

並列画像・グラフィック処理プロセッサ構成の検討

廉田 浩[†] 若谷 彰良[‡]

[†]九州大学大学院芸術工学研究院 〒815-8540 福岡市南区塩原 4-9-1

[‡]甲南大学理工学部 〒658-8501 神戸市東灘区岡本 8-9-1

E-mail: [†] kadota@design.kyushu-u.ac.jp [‡] wakatani@konan-u.ac.jp

あらまし 画像の相関計算とグラフィック計算という 2 種類の処理を行う再構成可能な演算ユニットのアレイと内部結合網からなる新しいプロセッサアーキテクチャを提案する。この 2 種類の処理は実時間顔画像認識処理に使用するものである。想定される SoC 中エンジン部分の主な構成は、オンチップフレームメモリ、並列演算ユニット、これら両者を結ぶデータ転送系、および並列演算ユニット間を結ぶ内部結合網である。特にこの内部結合網は、グラフィック処理の結果データを別の演算ユニットへと転送し、そこで相関計算を実行する場合に、非常に重要な役割を果たす。この結合網を含むプロセッサの動作確認のために、結合網の C シミュレータを構築しデータの転送特性と全体の処理性能を概略評価した。この結果、グラフィック計算による修正テンプレートの作成と相関計算によるマッチング処理を 1000 回程度繰り返すアルゴリズムを使った顔認識処理の場合では、512 ユニットの並列構成で 250MHz クロックで動作させると、VGA サイズ 15fps の動画像中の顔認識を実時間で実行できる能力があることがわかった。

キーワード 並列プロセッサ、2次元トーラス網、顔認識、テンプレートマッチング、幾何学計算

A Study of Organization for Parallel Image and Graphic Processors

Hiroshi KADOTA[†] and Akiyoshi WAKATANI[‡]

[†] Faculty of Design, Kyushu University, Shiobaru, Minami-ku, Fukuoka, 815-8540 Japan

[‡] Faculty of Science and Engineering, Konan University, Okamoto, Higashinada-ku Kobe, 658-8501 Japan

E-mail: [†] kadota@design.kyushu-u.ac.jp [‡] wakatani@konan-u.ac.jp,

Abstract A new organization of processing engine is proposed, which can execute at least two types of massive parallel operations, image-correlation and graphics, in a reconfigurable manner. These two functions are necessary to construct a real-time face-recognition engine by using the adaptive template-matching scheme. The system consists of On-chip Frame Memories, Data-transfer Sub-system, Parallel PE's and Inter-PE Network. Especially, the Inter-PE Network plays the major roles in graphic operations and correlation. The operations and the performance of the system is simulated with C-modeled function units. Human faces in a VGA-size motion picture of 15 fps will be identified by the above mentioned engine with 512 PE's operated at 250MHz clock, if the number of the trials for template adaptation / matching is less than 1000.

Keyword Parallel-processor, 2D-Torus-network, Face-recognition, Template-matching, Geometry-operation

1. まえがき

顔画像を使った人の認識（顔検出と個人識別：合わせて顔認識）技術は、セキュリティ、知的

なマシンインターフェース、或いはロボットビジョンといった応用分野で非常に重要である。顔認識には多くのアルゴリズムが提案され[1]、幾つかのものは既に実際のシステムに実装されているが、

まだ使用環境に対する制限が強く、ある程度の悪条件でも実用的な認識特性が得られるという、所謂「処理頑強性」は必ずしも十分ではない。この頑強性不足の主な原因は、顔画像が同一人物のものであっても、その向きや照明条件によって大きく変化するためと考えられる。この対策を含んだ新しいアルゴリズムの研究も進んでいるが、その中で筆者らは、特に3次元顔モデルを使った認識手法に注目している。但し、この3次元モデルを使った認識を完遂するためには、従来からある3次元グラフィック計算と、画像のマッチング計算とを高速に繰り返し実行する必要があり、動画像中の人物の識別を実時間で行うことなどは現在入手できる高速汎用デバイスでは不可能である。

本稿では、上記3次元モデルによる顔認識アルゴリズムの有効性が確認された場合に必要、高速プロセッサ（エンジン）の並列アーキテクチャを提案し、特に並列演算ユニット間の結合網の動作を解析することを通して全体の特性の推定した結果について記述する。

ここで提案するエンジンは、オンチップフレームメモリ、並列再構成型の演算ユニット（PE）アレイ、これらの間のデータ転送を行う多段結合網による転送系、およびPE間のデータ転送を行う内部結合網等から構成される。実行する高速並列処理は、3次元グラフィック処理（等に幾何学計算）と画像処理用の相関計算・フィルタリング処理に大別され、前者は簡略浮動小数点または16ビット以上の固定小数点計算であるのに対して、後者は低精度の固定小数点加減算やMAC演算である。これらはいずれも並列PEで実行するため、PEは各演算に対応できるような再構成型になっている。

次節以降では、上記の3次元グラフィックス処理や2次元の相関計算等およびそれらのデータ配置・移動の概要を述べた後、提案システムの全体構成とPE内の再構成型演算器、およびPE間のデータ転送を実行する2次元トラス型の内部結合網の動作と特性について述べる。最後に、全体の処理性能の予測について議論する。

なお、認識処理の前半の顔検出処理については、筆者らが提案している適合型テンプレートマッチング手法[2][3]を今回提案しているアーキテクチャ（PE間結合網なしの状態）に実装し効率的に実行できることを確認している[4][5][6]。

2. 認識用3Dグラフィック処理と相関計算

特定すべき人物の3次元顔モデルの形状データは、

微小ポリゴン単位で円筒座標系で記述され、位置 $r(\theta, y)$ 、輝度 $Y(\theta, y)$ 、面方位 $dir(\theta, y)$ 等のデータが展開された形でオンチップフレームメモリに格納されている。

2.1. グラフィック処理

図1に示すように、3次元グラフィック処理では、このモデルデータを細かいタイルに分割し、多段結合網によるデータ転送系を介して、それぞれのタイルを担当する並列PEの1個に供給し、照明補正や表情補正を行った後、回転行列を使った幾何学計算および透視変換を行い参照画像データとする。これらのグラフィック計算は、処理結果の画像を人間が直接見るものではないので、それ程高精度である必要はないが、ある程度のダイナミックレンジが必要なため、簡略浮動小数点または16ビット以上の固定小数点積和演算器が必要である。回転等の計算結果は実画像とのマッチングを評価するため、2次元画像の対応する部分（別のタイル系）を担当するPEへ転送する必要がある。PE間の2次元トラス網からなる内部結合網を介してこの転送が行われる。

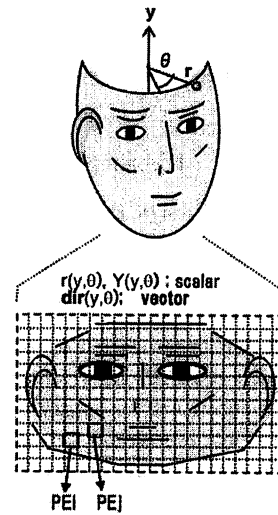
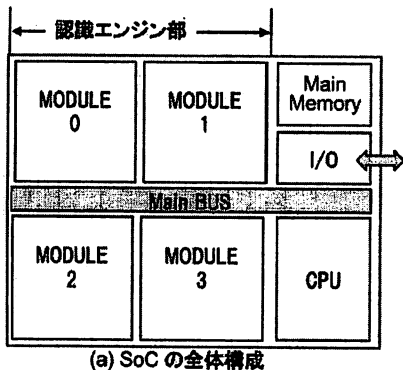


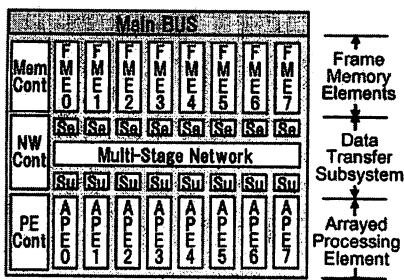
図1 3次元顔モデルとデータの持ち方

2.2. 画像の2次元相関計算

事前に顔の位置を検出してある被検査画像に対して、各検出位置付近で参照画像との2次元の相関計算を行いマッチングを評価する。こちらの計算は単純な固定小数点の加減算および低精度の積和計算である。相関値がある閾値に達した場合は特定すべき人物がその位置に存在することになり、



(a) SoCの全体構成



(b) エンジンモジュールの構成

図2 SoC全体とエンジンモジュールの内部構成

閾値に達しない場合は、適切なフィードバック情報をもとに、グラフィック計算に戻って、顔の回転角、照明方向、表情などを少し変化させた新しい参照データを作成し、そのデータとの間で再び相関計算を繰り返す。一定回数繰り返しても相関値が閾値に達しない場合、その顔検出位置には特定すべき人物が存在しないと判断する。

3. エンジン SoC の構成

全体のシステム(SoC)の構成を図2(a)に示す。この内 CPU は全体動作とデータ入出力や転送の制御を行う。メインメモリはCPUの主記憶であると同時に認識エンジン内の PE に対する共有メモリの役割も果たしている。一方、認識エンジン(プロセッサ)は4個のモジュールで構成されており、このブロックで、前節で述べたグラフィック計算や相関計算を実行する。エンジン部とCPUやメインメモリとはメインバスで結ばれている。

図2(b)に各モジュール内の詳しい構成を示す。この内部は、APE部:8列×16個のプロセッシング要素(PE)、FME:フレームバッファメモリ要素、およびこれらの間のデータ転送を実行する多段結

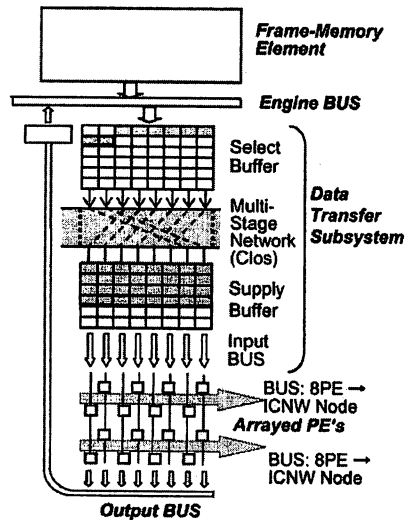


図3 モジュール内の1列中のデータ転送説明図

合網等を中心としたデータ転送系からなる。

この構成は、筆者らが従来から提案している画像処理並列エンジンとほぼ類似なものであるが、一点重要な追加要素がある。それは PE 間をつなぐ内部結合網である。また、PE 内部の演算器等の接続も扱うべき演算の種類にあわせて変更を加えている。フレームバッファ要素は SoC 全体で 32 個あり、合計では大きなフレームバッファとなる。

3.1. 各 PE 列の中のデータ転送

図3に1モジュール内のデータ転送の説明図を示す。CPUやメインメモリからは、エンジンバス(メインバスの一部)を通してデータが入力され、フレームメモリ要素に一旦格納される。ここから、再度このバスを通してデータ転送系に入力され、そこを通して適切な PE へとデータが供給される。

3.1.1. データ転送系

データ転送系は次の3要素からなる。即ち、Se: 選択バッファ、Multi-Stage-Network: 多段結合網、および Su: 供給バッファである。フレームメモリ要素側から入力されたデータはまず、選択バッファに入る。選択バッファでは、上下に隣接した2行分(16×1バイト)のデータ中選択された8バイトのデータを同時に多段結合網に供給することができる。(図3参照)このような"Shift&Rotate"型の選択を行うと、特徴抽出用のフィルタ処理を効率化することができる。多段結合網は Clos 網等の非閉塞型である。また、この多段結合網の出口

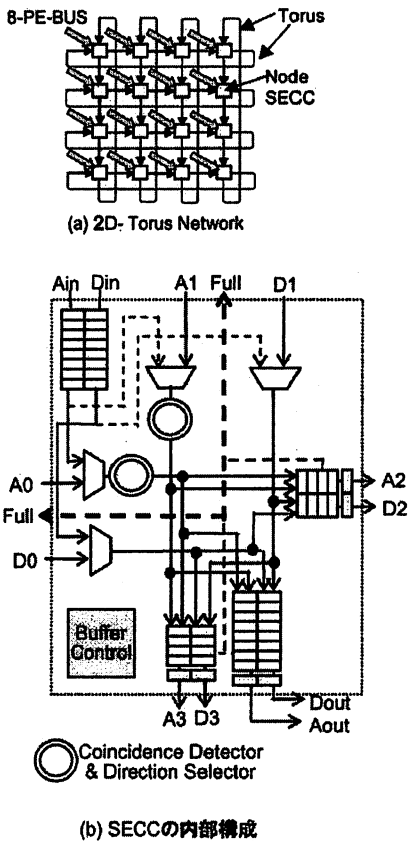


図4 2Dトーラス網と通信制御スイッチ要素

にある供給バッファから各PEまでは、8列x8本の入力バスを介してデータが供給される。

3.1.2. PE間結合網

PE間の内部結合網は、4x4の2次元トーラス形のメッシュである。この結合網は前述の通り、グラフィック処理から相関計算へと移る時のデータ転送（データ再配分）時に有効である。結合網のメッシュの各交点（ノード）間には、8個または4個のPEとの入出力可能なPE-BUSとの接続、上下左右に隣接する4ノードとの単方向性の接続信号線がある。また内部には、上や左隣のノードから転送されたデータの目的地アドレスを判断し、現ノードが目的地の場合ここで出力し、そうでない場合は、更に右や下に転送するような制御論理回路、上記PE-BUSの入力部・出力部各々にあるバッファ(8w)、右と下の隣接するノードへの各々の出力端にあるバッファ(4w)、選択回路等からの通信制御スイッチ要素(SECC)がある。

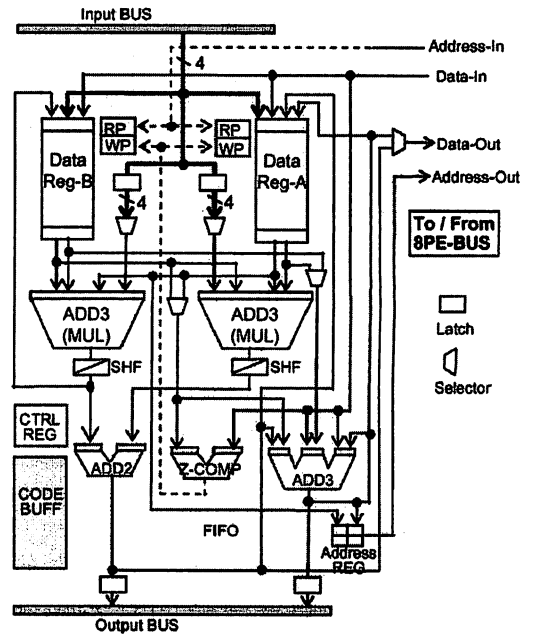


図5 PE内の演算器構成等

各バッファは、その先がフルの場合に一旦データを格納するために使う。また、PE-BUSを介して1ノードに8PEが接続される場合は（通常この接続を仮定している）、4x4の内部結合網1系統で1モジュールをカバーすることになる。

この内部結合網は、各ノードから目的地アドレスとともに入力されるデータを、その目的地まで転送する機能があり、支障がなければ1クロックで隣のノードまでデータ転送が可能である。

また、あるデータの入力ノード(NS)と目的地ノード(ND)との位置関係から、複数の転送ルートが考えられるが、ここでは、転送ルートを確認するための2種類の転送規則を考え、各々の特性を検証し、ハードウェアがさほど複雑にならない範囲で、より動作マージンの大きいものを採用する。
転送規則1: NSとNDとが異なる行のときは必ず縦方向への転送を優先させる。横転送は必ず縦転送のあとで行なう。

転送規則2: NSの行位置をRs、列位置をCsとして、(Rs+Cs)が偶数のときで、NSとNDの行が異なる時は縦転送、逆に(Rs+Cs)が奇数のときで、NSとNDの列が異なる時は横転送。転送の中継ノードでは同一方向を継続する方を優先するが、右でも、下でも転送可能の場合で、どちら一方がフルの場合は、転送方向の変更も行なう。

3.2. PE の内部構成

PE 内部には、演算器として、シフタ・選択器付 3 入力の加減算器が 2 個、単なる 3 入力加減算器 1 個と 2 入力加減算器が 2 個、コンパレータが 1 個存在する。加減算器はいずれも 12 ビット精度である。16 ビット以上の加減算はこれらを 2 個連結して行なう。また、レジスタ・バッファ、FIFO 類として、2 個の 2Bx32 ワードデータレジスタファイル、120 ワードコードバッファ、データの出入口部にラッチまたはレジスタ、および各種の制御レジスタがある。

3 入力のシフタ付加減算器は、通常の加減算の他に、4b x 8b の乗算を実行する場合に使用される。この乗算は基数 4 のブースアルゴリズムを使って実行される。参照画像と被検査画像との相関を計算する場合は、4b x 8b の乗算がそのまま使われるが、グラフィック計算の場合には 8b x 8b の積和が必要なので、この 4b x 8b の加減算器を 2 個使用して実行される。

4. 動作シミュレーションと性能概略予測

今回新たに追加になった PE 間結合網に関して C 言語によるシミュレーションを行い、データ転送特性とエンジン部全体の概略性能を予測した。

4.1. PE 間内部結合網中のデータ転送

PE 間のデータ転送を行う内部結合網では、以下の 3 点が重要である。

- ① どのような入出力パターンに対しても、入力頻度が制限以下ならデッドロックを起こさない
- ② 通常入出力パターンに対して、1 回の実効転送遅延時間のほとんどが入力時間間隔より小さい。
- ③ 通常入出力パターンに対して、入力待ちの行列の最大値が、制限値を超えない。

シミュレーションは、NS と ND との位置関係を乱数で設定し、一定の時間間隔（ピッチ：Pi）で全ノードから一斉にデータを入力する条件で行い、各種の特性値を、各転送規則について測定する。

最も重要な、転送遅延サイクルの分布および最大遅延サイクルと最大入力待ちサイクルを各々、図 6 および図 7 に示す。今回想定しているグラフィック計算（各 PE が 40-50 サイクルで 1 データ生成）から相関計算へ移る場合、結合網の各ノードへの入力ピッチは、8 PE 分を合わせると、5 ないし 6 サイクルとなる。転送規則 1 と 2 では明らかに有意差があり、入力時間ピッチが 4 サイクルで、

前者は明らかに劣化が始まっているので特殊な入力パターンが連続すると急激な遅延時間の増加ないしデッドロックの発生があり得るが（実際 7 サイクル以下のピッチで入力するとデッドロックを起こすパターンが存在する）、後者では劣化が始まる直前の段階である。一方、入力の待ち行列については両者とも問題ない。

以上の結果から、転送規則 2 を選択する。これによって、データの転送時間は、ほぼ計算時間に

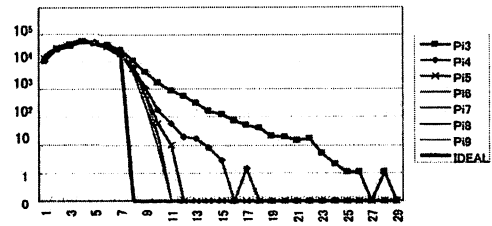


図 6(a) 転送規則 1 の遅延サイクル分布 (Pi:ピッチ)

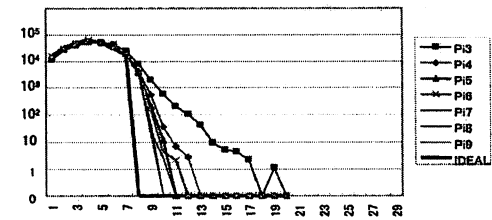


図 6(b) 転送規則 2 の遅延サイクル分布 (Pi:ピッチ)

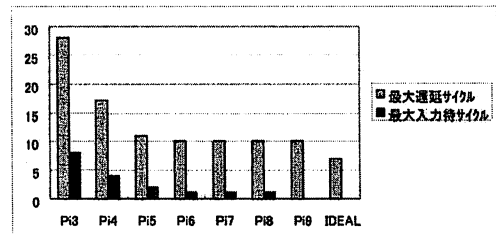


図 7(a) 転送規則 1 の最大遅延・入力待ちサイクル

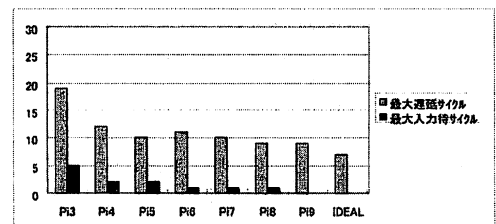


図 7(b) 転送規則 2 の最大遅延・入力待ちサイクル

隠れる形になり、総合の処理時間と無関係になる。

4.2. 認識エンジン部全体の性能予測

顔検出処理の部分も含めて、データ転送の時間がほぼ演算時間に隠れると仮定すると、顔認識全体の処理時間はほぼ演算時間で決まる。各処理の演算量の一覧を表1に、それらの概略実行時間を図8に示す。これは、VGAサイズの画像に対して顔検出を行い、10人が検出され、次に各顔に対して1000回のグラフィック処理-相関計算のイテレーションを実行する場合についての実行時間である。但し、クロック周波数は250MHzで、PE再構成のためのコード入れ替え時間も含まれている。

もしも識別処理が平均250回のイテレーションで1箇所1人物の特定が終了すると仮定すると、この結果は、VGAサイズの動画像中最大10人の人物が検出された場合、特定すべき4人(3次元顔モデルがある人物)と照合し識別するのに要する時間を意味し、約15fpsの動画に対して実時間で人物の認識を実行できることが分かる。

表1 顔認識に必要な各処理と演算量

フレームサイズ (P _F)	307,200	480×640
有効ピクセル数 (肌色領域)	76,800	0.25×P _F
特徴抽出フィルタ演算	22,520,000	150 演算/ピクセル
4-bitMAC (顔検出相関)	4,915,200,000	1000 テンプレート × 256 点
8-bitMAC (グラフィック用)	100,000,000	2000 点 × 1000 テンプレート
検出領域ピクセル数	10×256	10 人
4-bitMAC (識別用)	2,621,440,000	1000 テンプレート × 1024 ピクセル

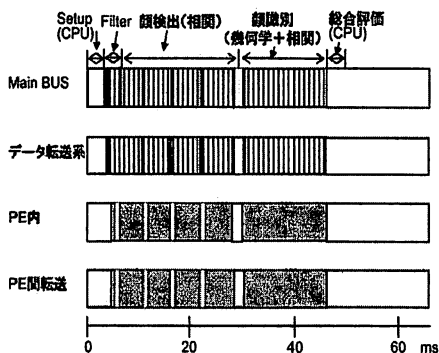


図8 顔認識処理中のデータ転送と処理経過時間

5. まとめ

画像の相関計算とグラフィック(幾何学)計算という2種類の処理を行う再構成可能な演算ユニットのレイアウトと内部結合網からなる新しいプロセッサアーキテクチャを含むSoCの構成を提案した。この2種類の処理は実時間顔画像認識処理に使用するものである。この構成中、特にPE間の内部結合網は、グラフィック(幾何学)処理の結果データを新たな別の演算ユニットへと転送し、そこで相関計算を実行する場合に、非常に重要な役割を果たす。この結合網を含むプロセッサの動作確認のために、結合網のCシミュレータを構築しデータの転送特性と全体の処理性能を概略評価した。この結果、幾何学計算による修正テンプレートの作成とマッチング処理を1000回程度繰り返すアルゴリズムを使った顔認識処理の場合では、512ユニットの並列構成で250MHzクロックで動作させると、VGAサイズ15fpsの動画像中の顔認識を実時間で実行できる能力があることがわかった。

文 献

- [1] M-H.Yang, D. J. Kriegman and N. Ahuja, "Detecting Faces in Images: A Survey," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.24, No.1, pp34-58, Jan. 2002.
- [2] 笠木伸悟, 若谷彰良, 廉田浩, "セグメント自己適合型テンプレートによる顔検出手法の基礎検討" 信学技報, Vol.105 No.98 pp.31-36, PRMU2005-98, 2005年10月
- [3] 廉田浩, 若谷彰良, "セグメント自己適合型テンプレートによる顔検出手法の検討(2)" 信学技報, Vol.106 PRMU2006, 2006年9月
- [4] H.Kadota, Y.Hori, A.Wakatani, "A New Reconfigurable Architecture with Smart Data-Transfer Subsystems for the Intelligent Image Processing," Proc. FPT04, pp.429-432, Dec., 2004
- [5] 若谷彰良, 廉田浩, "並列再構成画像処理システム用データ供給系" 信学技法, Vol.105, No.43, pp.25-28, RECONF2005-19, 2005年5月
- [6] 廉田浩, 笠木伸悟, 若谷彰良, "オブジェクト認識処理用高速リコンフィギャラブルシステムの検討", 信学技報, vol. 105, No. 451, pp. 73-78, Nov. 2005, RECONF2005, 2005年11月