coarse to Fine テンプレートマッチングの模式図を示す。図 2 では、目を探索している。まず、低解像度画像の状態で顔領域を探し出し、次のレベルの画像でその範囲を更に絞り込む。そして、高解像度画像を用いて、目を特定するというものである。

閾値の実験結果を図 3 に示す。これによると、注視領域の



図 2 多重解像度表現による領域の絞り込み

絞り込みを行う (図 3 右側) を行うことで、各解像度において、正しく目の位置を捉えることに成功している。図 3 左側は、絞り込みが行われていないために、目の探索に失敗している。

これをもとに追加実験すると、Coarse to Fine における解像度レベルの移動は、複数のレベルを飛び越しても同等の効果がある

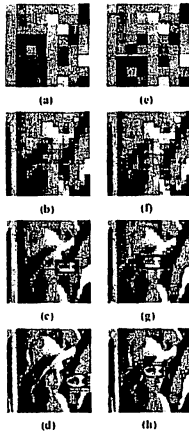


図3 目のテンプレートによる注視領域の探索 [1]

得られることがわかった。Wavelet 変換レベル 3 における探索結果を、原画像に反映させる場合の有効性も確認済みである。

本研究では、低解像度画像を得るために、前項で説明した Haar 関数を基底とした Wavelet 変換を使用した。これを用いた理由は、演算が加算とシフト演算のみで表現できるため、ハード化しやすいということである。

## 4. hw/sw 複合体

### 4.1 hwModule

本研究で想定した画像認識システムを実現するために、本研究室で開発した再構成可能な LSI である FPGA が搭載された PCI デバイス、hwModule (ハードウェアモジュール、図 4) を用いてシステムの構築を行った。図 5 には、hwModule を用いた hw/sw 複合体の実装図を示す。図 5 をみればわかるように hwModule には、hwObject の要素である hwNet (回路) と Data が置かれる。hwObject に関しては、次のセクションで詳しく説明する。

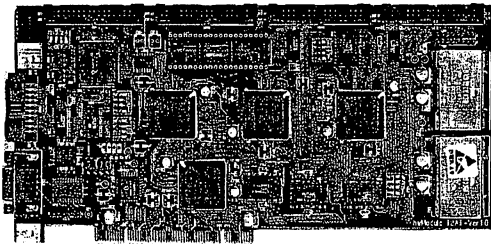


図4 hwModule

hwModule は FPGA、メモリ、マイクロプロセッサ、PCI バスコントローラにより構成される。図 6 には hwModule のブロック図を示す。hwModule は、ボード制御の PCI バスコントローラ用 FPGA を 1 個、仮想回路用 FPGA を 3 個搭載している。更に、処理データを格納するための LM (ローカルメモ

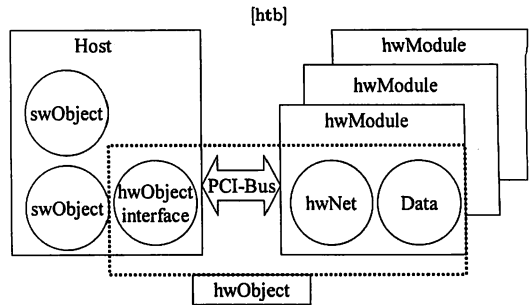


図5 hwObject と swObject の実装

り) を 3 個搭載している。各 FPGA は 2 つの LM バスを持っている。また、隣り合う FPGA 同士で LM を共有しているため、複数の仮想回路が動作する状況でもデータのコピーの必要がないという特徴を持つ。

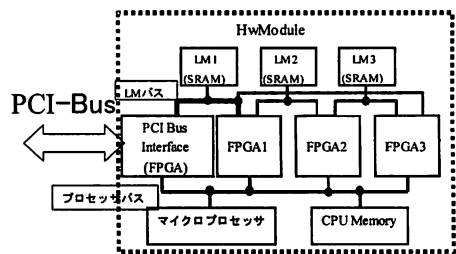


図6 hwModule 内部構成ブロック図

### 4.2 hwObject

hwObject とは、複雑なハードウェアの処理をカプセル化したもので、hwModule 上の FPGA に実装される hwNet により、目的の処理を行う。hwObject に於ける Host 上のメンバー関数は、対象の仮想回路と Host との結合を解決して動作させるための手続きのみである。つまり、従来の C++ 言語によるソフトウェアの記述を用いて、容易にハードウェアが扱えるようになる。hwObject には、ビット処理、信号処理など回路が得意とする機能を割り当てる。[3]

図 7 には、hwObject モデルの概要を示す。hwObject interface は、swObject サイドでは C++ 言語によるメンバー関数呼び出しによるアクセスを行い、hwModule サイドでは PCI バスを通したハードウェアとのデータ転送と制御を行う。

## 5. 回路

### 5.1 主要回路の概略

本研究では、カメラ回路、Wavelet 変換回路、テンプレートマッチング回路、画像切り出し回路 (Extractor 回路) の 4 つの回路を用いた。カメラ回路、Wavelet 変換回路、テンプレートマッチング回路に関しては、共同研究者が開発したもので、その詳細に関しては別途論文にて報告する予定である。本稿では、その概略を記載する。

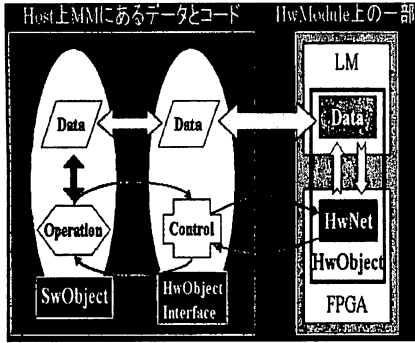


図7 hwObject モデルの概要

### 5.1.1 カメラ回路

CMOS センサ (図8) から来る信号を合成して、640\*480 の画像を生成する。出力画像フォーマットは、8bit グレースケール、24bit/16bit の YUV/RGB カラー形式である。フレーム数はグレースケールで 15[fps]、カラーで 30[fps] である。

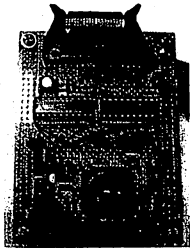


図8 CMOS カメラ

### 5.1.2 Wavelet 変換回路

haar-Wavelet 変換をレベル 1~3 まで行う回路である。パラメータによる記述で、入力画像は任意に設定が可能である。本システムでは、640\*480 が入力画像になるように設定してある。

### 5.1.3 テンプレートマッチング回路

差分絶対値和演算によるテンプレートマッチングを行う回路である。入力画像やテンプレートのサイズは、任意に設定可能であり、本システムでは、入力画像サイズは coarse 画像、fine 画像共に 80\*60、テンプレートサイズは 16\*16 である。

### 5.2 画像切り出し回路 (Extractor 回路)

Extractor 回路は、coarse 画像におけるマッチングの候補座標をローカルメモリから取得し、その座標に対応する領域を fine 画像中から切り出すという機能を持つ回路である。切り出し対象画像、切り出し画像のサイズ共に任意に設定することが可能で、本システムでは、対象となる fine 画像が 640\*480、切り出し画像が 80\*60 に設定してある。

図9に、Extractor のブロック図を示す。候補座標から対象領域の先頭アドレスと終了アドレスを計算し (HeightAddress-Creator)、ローカルメモリから画像を取得し、回路内をスルーさせ再びローカルメモリに書き込むという動作を行う。

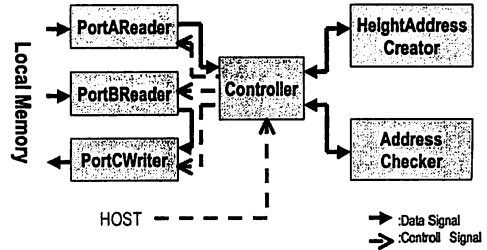


図9 Extractor のブロック図

回路のステートを、図10に示す。1つの画像に対して、座標の取得とアドレス演算 (画像の先頭アドレスを算出) は1回のみ実行される。その後は、アドレスセット、ReadImage、WriteImage が画像サイズ分だけ繰り返されるという動作をする。アドレスセットが実行された直後に、次のラインのアドレスを演算することで、WriteImage の終了と同時に、次ラインの Read が開始されるようにした。

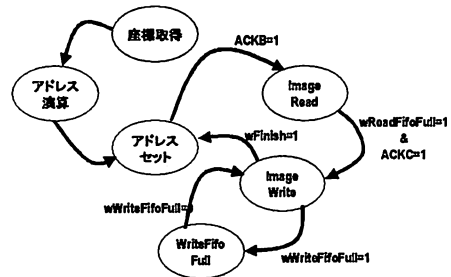


図10 ステータ

表1は作成した Extractor のスペック表である。ユーザー側の回路 (User hwnet) は、MClock (周波数: 66MHz) で動作させる仕様である。

表1 Extractor のスペック (周辺回路は含まない)

スライス数	423 (Max:2352)
最大動作周波数	70.912MHz

## 6. アプリケーション

### 6.1 構成

前節で解説した回路の hwModule 上における配置を、図11に示す。FPGA1にカメラ回路と Wavelet 変換回路、FPGA2にテンプレートマッチング回路、FPGA3に Extractor 回路をダウンロードする。演算に必要なデータはローカルメモリにあるわけだが、その割り当ては図11の上部に記載している。ホスト側には、Wavelet 変換をした 80\*60 サイズの coarse 画像と、fine 画像から切り出された 80\*60 の fine 画像、そしてマッチングの結果の3種類を転送する。

図11の回路を用いたアプリケーションの全体の構成図を、図12に示す。認識に関する基本演算は、全てハードウェアで行う

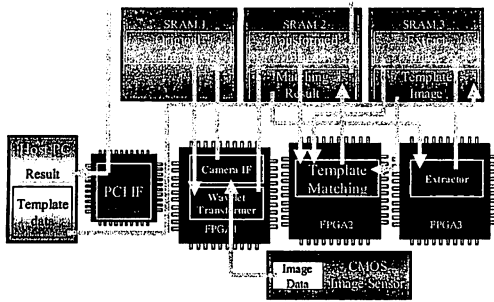


図 11 回路配置とメモリの割り当て

ため、ソフトウェアはコントロールと結果表示のみである。回路で演算した結果は、ローカルメモリを経由して、別回路やソフトウェアに渡される。各 Object は、ハンドルと呼ばれる一種のポインタにより結ばれ、Object が有するリクエストメソッドを用いることで、動作をコントロールすることができる。画像やマッチング結果の取得は、Object に結果取得メソッドを呼び出すメッセージを送ることによって行われる。

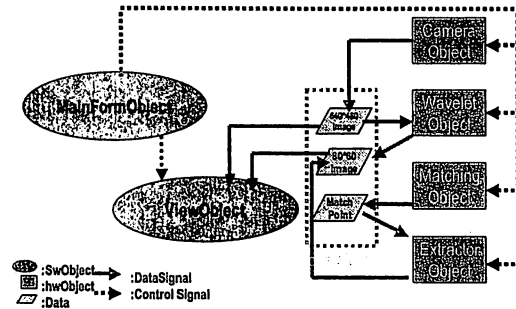


図 12 全体図

## 6.2 シーケンス

図 13 は、アプリケーションの処理シーケンスである。Wavelet 変換により得られた圧縮画像と、Extractor により切り出された画像が同一のマッチング回路に入力されることにより、マッチング回路に変更を加えることなく coarse と fine の画像がマッチングできるような仕様である。Wavelet 変換回路の仕様では、同時にレベル 1~3 までの画像が取得できるが、本システムではレベル 3 のみ、マッチング回路に渡す。動画動作が選択された場合は、座標出力の後に、再度カメラが呼び出され一連の動作がスタートする。

## 6.3 GUI

アプリケーションは、カメラをスタートすると全体が動き出す仕様になっている。(図 14 左上の再生ボタン) マッチングの結果は、coarse 画像と fine 画像それぞれの画像に赤色の矩形を描画することで通知する。画面右下にあるのは、マッチング結果の第 4 候補までの座標と、差分絶対値和である。

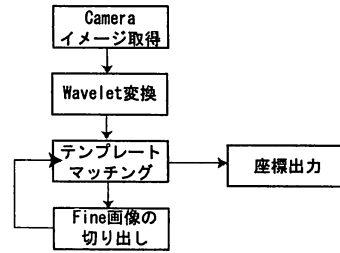


図 13 処理シーケンス

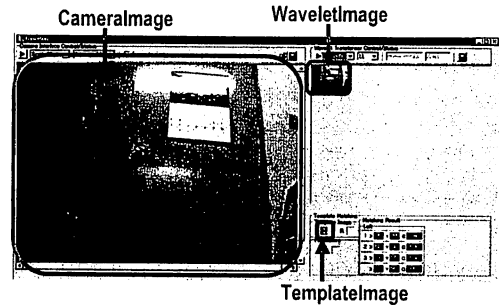


図 14 アプリケーション GUI

## 7. 検証

### 7.1 Extractor

Extractor のシミュレーションの様子を図 15 に示す。座標取得、アドレス計算、LM アクセスが、設計通りに動作していることが確認できた。

しかし、実装段階で正常動作しないことが判明した。LM ア

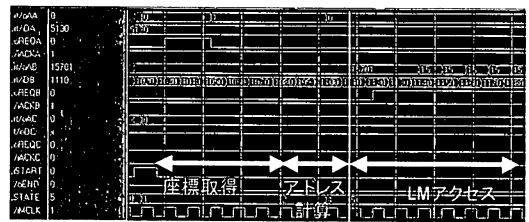


図 15 Extractor のシミュレーション図

クセスが正常に行われなかったために、write 要求に対する ack が立たないため、出力側の fifo がフルになり、「WriteFifoFull」の状態ですテートがストップしていることが、原因であると考えられる。なぜ ack が立たないかは、現在調査中である。

### 7.2 Coarse to Fine テンプレートマッチングシステム

Coarse to Fine テンプレートマッチングシステムを、Extractor 回路部分をソフトウェアによる代理処理によって構築した。CameraObject と WaveletObject の動作は確認されたが、MatchingObject の動作が正常に行われなかったという問題が確認された。テンプレート画像、サーチウィンドウ画像のローカルメモリ内での参照先が正しいか等、現在調査中である。

## 8. む す び

本報告書では、Coarse to Fine テンプレートマッチングシステムの構築手法について述べた。報告書では、各種トラブルからシステムの能力に関する部分の報告を載せることができなかつたが、Extractor 回路、MatchingObject 周辺の問題点を修正し、報告会では追加でその報告をする予定である。

現状の Coarse to Fine テンプレートマッチングには、絞り込みミスが起きた場合の対処法が組み込まれていない。マッチング結果である差分絶対値和に対して、適当な閾値を設けることで、coarse 画像におけるマッチング精度の判断を行う。そして、fine 画像への反映時に探索範囲の伸張を行うことで、絞り込みミスの低減が可能と考える。この機能を本システムに導入し、より頑強なシステム構築を目指す予定である。

### 文 献

- [1] 関根優年, 金丸隆志, 伊藤光, “多重照合による顔領域の特定” 信学論, vol.J86, no9,pp.969-973, Sep 2003
- [2] ウェーブレットビギナーズガイド, 榊原進 著, 東京電機大学出版局
- [3] 工藤健慈, 今中晴記, 志賀裕介, 関根優年, “hw/sw 混載システムにおける hwObject モデルとその制御手法”, 情報科学技術フォーラム一般講演論文集第 1 分冊 FT2002, C-3,pp.193-194, 2002/9.