

# 浮動小数点 DSP による高並列アレイプロセッサシステム†

小 畑 正 貴\*\*

本論文は高並列計算機 SPA の設計と実装, 基本動作, 性能評価について述べている。SPA は浮動小数点 DSP を用いた同期動作型の 1 次元アレイプロセッサで, 画像処理や画像生成, 数値シミュレーションなどへの応用における価格性能比の優れたスーパーコンピューティングを目的とする。20 MFLOPS の浮動小数点 DSP を 256 台用い, 接続形態は双方向リングと放送機能を持つ 2 本のバスとの組合せで, プロセッサ間データ転送は 6 M ワード/秒の DMA 転送である。応用例として近接画素を用いた画像フィルタ, 2 次元 FFT, 連立方程式の計算 (掃き出し法) という種類の異なる 3 問題を実行して, 計算速度やデータ転送の割合を測定した。

## 1. はじめに

大規模高速数値計算に対する要求の拡大と, マイクロプロセッサレベルでの数値計算能力の向上を背景として, 高並列型スーパーコンピュータの開発が盛んになっている<sup>1)~4)</sup>。これらは, 数十 MFLOPS の倍精度浮動小数点計算能力を持つマイクロプロセッサを要素プロセッサに用い, 100 台以上のプロセッサをハイパキューブや多次元アレイなどのネットワークで接続した MIMD 型並列計算機の形態をとっている。このような高並列型スーパーコンピュータは, 従来のベクトル型スーパーコンピュータに対して価格性能比が優れている点と, 超並列化によって絶対性能の点でも優位に立つ可能性がある点で注目されている。現状では, 並列ソフトウェア開発の難しさから主に特殊用途への利用に留まっているが, 今後の並列ソフトウェア研究により有効性が広く発揮できるものと期待できる。

一方, 数値計算目的のプロセッサの一つに, 音響機器などにおけるデジタル信号処理を主な用途として発達してきたデジタルシグナルプロセッサ (DSP) がある。最近では, 精度・速度ともに飛躍的に向上してきており, 単精度ではあるが数十 MFLOPS の浮動小数点計算能力を持つものが登場している。また, ハードウェア・ソフトウェアともに汎用性が増し, 使いやすくなってきている。このような浮動小数点 DSP は高並列型スーパーコンピュータの要素のプロセッサとして用いることが可能である。現状では精度の点で数値計算一般には対応できないが, 用途によっては価格性能比の優れた高速数値計算システムが実現できる。

並列計算機 SPA (Signal Processor Array)<sup>7)</sup> は, 上記の点を考慮して試作されたマルチ DSP システムであり, 同期動作型アレイプロセッサとして位置づけられる。汎用 DSP による数百台規模のマルチプロセッサはほかにも報告されているが<sup>5), 6)</sup>, 現在のところいずれもニューロネット・シミュレーションに応用が絞られている。SPA では画像処理, 画像生成, 数値シミュレーション分野での高速計算を目的とする。SPA は以前に開発したアレイプロセッサシステム (BC プロセッサアレイ)<sup>8)</sup> と基本構造は同じであるが, プロセッサを浮動小数点 DSP にし, プロセッサ間転送を DMA にしたことで性能は飛躍的に向上した。

以下, SPA のハードウェア構成と実装, データ転送と同期, 応用について述べる。応用では, 画像フィルタ, 2 次元 FFT, 連立方程式における掃き出し法というタイプの違う三つの問題を取り上げ計算方法と性能測定結果を示す。

## 2. ハードウェア

### 2.1 全体構成

システムの概要を表 1 に, ハードウェアの全体構成を図 1 に示す。ホストコンピュータにはパーソナルコンピュータを用い, SPA はホストの入出力機器として接続される。また, 画像の入出力用にカラービデオカメラと CRT, VTR を接続している。

接続形態は実装の容易な 1 次元アレイとし, 1 次元構造向きのアルゴリズムを中心に考えていくことにした。非同期並列処理や動的負荷分散を行う MIMD マシンの場合は任意のプロセッサ間での効率よい非同期通信が必要である。そこで, 数十台程度のプロセッサ数でも多次元構造を採ってプロセッサ間距離を短縮し, 通信の平均多重度を上げる必要がある。一方 SPA では大規模行列を静的に分割し, 隣接通信とバスによ

† A Highly Parallel Array Processor System with Floating-point Digital Signal Processors by MASAKI KOHATA (Department of Electronic Engineering, Faculty of Engineering, Okayama University of Science).

\*\* 岡山理科大学工学部電子工学科

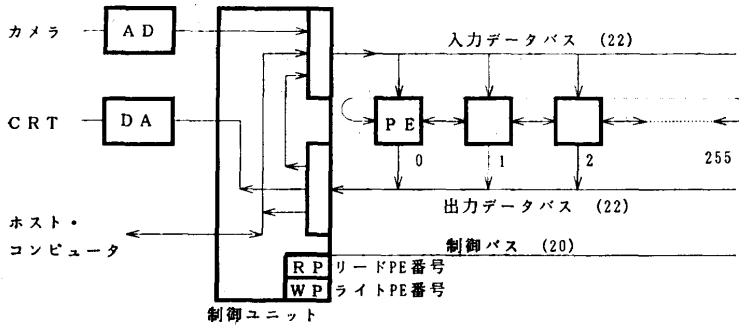


図 1 SPA システム構成  
Fig. 1 System configuration of SPA.

表 1 SPA システム概要  
Table 1 Outline of SPA system.

PE	
データ形式	22 ビット浮動小数点 (16+6)
サイクルタイム	100 ns
計算速度	20 MFlops (積和計算)
メモリ容量	32 KB
全体	
プロセッサ数	256
ピーク性能	5.12 GFlops
メモリ容量	8 MB
接続	バス 2 本, リング
PE 間転送速度	6 MW/sec (× 通信多重度)
ホスト	PC 9801 RA
その他	画像用 AD/DA 変換器

る放送により同期的に計算を進める。この場合、全プロセッサは競合無しに通信を行うので台数が増えても通信効率が落ちることはない。

SPA では、転送の高速化や放送によるバスの効率的利用を計り、データ転送能力を上げるようにしている。PE 間接続には 2 本のバス (それぞれ単方向の入力バスと出力バス) 接続と双方向のリング接続とを組み合わせている。すべての転送はデータ幅 22 ビットの並列転送で、転送速度は (6 M ワード/秒) × 多重度である。PE 間転送時のメモリアドレスは各 PE 内のアドレスレジスタから供給するので、接続はデータバスのみとなっている。

### 2.2 PE (プロセッシングエレメント)

PE の構成を図 2 に示す。DSP のデータ形式は 22 (16+6) ビットの浮動小数点で、サイクルタイムは 100 ns である。DSP の主な構成要素は乗算器と ALU, 512 ワードの内部データメモリであり、積和計算を 1 サイクルで行う。動作クロックは各 PE で独立に発生する。ローカルメモリはプログラムとデータ

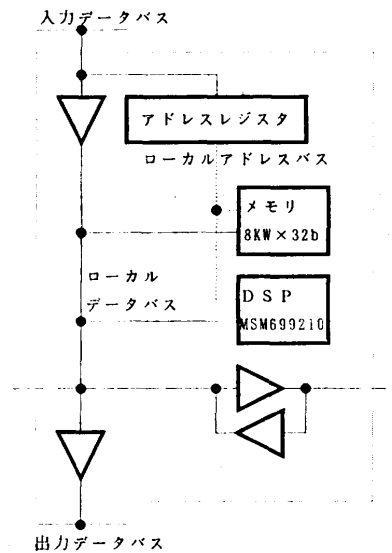


図 2 PE の構成  
Fig. 2 Processing element of SPA.

共用で、アドレスレジスタは DMA 転送時のメモリアドレスを生成する。

PE-PE 間およびホスト-PE 間のデータ転送は DMA により行う。DMA の間 DSP は停止し、ローカルデータバスは入力バスまたは出力バス、隣接 PE のローカルデータバスのいずれかに接続される。ローカルアドレスバスはアドレスレジスタに接続される。

アドレスレジスタの値は入力バスを通してホストから初期設定され、1 回のメモリアクセスごとにインクリメントされていく。n 個のデータを転送する場合は、転送元 PE と転送先 PE のアドレスレジスタに先頭アドレスをそれぞれセットし、両 PE のデータバスを接続して n 個のパルスを発生させればよい。DMA 時の転送タイミングは、制御ユニット内にあるパルス発生器 (6 MHz) に同期する。

### 2.3 制御ユニット

制御ユニットは、PE の動作の制御、バスの接続切り換え、PE 間での DMA によるデータ転送の制御などを行う。前述のようにデータ転送時には、制御ユニットまたはホストコンピュータが直接 PE 内のメモリをアクセスする。入力バスにはマルチプレクサによって、AD 変換器、ホスト、出力バスのいずれかが接続される。出力バスを選択した場合、出力バス上のデータが直接入力バスに送られることになる (バイパス)。

制御ユニット内には、PE 番号を保持する二つのレジスタ (RP と WP) がある。RP はデータを読み出

す PE を指定し, WP は書き込む PE を指定する. これらのレジスタにはオートインクリメントモードがあり, このモードにするとデータ転送ごとに PE が切り替わるので, 全 PE へのデータの分配を効率よく行うことができる.

また, 画像データの入出力用に, ビデオ帯域の AD/DA 変換器を用意している. ビデオカメラからの画像は AD 変換され, 直接 PE 内のメモリに転送される. また, CRT への表示データは直接 PE 内メモリから読み出され DA 変換され表示される.  $256 \times 256$  ドットの画素数で, カラー 21 ビット (各色 7 ビット) の画像のリアルタイム入出力が可能である.

## 2.4 実 装

48 cm  $\times$  47 cm の 4 層基板を用いて 8 PE を 1 ボードに実装した. バックプレーンで 32 枚の SPA ボードと制御ユニットを接続し, 制御ユニットをホストと接続する. 本体の大きさは約 50 cm 立方で, パーソナルコンピュータの横に置かれる (図 3). 基板内では 8 PE を横一列に配置し, 奇数枚目と偶数枚目で PE 番号を逆向きに付けることにより, 基板間での隣接

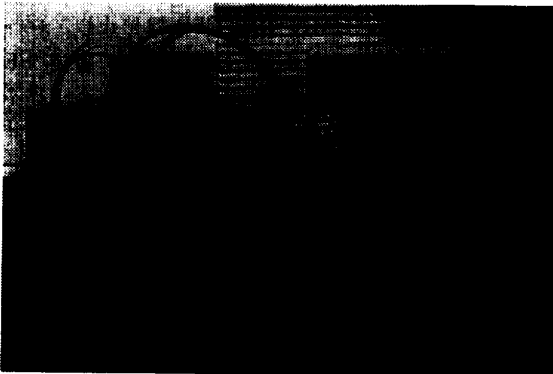


図 3 システム全景

Fig. 3 Whole view of system.

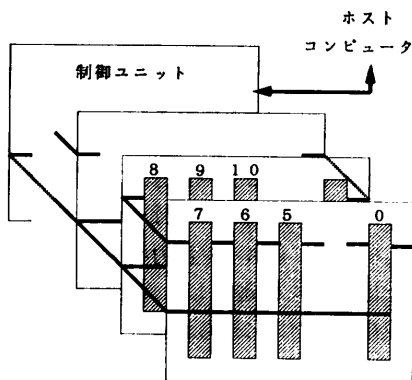


図 4 ハードウェアの実装

Fig. 4 Implementation of the hardware.

PE 接続とバス接続における配線長を抑えている (図 4). 隣接 PE 接続が双方向性であることから, 方向制御信号を 1 枚おきに逆転することで, 基板の向きは変えずに実現できる.

## 3. データ転送と同期

### 3.1 転送モード

以下に, 基本的な 5 形態の PE 間データ転送方法を示す. 前者 2 形態はホスト・AD・DA と PE 間でのデータ転送で, 後者 3 形態は PE-PE 間のデータ転送である. データ転送は常にホストの制御により行われ, PE が自分からデータを出すことはない. したがって, PE 間データ転送のプログラムはホスト上に存在して, PE 上には存在しない.

#### (a) ダイレクトリード/ライト (図 5 a)

RP で指定した PE からの読み出しと WP で指定した PE への書き込みを行う. PE が別であれば読み書きを同時に行うことができ, 実時間画像処理で前画面の出力と次画面の入力とを並行して行う場合に有効である. この場合, 通信多重度は 2 となる. ここで,

src: 転送元メモリアドレス  
dst: 転送先メモリアドレス  
srcPE: 転送元 PE 番号  
dstPE: 転送先 PE 番号  
size: 転送データ数

とすると, 転送手順は以下のようになる.

#### [転送手順]

- ① RP に srcPE, WP に dstPE, を設定.
- ② srcPE と dstPE のアドレスレジスタに, それぞれ src と dst を設定.
- ③ ダイレクト指定. (srcPE バス-出力バス, dstPE バス-入力バスを接続, srcPE メモリを read, dstPE メモリを write, に固定)
- ④ size 個の転送パルスにより読み書きを同時に行う. ホストの通信では, リード/ライトを分離する.

#### (b) ブロードキャスト (図 5 b)

入力バスによるデータ書き込み時, 同一データを全 PE に同時転送する (放送). 全 PE へのプログラムのロードもこれで行われる. アドレスレジスタの設定にも有効で, 全 PE に同一の先頭アドレスが与えられる. PE 数  $n$  に対して, 通信多重度は  $n$  となる.

#### [転送手順]

- ① 全 PE のアドレスレジスタに dst を設定.
- ② ブロードキャスト指定. (全 PE のバスは入力バス

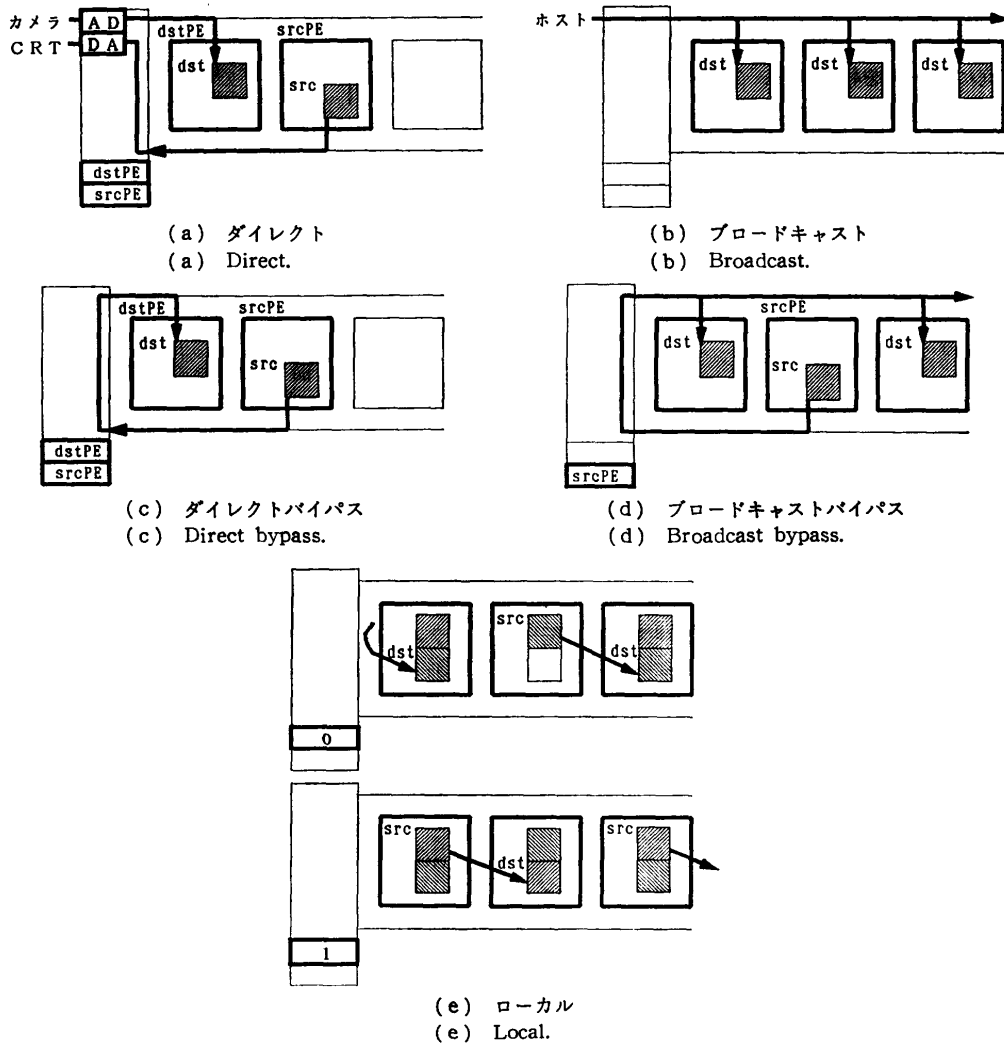


図 5 データ転送モード  
Fig. 5 Data transfer modes.

- に接続, 全メモリは write に固定.)
- ③ホストの src 番地から size の数だけデータを出力.  
(c) バイパスダイレクト (図 5 c)  
出力バスを入力バスにバイパスし, 転送元 PE からの読み出しデータを直接転送先 PE に書き込む. 多重度は 1 である.
- [転送手順]**
- ①RP に srcPE を, WP に dstPE を設定.
  - ②srcPE と dstPE のアドレスレジスタに, それぞれ src と dst を設定.
  - ③ダイレクトバイパス指定. (srcPE バス-出力バス-入力バス-dstPE バスを接続. srcPE メモリを read, dstPE メモリを write に固定.)
  - ④size 個の転送パルスを発生.

- (d) バイパスブロードキャスト (図 5 d)  
指定 PE から読み出したデータを入力バスにバイパスし, 自分以外の全 PE に放送する. src=dst とすれば転送後の全 PE の同一アドレスには同一データが保持されることになる. 多重度は  $n-1$ .
- [転送手順]**
- ①RP に srcPE を設定.
  - ②全 PE のアドレスレジスタに src を設定.
  - ③ブロードキャストバイパス指定. (srcPE バス-出力バス-入力バス-他の全 PE のバスを接続. srcPE メモリは read, 他の全メモリは write に固定.)
  - ④size 個の転送パルスを発生.
- (e) ローカル (図 5 e)  
全 PE は PE 番号によって偶数 PE グループと奇

数 PE グループに分かれ、偶数 PE から隣合う奇数 PE (または奇数 PE から偶数 PE) に対して右方向または左方向にデータを転送する。アレイの両端は接続されリング状になっているので、アレイ内データが回転することになる。同時に全 PE 数の半数の転送が行われることになり、アレイ全体のデータを1方向に回転するには2回の転送サイクルが必要となる。したがって、多重度は  $n/2$  となる。ローカル転送は全 PE で一斉に行われるので1カ所だけでの転送はできない。

#### 【実行手順】

- ① WP に 0 を設定。(奇数グループが転送元に、偶数グループが転送先になる。)
- ② 転送元グループのアドレスレジスタには `src` を、転送先グループのアドレスレジスタには `dst` を設定。(アドレスレジスタの設定には、偶数 PE のみまたは奇数 PE のみに対する放送機能がある。)
- ③ 右(左)ローカル指定。(転送元 PE のバスを隣の転送先 PE のバスに接続。転送元グループのメモリは `read`、転送先グループのメモリは `write` に固定。)
- ④ `size` 個の転送パルスを発生。
- ⑤ WP を 1 にしてグループを替え、②③④ を繰り返す。

#### 3.2 プログラミングスタイルと同期

SPA の各 PE はローカルメモリ上のプログラムにより非同期独立動作するが、PE 間通信はホストにより一斉同期をとって行われる。SPA の並列処理形態はマクロ SIMD 型といえることができる。ホストはまず全 PE に同一プログラムをブロードキャストし、つぎに比較的大きなプログラムブロックを単位として一斉同期をとりながら処理を進めていく。

各 PE は PE 間同期のために2ビットの出力フラグと2ビットの入力フラグを持つ。全 PE の出す2ビットの出力は制御バス上でそれぞれ論理積と論理和がとられ、ホストによってテストされる。またホストの出す2ビットの信号は入力フラグとして全 PE にテストされる。論理積のビットによって全 PE 間での同期がとられる。

ホストと SPA の同期動作を図6に示す。PE には同期のマクロ (`Sync`) を、ホストには同期と再開二つの関数 (`Sync()`, `Restart()`) を用意した。各 PE は1ブロックの実行が終わると出力フラグをたて、その後入力フラグをテストしながら待ち状態に入る。ホストは論理積ビットを見ており、これがたった時点で全

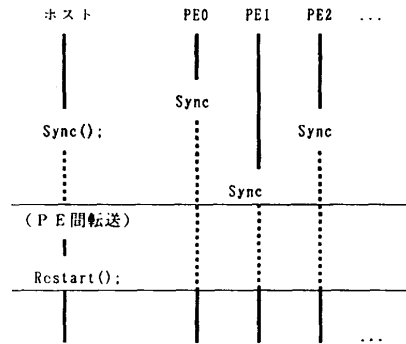


図6 プロセッサ間同期

Fig. 6 Synchronization between processors.

PE での実行の終了を知る。ホストはさらに必要ならばデータ転送を行い、全 PE の入力フラグに信号を送ることにより一斉に再起動をかける。

以上のように、同期とデータ転送はすべてホストの制御によって行われる。SPA では、PE 上の計算プログラムとホスト上の同期・転送プログラムとが協調して動くことにより、全体での計算が進められる。

## 4. 応用プログラムと計算時間

### 4.1 近接画素による画像フィルタ

ある画素を中心とした  $m \times m$  の近接画素と  $m \times m$  の係数行列の要素との積和計算を、全画素に対して行う。画像サイズは  $256 \times 256$  または  $512 \times 512$  で、それぞれ1ラインまたは連続する2ラインを1 PE に割り当てる。係数行列は  $3 \times 3$  または  $5 \times 5$  で実行時間を測定した。

画像サイズが  $256 \times 256$  で、係数行列が  $3 \times 3$  の場合では、あるラインの計算にはその上下1ラインずつの画素データが必要となる。この場合のプログラムを図7に示す。全 PE が計算に必要なデータを得るには1ライン256画素の右回転と左回転を行えば良い。ホスト側のプログラムは C、SPA 側のプログラムはアセンブラによって記述されている。

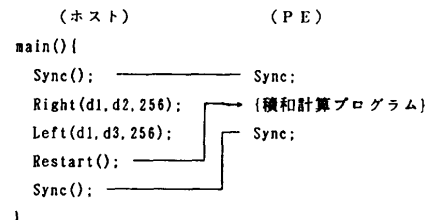


図7 フィルタ処理のプログラム

Fig. 7 Image filtering program for SPA.

実行時間 (PE 間データ転送時間を含む) を表 2 に示す。下段は乗算と加算の全回数を実行時間で割ったもので、システムの平均性能 (GFLOPS 値) を示す。画像サイズと係数行列のサイズによって、データ転送と計算との比率、および DSP 内でのパイプライン効率が変わってくるため平均性能に違いがでている。データ転送時間の実行時間に対する割合は、(画像サイズ, 係数行列) が (256×256, 3×3) の場合が最大で 20%, (512×512, 5×5) の場合が最小で 7% である。

(256×256, 3×3) の場合の実行時間 1.03 ms のうち、計算時間と転送時間はそれぞれ 0.82 ms と 0.21 ms である。同じ問題を 1 PE で処理する場合は、計算時間が 0.82×256=210 ms となり、転送は必要なくなる。1 PE に対する 256 PE の速度比は 210/1.03=204 となり、この時の効率は 204/256=80% と計算できる。

この例は負荷が完全に均質である点では並列処理に有利であるが、通信に対して計算処理が軽いので台数効果が得られにくいという不利な点もある。ここで 80% の効率を得られたことは、画像処理に限らず近接データによって計算を進める問題には有効性があることを示している。

動画像の連続処理における画像の入出力と計算の関係を図 8 に示す。3 組のデータ領域を用意することにより計算・前画像の出力・次画像の入力をオーバーラップさせる。この場合、出力は 2 画面時間ずつ遅れることになる。入出力はダイレクトモードで行われる。左

右の PE へのローカル転送はビデオ信号の同期信号中に終了するので画像入出力に影響しない。1 画面の計算に要する時間が 1 画面時間 (NTSC 方式で 1/30 秒) 以内であれば画像入出力も含めて実時間連続処理が可能である。例えば、ビデオカメラで撮影している動画像 1 枚 1 枚に上記処理を行いながら処理画像を VTR に録画していくことができる。

4.2 2次元 FFT

$n \times n$  点の 2 次元 FFT は各行への  $n$  点 FFT を全行に対して行った後、各列への  $n$  点 FFT を全列に対して行うことで完了する。SPA では最初 1 PE に 1 行を持たせるので行方向の FFT は各 PE で独立に実行できる。次に列方向の FFT を行うためにデータの転置操作を行う。基本となる 1 行の転置は 1 PE の持つ連続データを読み出しながら全 PE に順に転送することである。これはダイレクトバイパス転送を使い、読み出し PE 番号を固定して書き込み PE 番号をオートインクリメントにすることで実行できる。転置の後、各 PE が再び独立に FFT を行えば計算は終了する。計算結果は転置された形になるので、出力時には列方向を先に読み出す。

256×256 点の複素 FFT の計算時間を表 3 に示す。1/30 秒で処理できているが処理時間のうち 78% がデータの転置に費やされている。ダイレクトバイパス転送において 1 度に 1 データしか転送を行えないことがボトルネックになっている。効率を上げるためにはバスの高速化と多重化が必要である。多重化では、各 PE のメモリを  $m$  バンクに分割し、バスを双方向化して  $m$  本用意すれば多重度を  $m$  にできる。例えば  $m=4$  とするとデータ転送時間は 1/4 になり、全計算時間は 14 ms 程度に抑えられる。

4.3 掃き出し法による連立方程式の計算

掃き出し法の計算手順を図 9 に示す。ここでは係数行列のサイズを 256 以下とし、1 行を 1 PE に割り当てて並列計算を行う。ピボット選択は行っていない。ピボット行 (図の第  $k$  行) をブロードキャストバイパス転送で全 PE に転送すれば、全 PE は並行して消去計算を実行できる。この操作を行数だけ繰り返せ

表 2 計算時間と計算速度  
Table 2 Execution time and computing speed.

画像 係数	256×256	512×512
3×3	1.03 (1.15)	3.45 (1.37)
5×5	1.94 (1.69)	6.40 (2.05)

単位 ms  
(GFlops)

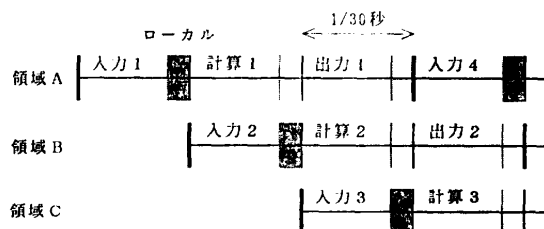


図 8 動画像の連続処理

Fig. 8 Continuous processing of video signal.

表 3 2次元 FFT の実行時間 (256×256)  
Table 3 The execution time of two-dimensional FFT.

	全計算	計算	通信
時間 (ms)	33.0	7.2	25.8
比率 (%)	100	22	78

```

float a[N][N+1]; /*係数行列*/
for(k=0;k<256;k++){
  p=1/a[k][k];
  for(j=k+1;j<=N;j++){
    a[k][j]=a[k][j]*p;
    for(i=0;i<N;i++){
      if(i==k) continue;
      for(j=k+1;j<=N;j++){
        a[i][j]=a[i][j]-a[i][k]*a[k][j];
      }
    }
  }
}

```

図 9 掃き出し法のアルゴリズム

Fig. 9 The algorithm of Gauss-Jordan method.

表 4 掃き出し法の実行時間 (256 変数)

Table 4 The execution time of Gauss-Jordan method.

	全計算	計算	通信
時間 (ms)	49.4	42.8	6.6
比率 (%)	100	87	13

ばよい。上三角部分の消去も行われるので後退代入が不要になる。

係数行列として  $256 \times 256$  の密行列を用いた場合の計算時間を表 4 に示す。ブロードキャスト転送を使うことで転送時間は全体の 13% となっている。図 9 での計算回数は行列のサイズ  $N$  に対して、逆数計算  $N$  回、積算  $N^2(N+1)/2$  回、減算  $N(N^2-1)/2$  回となる。SPA において  $N=256$  での全計算回数を計算時間 (49.4 ms) で割った平均性能は 340 MFLOPS となる。

## 5. おわりに

浮動小数点 DSP を 256 台用いたマルチプロセッサシステム SPA について述べた。単位プロセッサに DSP を使い 1 次元接続にしたことは、応用に制限を受けるといふ不利がある反面、価格性能比に優れていて実装が容易であるという有利さもある。掃き出し法の実行においても高い実効性能が得られた点で信号処理以外の応用も期待できる。

転送モードに関しては、画像フィルタではダイレクトと転送とローカル転送、FFT ではダイレクトバイパス転送、掃き出し法ではブロードキャストバイパス転送が利用された。ただし、FFT では単一バスでのデータ転送のボトルネックが問題であり、バスの多重化を考える必要がある。

今後の課題としてはソフトウェア開発環境の整備、特にホストと PE のプログラムを一体化して記述できるプログラミング言語の設計と処理系の実現があげ

られる。また、応用として、

- コンピュータグラフィックス
- 数値計算 (シミュレーション)

などの問題を取り上げ、SPA 上に実現していく予定である。特に連立方程式の計算ではピボット選択の導入や PE 数より大きい行列への対応、反復法の実装を行ってきたい。

**謝辞** 日頃から有益な示唆をいただき、本学工学部宮垣嘉也教授、神戸大学工学部金田悠紀夫教授に感謝します。また、SPA の製作と性能測定に協力された本学大学院生の野田良作氏 (現三菱電機(株))、伊藤拓氏 (在学中) に感謝します。

## 参考文献

- 1) 石畑ほか: 高並列計算機 CAP-II の構成とメモリシステム, 情報処理学会研究会報告, 90-ARC-83, pp. 217-222 (1990).
- 2) 谷川ほか: 並列計算機 ADENA-ADENA システムの実現, 電子情報通信学会研究報告, CPSY 88-11 (1988).
- 3) 白川ほか: 並列計算機 QCDPAX とその並列処理プログラムディバグ, 並列処理シンポジウム JSPP '90 論文集, pp. 217-224 (1990).
- 4) 中田ほか: 並列回路シミュレーションマシン Cenju, 並列処理シンポジウム JSPP '90 論文集, pp. 353-360 (1990).
- 5) 平岩ほか: ニューラルネットのための RISC プロセッサアレイ, 並列処理シンポジウム JSPP '90 論文集, pp. 385-392 (1990).
- 6) 吉沢ほか: 高並列リングアーキテクチャ, 情報処理学会研究会報告, 90-ARC-83, pp. 67-71 (1990).
- 7) 小畑: シグナルプロセッサアレイ-SPA-, 情報処理学会研究会報告, 89-ARC-79, pp. 101-106 (1989).
- 8) 小畑ほか: 試作 BC プロセッサアレイとその評価, 情報処理学会論文誌, Vol. 27, No. 5, pp. 909-915 (1986).

(平成 2 年 10 月 11 日受付)

(平成 3 年 6 月 13 日採録)

### 小畑 正貴 (正会員)

1957 年生。1980 年神戸大学工学部電子工学科卒業。1985 年同大学院自然科学研究科博士課程修了。学術博士。1984 年岡山理科大学理学部助手, 1990 年同大学工学部助教授, 現在に至る。計算機アーキテクチャ, 並列処理の研究に従事。電子情報通信学会会員。