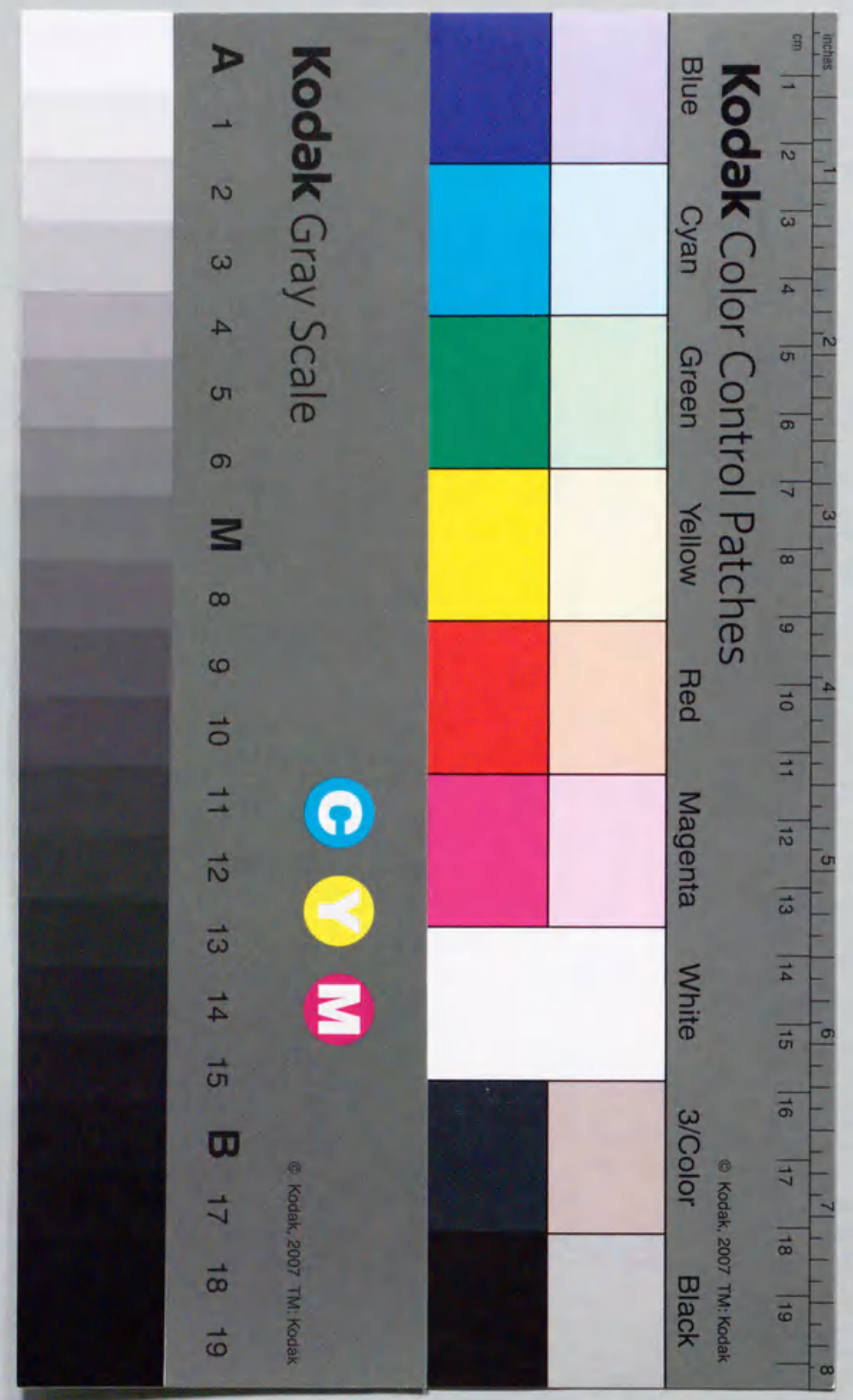


セルラ配列型SIMDプロセッサの
構成法と応用に関する研究

近藤 利夫



報告番号 乙 第 5067 号

セルラ配列型SIMDプロセッサの
構成法と応用に関する研究

近藤 利夫

第1章 序論	1
1.1 研究の背景と目的	1
1.2 セルラ配列型S IMDプロセッサの位置付けと問題点	2
1.2.1 セルラ配列型S IMDプロセッサの位置付け	2
1.2.2 MIMDプロセッサとの対比	2
1.2.3 セルラ配列型S IMDプロセッサの問題点	4
1.3 対象とするセルラ配列型S IMDプロセッサの基本アーキテクチャ ..	5
1.3.1 代表的なS IMDプロセッサの概観	5
1.3.2 研究の対象とする基本アーキテクチャ	6
1.4 研究の課題とそれに対する解決方針	8
1.4.1 目的達成上の課題	8
1.4.2 課題解決のための基本的な方針	8
1.4.3 PE配列部構成上の課題に対する具体的な取り組み方針 ..	9
1.5 本論文の構成	10
第2章 2次元隣接結合ネットワークを補強する伝搬演算機構	12
2.1 はじめに	12
2.2 伝搬演算の概要	13
2.3 基本的な伝搬演算方式	14
2.3.1 逐次伝搬方式	14
2.3.2 バイパス伝搬方式	18
2.3.3 セレクティブ伝搬方式	20
2.4 N進ツリー構成の伝搬演算機構	21
2.4.1 バイパスN進ツリー構成導出	22
2.4.2 セレクティブN進ツリー構成導出	24
2.4.3 N進ツリー構成の演算時間とPOE数	25
2.5 性能、ハードウェア規模の評価	26
2.5.1 速度性能	26
2.5.2 ハードウェア規模	27
2.6 N進ツリー構成伝搬演算機構を付加した1次元PE配列の チップレイアウト	29
2.6.1 N値の選択	29
2.6.2 4 ² 進POUの構成	30

2. 6. 3	4 ² 進ツリー構成のチップレイアウト	30
2. 6. 4	性能推定	30
2. 7	むすび	31
第3章	可変構造と演算アルゴリズム	33
3. 1	はじめに	33
3. 2	修飾型SIMD制御による適応的な可変構造の実現	33
3. 2. 1	PE配列への適合的なPU割付け	34
3. 2. 2	指定PEでの直並列変換	35
3. 2. 3	指定PEでの転送バス間の乗り換え	36
3. 2. 4	PE毎の転送方向設定	38
3. 2. 5	PE毎のALUファンクション設定	38
3. 3	可変構造を活用する効率的な演算法	39
3. 3. 1	伝搬加算を巧みに利用する矩形の局所領域の総和演算	39
3. 3. 2	論理回路シミュレーション	41
3. 4	むすび	43
第4章	データの転送、回転、およびメモリアクセス機構	44
4. 1	はじめに	44
4. 2	従来のメモリアクセス機構の問題点	45
4. 2. 1	メモリアクセス方式とその有用性	45
4. 2. 2	実現上の問題点	46
4. 3	提案する転送回転型メモリアクセス機構の構成	46
4. 4	メモリアクセス法	49
4. 4. 1	間接アドレッシングによるデータアクセス	49
4. 4. 2	従来型2次元アクセス	51
4. 4. 3	疑似2次元アクセス	52
4. 4. 4	補正無しでの2次元アクセス	54
4. 4. 5	間接的な2次元アクセスとそれに用いる90度回転	55
4. 5	評価	58
4. 5. 1	構成の規則性, 単純性	58
4. 5. 2	従来型2次元アクセス機構との高速性比較	59
4. 5. 3	各アクセス法の高速性比較	60

4. 6	むすび	62
第5章	1ビットセルラ配列型SIMDプロセッサシステムの設計と試作	64
5. 1	はじめに	64
5. 2	PE配列とその構成要素であるセルラ配列型LSIの設計	64
5. 2. 1	PE配列とLSIの構成	64
5. 2. 2	PE構成	66
5. 2. 3	伝搬演算機構の構成	66
5. 2. 4	回路技術	67
5. 2. 5	チップレイアウト	69
5. 2. 6	設計・試作結果の評価	72
5. 3	小型システムLISCARIの構成	73
5. 3. 1	全体構成	74
5. 3. 2	1次元/2次元両用のPE配列	74
5. 3. 3	90度回転機構	75
5. 3. 4	制御構成	76
5. 3. 5	装置化	77
5. 3. 6	ソフトウェア開発環境	79
5. 3. 7	基本性能評価	80
5. 4	大規模システムAAPの構成	81
5. 4. 1	全体構成	82
5. 4. 2	制御スカラ演算ユニット	82
5. 4. 3	PE配列とその実装技術	82
5. 4. 4	ソフトウェア開発環境	85
5. 4. 5	基本性能	85
5. 5	むすび	85
第6章	8ビットセルラ配列型SIMDプロセッサシステムの設計と試作	87
6. 1	はじめに	87
6. 2	小型経済化のためのアプローチ	87

6. 3. 1	PE配列部の構成	89
6. 3. 2	プロセッシングエレメント (PE) の構成	89
6. 4	PE配列LSIの設計	90
6. 4. 1	全体設計	92
6. 4. 2	PE設計	92
6. 4. 3	セレクトティブツリー構成伝搬演算機構の設計	92
6. 5	並列プロセッサシステムの実現	93
6. 5. 1	ボード	93
6. 5. 2	システム	94
6. 6	基本性能評価	94
6. 6. 1	基本演算	94
6. 6. 2	90度回転処理	95
6. 7	むすび	97

第7章	大規模セルラ配列型SIMDプロセッサAAPの LSI-CADへの応用	99
7. 1	はじめに	99
7. 2	論理シミュレーション	99
7. 3	自動配線	101
7. 3. 1	新配線アルゴリズムの概要	102
7. 3. 2	AAP1の配線処理速度評価	104
7. 4	むすび	106

第8章	小型セルラ配列型SIMDプロセッサLISCAR1の文字 認識処理への応用	107
8. 1	はじめに	107
8. 2	手書き英数カナ認識処理	107
8. 2. 1	認識処理の構成	108
8. 2. 2	並列処理と逐次処理の切り分け	109
8. 2. 3	並列処理の内容	109
8. 3	印刷漢字読み取り	111
8. 3. 1	認識処理の構成	111

8. 3. 2	並列処理の概要	112
8. 4	手書き漢字読み取り	115
8. 4. 1	認識処理の構成	115
8. 4. 2	並列処理の概要	115
8. 5	速度性能の改善	117
8. 5. 1	高速化手法	118
8. 5. 2	高速化の結果と考察	119
8. 6	むすび	122

第9章	LISCAR2の漢字認識処理への応用	123
9. 1	はじめに	123
9. 2	認識処理の構成	123
9. 3	並列処理の概要	124
9. 3. 1	特徴抽出処理	124
9. 3. 2	次元圧縮処理	129
9. 3. 3	大分類処理	129
9. 3. 4	識別	132
9. 4	速度性能の評価	132
9. 5	むすび	134

第10章	結論	135
10. 1	研究の内容と主な成果	135
10. 2	今後の研究の方向と課題	138

参考文献	144
------	-----

研究業績リスト	149
---------	-----

略語、定義語一覧	151
----------	-----

第1章 序論

1.1 研究の背景と目的

逐次型の汎用プロセッサは、誕生以来長足の進歩を遂げ、この20年間で2桁から3桁の小型化、高速化が達成されている。それにもかかわらず、より高性能な並列プロセッサに対して、今日、大きな期待が寄せられている。これは、常に逐次プロセッサの処理能力を何倍も上回る超高速処理の要求が存在するばかりか、性能向上が進めば進むほど応用範囲、適用分野が広がり、その結果、性能向上あるいは小型化、経済化の要求がますます強まり、従来からの逐次プロセッサの改良だけで対応できる見通しが立たなくなっているからである。

この逐次プロセッサの性能限界の克服をねらった並列プロセッサの研究は、種々のレベルで古くから行われてきた。中でも、多数のプロセッサを用いるプロセッサ配列型並列プロセッサは、原理的に非常に高い並列度が達成できるため、超高速処理の実現が可能な技術として位置付けられてきた^{1,2,3,4)}。しかし、並列度が上がる分、プログラム開発が困難になるばかりか、ハードウェア規模も著しく大きくなるため、なかなか商用レベルのシステムとして普及するには至らなかった。

一方、プロセッサのハードウェアを支えるLSI製造技術の進歩は著しく、2年で2倍(論理LSI)ないし3倍(メモリLSI)のペースで着実に集積規模が増加している。その結果、現在では論理LSIで1000万トランジスタを越えるもの⁵⁾まで実現されるに到っている。しかし、CAD技術は、一時この進歩に追い付けず、1970年代から1980年代にかけて、論理LSIの分野で設計ネックが顕在化し、製造技術の進歩を活かし切れない傾向にあった。この現状を打開すべく、設計容易な論理LSIの構成法が求められて、1980年代には構成の単純化を追究したRISC型逐次プロセッサ⁶⁾と単純な演算ユニットの規則的配列からなるシストリック配列⁷⁾の研究が期を同じくして活発化した。

このような状況のもとで、本研究は、LSIの製造技術の進歩を活かしきれる並列プロセッサの実現をめざし、シストリック配列に匹敵する単純性と規則性を備えたセルラ配列型のSIMDプロセッサを取上げ、その構成法、設計法、処理アルゴリズムを追究する。そして、特定の応用分野で、従来不可能であったような超高速処理と経済性の両立に道を開くとともに、処理対象に適合させる専用ハードウェア並みの性能/価格を備える並列プロセッサシステムの実現を目指す。

1. 2 セルラ配列型SIMDプロセッサの位置付けと問題点

1. 2. 1 セルラ配列型SIMDプロセッサの位置付け

SIMD (Single instruction stream, multiple data-stream) プロセッサは、並列演算部が制御部から配られる単一の命令で一括制御される同一構成のプロセッシングエレメント (PE: Processing Element) 配列で構成されることから、最も単純な構成の並列プロセッサの1つといえる。その中でも、処理語長の短いPEを隣接結合ネットワークで結ぶタイプのSIMDプロセッサは、並列演算部が極めて単純で規則的な構成になることから、PEが細胞にたとえられて、セルラロジック配列、セルラ配列プロセッサ等と呼ばれてきた^{8,9,10)}。

この単純性、規則性ゆえに、その原形は30年以上も前に提案され¹⁾、1ビット構成のPEを用いるタイプを中心にいくつものシステムが開発されてきた^{2,3,11,12,13,14,15)}。1ビット構成のPEが好まれたのは、構成が最も単純であるために大規模なプロセッサ配列を比較的容易に実現できるばかりか、ビット演算を語長分繰り返すこと (ビット直列演算) によって短語長から長語長の演算まで柔軟に対応できるからである。それでも、大規模LSIの利用なくしては、コストに見合った性能が得られなかったのか、広く普及するには至らなかった。ところが、近年のLSI技術の進歩が状況を一変させた。1チップに多数のPEを搭載したVLSIを活用した大規模なシステムが実現され、科学技術計算からAIやニューラルネットの応用にまで際だった性能が発揮されることが示された^{16,17)}。その結果、現在ではセルラ配列型のSIMDプロセッサがスーパーコンピュータの一角を占める代表的なSIMDプロセッサとして位置付けられるに至っている。

1. 2. 2 MIMDプロセッサとの対比¹⁸⁾

SIMDプロセッサを議論する上で常に対比されるのが同じくプロセッサ配列型の並列プロセッサであるMIMD (Multiple instruction stream, multiple data-stream) プロセッサである。

MIMDプロセッサは、PE毎に個別に制御部を有しているためSIMDプロセッサに比べると並列処理の融通性が高く、応用の範囲が広いという利点を有している。また、1チップマイクロプロセッサを組込むことで、PEが比較的容易に構成される利点もある。このため、最近では汎用性が強く要求されるスーパーコンピュータの分野で、これまで主流であったCRAYで代表されるベクトルプロセッサを凌ぐ勢いにある¹⁹⁾。とはいえ、本研究の目的である専用ハードウェアを越えるような高性能/コストの実現、

あるいは既存の小型スーパーコンピュータ並みのコストでの超並列処理の実現は、現状では困難である。これは、MIMDプロセッサでは、各PEが個別にシーケンス制御部を必要とする分ハードウェア規模が大きく、1チップに複数のPEを規則的に並べるLSI向き構成により、並列度をかせぐことが難しいからである。従って、同一ハードウェア規模当たりの並列度は、SIMDプロセッサより低くなってしまふ。その上、各PEを独立あるいは非同期に動作させるためにプログラミングやデバッグが複雑化し、これを少しでも緩和するために、本格的なプログラミング支援ツールを開発しなければならなくなる。

このMIMDプロセッサに比較すると、SIMDプロセッサは、同じプロセッサ配列型の並列プロセッサでありながら、相反する性質を備えている。すなわち、MIMDプロセッサとは異なり、個別にシーケンス制御部を設ける必要がないために、PE構成は逐次型マイクロプロセッサに比べるとかなり単純化される利点がある。特に、セルラ配列型のSIMDプロセッサではネットワークが隣接結合をベースとした単純な構成で済むために、レジスタ付きALUにPE間転送パスを付加する程度の極めて単純な構成でPEが実現される。従って、シストリック配列同様、同一構成のPEを複数、規則的に配列するチップ構成が可能となり、最近のLSIの集積能力をほぼ完全に活かすことが可能になる。また、SIMDプロセッサは並列演算部全体が同期して動作する1つの演算器とみなすことができるので、プログラミングやデバッグについても、従来の逐次プロセッサに近い感覚で行うことができる利点がある²⁰⁾。

SIMDプロセッサの欠点は、先に述べたようにMIMDプロセッサほど並列処理に対する融通性が高くなく、このために高速化可能な処理が定型的な配列演算に限られる傾向にあることである。この欠点は、並列演算部全体を単一命令で一括制御するSIMD型の制御方式そのものに起因しており、SIMDプロセッサの利点とは裏返しに関係にある。従って、この欠点を根本的に解消することは不可能といえる。それでも、それを少しでも改善しようと、多少の並列度低下を伴っても各PEにMIMD的な自律制御機能を付加することで汎用性を高める努力が、これまで種々なされてきた²⁰⁾。今後も、この汎用性向上と並列度向上との両立の問題は、SIMDプロセッサにおけるアーキテクチャ設計上の最重要課題の一つであり続けるものと考えられる。

SIMDプロセッサとMIMDプロセッサとの間で比較すべき、もう一つのアーキテクチャ上の重要なポイントに、PE間ネットワーク構成があげられる。これは、大規模化と遠隔PE間の通信性能の確保を両立しようとする、構成次第ではプロセッサ配列自体を上回ってしまうほどPE間ネットワークのハードウェア規模が大きくなるからで

ある。このため、PE間ネットワークの構成法は、PE間の高速度転送とハードウェア規模低減の両立をめざして、階層バス、クロスバ、2次元隣接結合、ツリー、ハイパーキューブ、多段結合網等の多くのネットワークが、古くから研究されてきた²¹⁾。

MIMDプロセッサとSIMDプロセッサのネットワークの違いは、転送を分散的に制御する必要性の有無にある。すなわち、MIMDプロセッサではランダムに発生する任意のPE間の転送をサポートするために分散的な制御機構が必要なものに対し、SIMDプロセッサでは転送をプロセッサ配列全体で同期的に行えるので、分散的な制御機構を必要としない。従って、その分SIMDプロセッサのネットワークの方がハードウェア規模が小さくなるといえる。しかし、このようにいえるのはあくまでプロセッサ配列のサイズを同一と仮定した場合のことである。ネットワークのハードウェア規模がPEあるいはネットワークを構成する中継ノード間の結線の総数に強く依存することからすると、MIMDプロセッサ以上の並列度をねらうSIMDプロセッサの方が、PE間ネットワーク増大の問題は、より深刻と言える。

なお、ここでは、もう一つの代表的な配列型の並列演算器であるシストリック配列との比較は行っていない。この理由は、シストリック配列が、一般に汎用性がほとんどなく、プロセッサというよりも専用の高速演算器であり、専用ハードウェアの構成要素にはなっても、並列プロセッサそのものではないと判断したからである。一部には、かなりの汎用性を備えたシストリック配列も実現されてはいるけれども、それらはSIMDプロセッサ、あるいはMIMDプロセッサの一種ともみなせ、SIMDプロセッサとMIMDプロセッサの比較のみで十分と考えたからである。

1. 2. 3 セルラ配列型SIMDプロセッサの問題点

セルラ配列型SIMDプロセッサは、当然SIMDプロセッサの一種であり、処理の柔軟性が低い、並列度が大きい分ネットワークのハードウェア規模低減の制約が強い等、抱えている問題自体が異なるわけでない。違いは、他のSIMDプロセッサに比べ、これらの欠点を補うための対策がとりにくいことにある。この理由は、問題解決に効果的だからといって、PE単体あるいはPE間ネットワークの複雑化をまねくような構成を採れば、単純、規則的かつ高並列なセルラ配列構成の利点そのものが損われてしまうことにある。ただし、どこまでの複雑さが許されるかについては、その時点でのLSIの製造技術、設計技術、さらには目的とするシステムの位置付けによっても変わるので注意が必要である。結局、このような周囲の状況をみながら、いかにSIMDプロセッサの問題点を克服し、性能/コストの高い、あるいは並列度の大きいシステムを実現するかが、アーキテクチャ構成上のポイントとなる。

1. 3 対象とするセルラ配列型SIMDプロセッサの基本アーキテクチャ

1. 3. 1 代表的なSIMDプロセッサの概観

前節で述べたように、セルラ配列型SIMDプロセッサの歴史は古く、その発展形としてILLIAC IV³⁾をはじめとする各種のSIMDプロセッサが開発されてきた。その結果、本研究の開始時点には、すでに近年のSIMDプロセッサの代表的なアーキテクチャが出そろっていた。本節では、これらのうち本研究を進めるに当たって特に参考としたプロセッサであるILLIAC IV、グッドイヤーSTARAN^{11,14,15)}、ICL DAP¹²⁾、CLIP-4¹³⁾を概観し、セルラ配列構成を前提としたアーキテクチャの利害得失を明確にする。

ILLIAC IVは、現在のSIMDプロセッサのルーツとも呼ぶべきSIMDプロセッサであり、図1.1に示す単一段の2次元隣接結合ネットワーク、PE毎に個別のアドレス生成を可能とするレジスタ間接アドレッシングによるデータアクセス、PE毎の個別の制御レジスタによるPEの活性制御等の現在のSIMDプロセッサを構成する上での基本技術のほとんどが組込まれている。しかし、PE自体は64/32ビットの浮動小数点演算、48/24/8ビットの固定小数点演算等をサポートする大掛かりなもので、本研究の目標とする複数PE搭載のチップを並べるだけのセルラ配列構成を実現するには程遠い複雑な構成をとっている。

ICL DAPは、商用化された数少ない汎用のSIMDプロセッサの一つである。ILLIAC IVとは単一段の2次元隣接結合ネットワークでPE間を結合するPE配列構成を採る点では共通ながら、4,096個の1ビット構成PEからなる典型的なセルラ配列構成を採る点では対極的である。PEの構成は極めて単純で、1チップに多数のPEを搭載することも比較的容易で、LSI化によりセルラ配列構成の利点を引出すのに十分な構成と言える。

CLIP-4は、先駆的なセルラ配列型SIMDプロセッサであり、実際に4PEを搭載したLSIで96x96のPE配列を構成している。PEの構成は単純でPE対応の局所メモリの容量も小さいが、セルラ配列型SIMDプロセッサにおける伝搬演算機能の有用性を示した最初のシステムである。

グッドイヤーのSTARANも商用化された数少ないSIMDプロセッサの一つである。極めて単純な構成の1ビットプロセッサ256個でPE配列を構成する点ではICL DAP、CLIP-4と共通である。違いは、2次元の隣接結合ネットワークの代わりに、一種の多段結合のネットワークであるフリップネットワークを装備している点と局所メモリの配列をMDA (Multi-Dimensional Access memory) 構成としている点

にある。これらのフリップネットワークとMDAとを用いることで、MDA内の2次元配列データに対し行単位、列単位の両方向のデータ列に対するアクセス機能（2次元アクセス機能）を実現している。また、フリップネットワークは単体で置換やシフトも実現できる。このようにSTARANの構成は高い機能を提供できるものの、MDAの構成にはメモリのビット毎にアドレスを修飾できる必要がある、フリップネットワークの構成には多数のスイッチ、多数の遠隔PE間の結線が必要になる等、単一段の2次元隣接結合ネットワークに比べるとハードウェア規模はかなり大きくなってしまふ。

1. 3. 2 研究の対象とする基本アーキテクチャ

(1) 2次元セルラ配列構成

STARANで採用された多段の結合ネットワークは、機能の高さからは極めて魅力的である。実際、STARANが実現している2次元アクセス機能は、1次元モードで動作させるPE配列で2次元の配列データを処理する場合に必須とも言える機能である。しかし、STARANの採用している多段の結合ネットワークは、ハードウェアに対する負担が大きく、大規模化あるいは小型経済化を第一義的な目標とするセルラ配列構成の考え方とは相容れない。

そこで、ILLIAC IV、ICL DAP等にならひ、M個からなるPE配列に対してPE間最大距離（PE間距離）が $O(\sqrt{M})$ と比較的小さく、かつLSIチップの平面構成に整合する点でLSI化に適している図1.1の2次元隣接結合型をベースとしたネットワークを採用する。

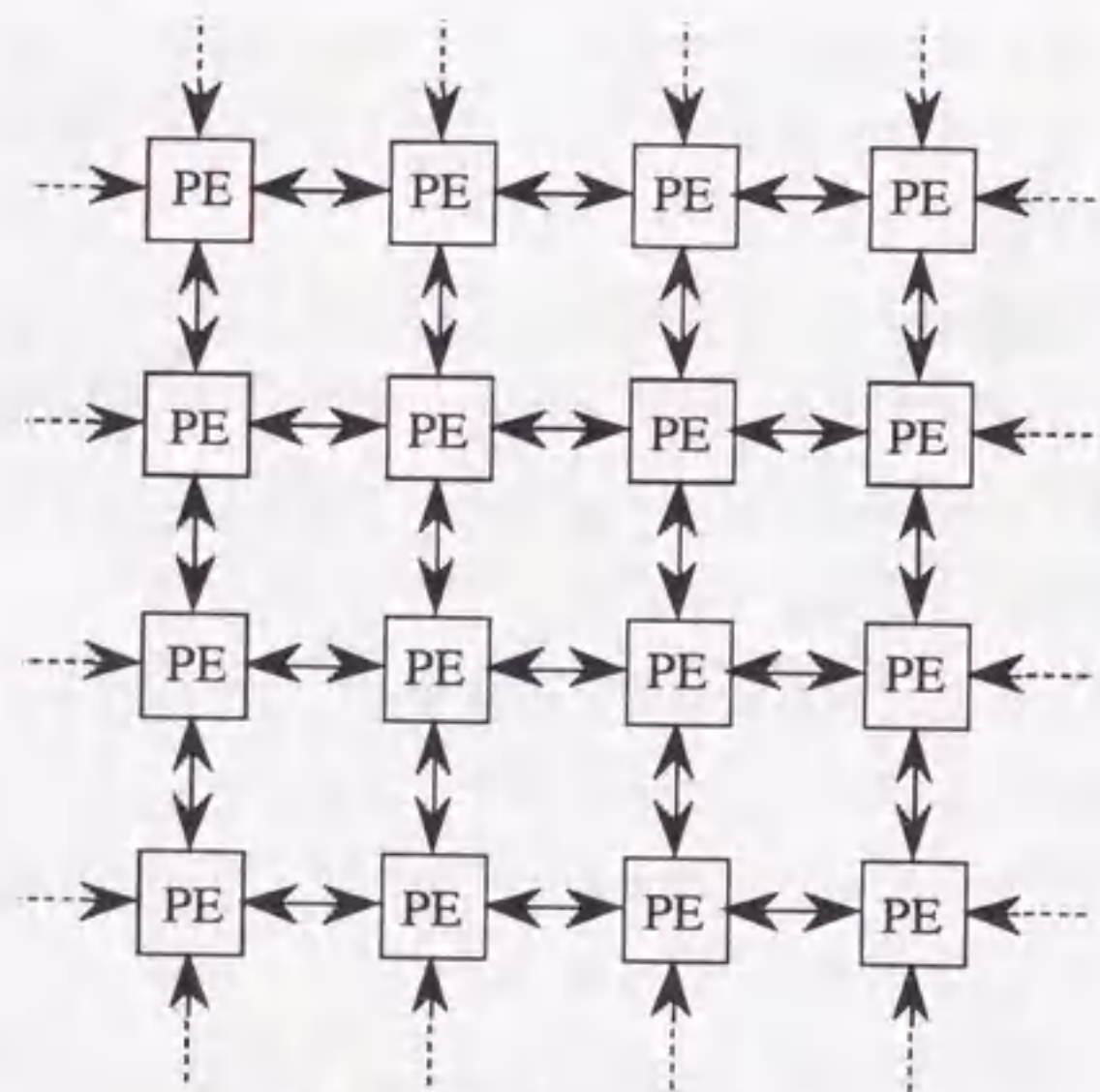


図1.1 PE配列の構成

ただし、 $O(\sqrt{M})$ のPE間距離は大規模なPE配列において遠隔PE間の転送を行うのに十分高速とはいえず、なんらかの補強手段を考える必要がある。PE配列を1次元モードで動作させる場合の2次元アクセス機能についても、多段の結合ネットワークを用いない単純な方法を検討する必要がある。

(2) 修飾型SIMD制御

本研究で対象とするSIMDプロセッサの全体構成を図1.2に示す。図1.1を機能分担が明確になるように書き直し、制御部を付加した構成であり、上部の制御部で発生した信号により、その下に配置する2次元の隣接結合ネットワーク、PE配列、メモリ配列等からなる並列演算部全体を制御する典型的なSIMD制御方式を採用する。また、この図に示されるように、メモリユニットは局所メモリとしてPEと1対1で接続する標準的な構成（各PEが1個の局所メモリを個別にアクセスする構成）を採る。異なるメモリユニットとの通信は、PE配列の上側に示される隣接結合ネットワークを介して行う構成とする。

PE単位の自律的な制御手段としては、ILLIAC IV以来、各種のSIMDプロセッサに採用されてきた活性制御機構を採用する。この制御機構は、演算結果をデスティネーションのレジスタへ書込むか否かをPE毎に設ける1ビットのレジスタ（活性

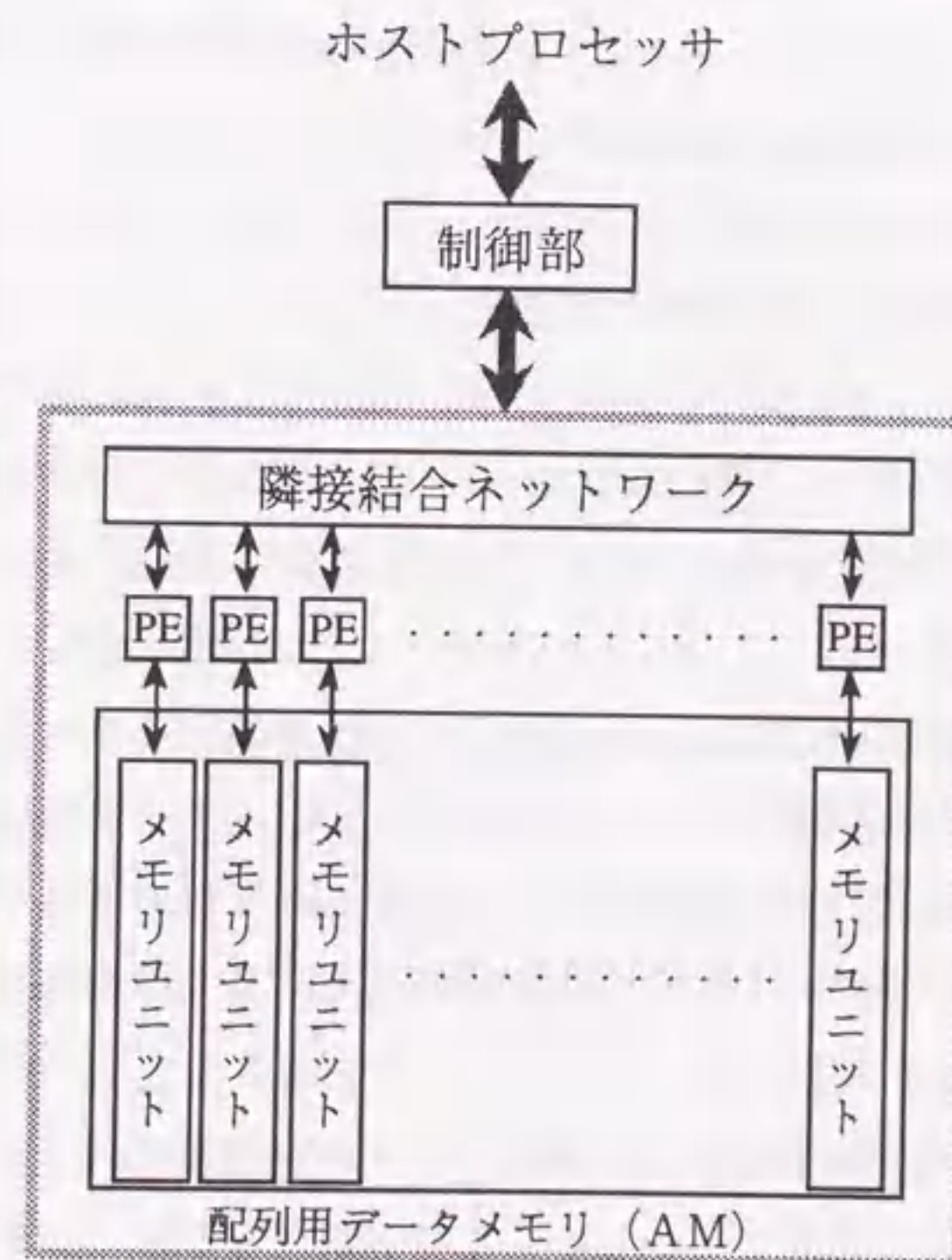


図1.2 セルラ配列型SIMDプロセッサの構造

制御レジスタ)で制御することにより等価的にPEの活性制御を行うものである。さらに、制御部からの全PE共通の命令を、PE単位で修飾するための制御レジスタを個別に設け、PE配列の機能の向上をはかる。ただし、この制御レジスタによる命令修飾機能の一つであり、ILLIAC IVで採用されたレジスタ間接アドレッシングによりPE単位で独立のアドレスを生成する機能については、そのままの構成では搭載しない。PE単位に個別にアドレス端子が必要で、LSIやボードに対する負担が大きいからである。

1. 4 研究の課題とそれに対する解決方針

1. 4. 1 目的達成上の課題

本研究では、1. 1で述べた目的を達成するために、最近のLSIの集積能力を完全に引出せるセルラ配列構成との両立が可能な以下の3つのPE配列部構成法、

- ①効率的な遠隔PE間転送が可能なPE間ネットワーク構成法
- ②2次元アクセス、PE単位の自律的なアドレッシング機能等の多様なアクセスモードをサポートできる局所メモリアクセス機構構成法
- ③高速性と処理データの構造変化に適宜対応可能な適応的なプロセッサ配列構成法を示すこと、さらにはこれらに基づいてPE配列部を実現する場合の
- ④PE配列部用LSIおよびSIMDプロセッサシステムの構成法、設計法
- ⑤実際の応用でのPE単体あるいはPE配列部全体の演算法、処理法

を明らかにすることを主要な課題とする。

1. 4. 2 課題解決のための基本的な方針

課題解決のために、2つの基本方針を採る。第1の方針は、ハードウェア化が性能向上に寄与しない部分については極力ソフトウェアで実現することである。これは、性能向上と開発工数低減の両立をはかるためである。この方針は、デバッグが困難でかつ設計・開発の一巡の期間(TAT)の長いハードウェアに比べると、同一機能の実現ならば、ソフトウェアの方が開発工数が数段小さいことに基づく。ただし、多数のアプリケーションプログラムを開発しようとする場合、ハードウェアの機能が低いとその分、個々のプログラムの記述負担が増加して、全体の開発工数が増加する可能性については、コンパイラやマクロライブラリを必要に応じて構築することで回避できるものとして考えない。

第2の方針は、応用を特定の分野に限定し、その中で専用ハードウェア並みの高い性能/コストの確保を目指すことである。これが実現できれば、並列プロセッサの持つ処理アルゴリズムや処理方式の可変性により専用ハードウェアに対し十分優位にたてるか

らである。また、これは、逆に不特定多数の分野にまで適用できるだけの汎用性の確保までは目指さないことを意味している。セルラ配列構成との両立が難しいだけでなく、多くのアプリケーションの開発と、そのための汎用コンピュータ並みのプログラム開発環境の提供が必要となり、全体の開発負担が大きくなりすぎるからである。

特定の分野としては、文字認識とLSI-CADの2つを選ぶ。身近な応用の中では、2次元の配列データを直接並列に演算することで処理が進められる割合が大きく、SIMDプロセッサ向きの応用と考えられるからである。また、一方の文字認識処理が、OA機器に組込むために小型経済性を要求される分野、もう一方が大規模なLSIを効率良く設計するために超高速処理が要求される分野と、セルラ配列型SIMDプロセッサの能力をはかる最適な応用と考えられるからである。

1. 4. 3 PE配列部構成上の課題に対する具体的な取り組み方針

本研究の最重要課題であるPE配列部構成法に対しては、セルラ配列構成の利点を享受するため、単純性、規則性を優先することとし、以下の3つの方針のもとに取り組む。

- ①直接サポートする算術演算機能は、短語長の固定小数点演算に限定する。
- ②PE間ネットワークの構成は2次元の隣接結合にCLIP-4で用いられたPE間の伝搬演算機構を組合わせたものとする。
- ③制御レジスタによる各PEの個別の命令修飾機能、またはこれを利用した複数PEの協調動作により対応できる範囲で、可変のPE配列構造を実現する。

ここで、①の方針は、1. 4. 2で選定した応用に対しては短語長の固定小数点データ主体の処理で十分であるとの観点に基づき、PE構成の単純化を指向したものである。従って、低速なソフトウェアエミュレーションで実行することになる長語長の固定小数点演算、浮動小数点演算等の頻度の高い応用では、十分な高速性能が得られなくなる点に留意する必要がある。

②の方針は、ハードウェア規模低減をはかるために、遠隔PE間の転送バスの負担の小さいネットワークとすることを目指したものである。伝搬演算機構は、演算結果をPE間で順次伝搬させることで結果の得られる演算(走査演算[scan]とも呼ばれる)を対象とした高速化機構であり、効率の良い遠隔PE間の転送演算を実現するセルラ配列構成向きの転送手段の1つと考えられる。しかし、高速化可能な転送・演算に限られるわけで、処理アルゴリズム自体をこれに適合させることによって実行性能の向上を図る必要が生じる。

③の方針は、PEの構成をすべて同一とし、PE配列の規則性を確保するための方針

である。従って、PEの構成上の工夫により、いかに柔軟な可変構造を実現するかがポイントになる。

1.5 本論文の構成

本論文は10章からなる。以下、各章の概要を示す。

第2章では、本研究で前提とした隣接結合ネットワークに対し、遠隔PE間の演算あるいは転送を補強する伝搬演算機構の構成方式として、バイパス方式とセレクトティブ方式を提案し、高速化のためにツリー型の階層構成を導くとともに、そのハードウェア規模、速度性能を評価する。

第3章では、PE配列の一層の稼働率向上を目指し、全PE共通の命令を、PE毎に付加する制御レジスタで個別に修飾することによって実現されるPE配列の各種の可変構造的な機能について述べる。さらに、前章で述べた伝搬演算と組み合わせることによって、汎用化に有効なPE配列内の矩形の局所領域毎の総和演算やLSI-CADの論理シミュレーションの中核部分等が効率的に実行できることを示す。

第4章では、LSIやボードの構成に負担の小さいメモリアクセス機構として、基本のSIMDプロセッサにわずかな改造を加えるだけで実現される転送回転型メモリアクセス機構を提案し、その各種のアクセス法を明らかにするとともに、従来型の2次元アクセス機構と比較して、端子数、アクセス速度等を評価する。

第5章では、セルラ配列構成の利点をすべて引出すことを狙って開発した1ビット構成の大規模セルラ配列型SIMDプロセッサAAP (Adaptive Array Processor)、1ボードSIMDプロセッサLISCAR (Line Scannable Cellular ARray processor)、およびこれらのPE配列部の構成要素であるセルラ配列型LSIの設計、構成について述べる。また、それらの基本性能も評価する。

第6章では、小型経済化と高性能化の両立をはかるために、第4章までの成果を積極的に取り入れるとともに、積和演算能力を強化した8ビット構成の1ボードSIMDプロセッサLISCAR2の設計、構成、基本性能評価等について述べる。

第7章では、身近な応用であるLSI-CADの中から、特にセルラ配列型SIMDプロセッサ向きと考えられるマスクパターンレイアウト設計の自動配線と、汎用コンピュータでは処理時間が大きくなりすぎる論理シミュレーションを取上げ、AAP向きの処理アルゴリズムを提案するとともに、性能を評価する。

第8章では、従来の文字認識処理ハードウェアでは実現困難であった小型高速性と搭載アルゴリズムの可変性の両立を目指し、複数の文字認識処理の搭載を試みた結果について述べる。具体的には、手書き英数かな文字認識処理、印刷漢字認識処理、手書き漢

字認識処理のLISCAR1上での処理方法を示すとともに、処理性能を評価する。さらに、高速化手法を明らかにし、速度性能改善をはかるとともにLISCAR1の問題点を示す。

第9章では、LISCAR1の処理速度が特に問題となった手書き漢字認識処理に対して、さらに高度化したアルゴリズムをLISCAR2で高速に処理することを試みる。具体的には、高速化のポイントになる特徴抽出処理、次元圧縮処理、大分類処理、識別処理の処理方法を示すとともに、処理性能を評価する。

第10章では、本論分の研究成果を総括するとともに、今後の研究の方向と課題を展望する。

第2章 2次元隣接結合ネットワークを補強する 伝搬演算機構

2.1 はじめに

1.3.2の基本アーキテクチャの項で説明した2次元隣接結合型のネットワークは、比較的小さなハードウェア負担で、頻度の高い近傍PE間通信を効率良く実行できる点で優れている。しかし、規則性の高い定型的な処理に限っても総和演算のように一定の割合で出現する遠隔PE間の通信を高速に実行することはできない。かといって、他の高度のネットワークのように複雑な構成をとることはセルラ配列構成では許されない。そこで、遠隔PE間転送については、汎用性は劣るもののハードウェア負担の小さい伝搬演算機構を、方針の項でも述べたように隣接結合ネットワークの中に組込むことにする。本節ではこの伝搬演算機構の最適な構成法を追究し、その性能、ハードウェア規模を評価する。

さて、伝搬演算とは、データ配列 $a_1, a_2, a_3, \dots, a_M$ に対し $a_1(a_1@a_2), (a_1@a_2@a_3), \dots, (a_1@a_2@a_3@a_M)$ を求める演算であり、配列要素がPE配列に分散配置されている場合に、途中の演算結果をPE間で順次伝搬させることによって求められる。ここで、@は、+、V、 \wedge 、min [最小値を求める演算]、max [最大値を求める演算] 等の演算子である。この演算は、放送、総和、部分和、部分コピー等の比較的頻度の高い遠隔PE間の転送、演算を実現できるため、セルラ配列型SIMDプロセッサの中核的な演算機能の1つとして位置付けられている^{13,22)}。また、ソート・マージ処理、グラフ・幾何学問題等の基本処理を高速に実行できることも、示されている^{23,24,25)}。

このような有用性から、セルラ配列型SIMDプロセッサでは、伝搬演算が、古くから、PE間の伝搬（縦続的に接続されるPE間で、途中、クロックで同期をとることなくデータを転送すること）を利用することで実現されていた^{13,22)}。しかし、この方式ではプロセッサの並列度に見合った伝搬演算性能は得られない。演算にかかわるPEが、伝搬の先頭の波面に位置するPEに限られ、PE配列の並列性を生かせないからである。

並列度を上げる比較的単純な方法の一つに、縦続的に連なるPE列の適当な間隔Nごとにバイパスを設ける方式がある。この方式によれば、長さMの配列データに対する伝搬演算時間を、 $O(\sqrt{M})$ 程度にまで短縮できる。それでも、Mが大きくなると十分な演算速度は得られなくなって来る。

これに対して、最近では2進ツリー構成の演算機構により、伝搬演算時間を $O(\log_2 M)$ 程度まで短縮する手法が提案されている²³⁾。ハードウェア量も、ハイパーキューブ型のPE間転送系で、同一の2進伝搬演算アルゴリズムをエミュレートする²⁶⁾のに比べると、かなり小さい。それでも、単純なPE配列のハードウェア規模に比べ無視できる規模ではない。ハードウェア量低減には2進からN進への拡張が有効と考えられるにもかかわらず、速度性能の低下をきらってか、十分な検討はなされていない。

そこで、ここでは高速性を損なわないN進構成の実現をめざし、バイパス型伝搬および加算器の高速化によく用いられるセレクトティブ伝搬のツリー状の階層化について検討する。そして、多段の階層化がN進ツリー構成の伝搬演算機構として一般化されること、Nを4に選ぶことで、従来の2進ツリー構成に比べ、1/2程度のハードウェア規模で2倍の高速性の得られるツリー構成の伝搬演算機構が実現できること等を示す。

本章では、伝搬演算の概要、基本的な伝搬演算方式による伝搬演算機構の概要、バイパス型・セレクトティブ伝搬の多段の階層化によるN進ツリー構成伝搬演算機構の導出、各方式の性能・ハードウェア規模の評価、LSIのチップレイアウトの基本的な検討結果等について述べる。

2.2 伝搬演算の概要

はじめに、1次元配列（ベクトル）Aに対する代表的な伝搬演算の例を図2.1に示す²³⁾。ここで、Segment-Flagsとは、配列Aが複数の部分配列に分割される場合の部分配列の始点か否かを示す1次元配列、+propagation(A)は、Aに対する加算伝搬結果、max-propagation(A)は、Aに対する最大値伝搬結果である。first-propagation(A)は、始点データの部分配列内での放送伝搬の結果である。

first-propagationでは、部分配列の始点に送信PEを、その部分配列のいずれかに受信PEを、それぞれ対応させることで、遠隔PE間のデータ転送を行なうことができる。

A	=	[7 1 3 9 4 2 5 0 6]
Segment-Flags	=	[1 0 0 1 0 1 0 0 0]
+propagation(A)	=	[7 8 11 9 13 2 7 7 13]
max-propagation(A)	=	[7 7 7 9 9 2 5 5 6]
first-propagation(A)	=	[7 7 7 9 9 2 2 2 2]

図2.1 代表的な走査演算の例

この転送機能は、転送区間が重複しなければ、それぞれに対応する部分配列内で並列にデータ転送が行なえる点で優れている。これに対し、一般によく用いられる共通バスでは、同時にバス全体で1組のデータしか転送できないので、転送区間が重複していても、データの組数分だけの転送を繰り返す必要がある。従って、転送速度が同等ならば、所要時間は繰り返しの回数分だけ長くなってしまふ。

伝搬演算の有用性の一端を、文字認識処理で特徴抽出に用いられる黒点連結の長さ(ランレングス)の抽出処理を例に示す。この場合の黒点連結の長さの抽出処理は、2値データの1次元配列Bに対して1の連なりの長さを求める処理であり、加算型の伝搬演算として図2.2のように簡単に実現される。ここで、Segment-Flagsの配列データは、配列Bについて1の連なりの左端、0の連なりの左端をそれぞれ検出することで生成することができる。

なお、以上の説明では、理解し易くするためにワード単位の演算例を示しているけれども、ビット単位の伝搬演算機構を利用する場合には、このワード単位の演算をビット直列処理により行うことになる。

B = [0011111000110]
 Segment-Flags = [0010000100101]
 +-propagation(A) = [0012345000120]

図2.2 +-propagationによる黒点連結の長さの抽出

2.3 基本的な伝搬演算方式

2.3.1 逐次伝搬方式

PE間の隣接結合をそのまま利用する最も基本的な伝搬演算方式である。図2.3に示すように、送信PEからの送出データを、途中、各転送PEでいちいち同期をとるこ

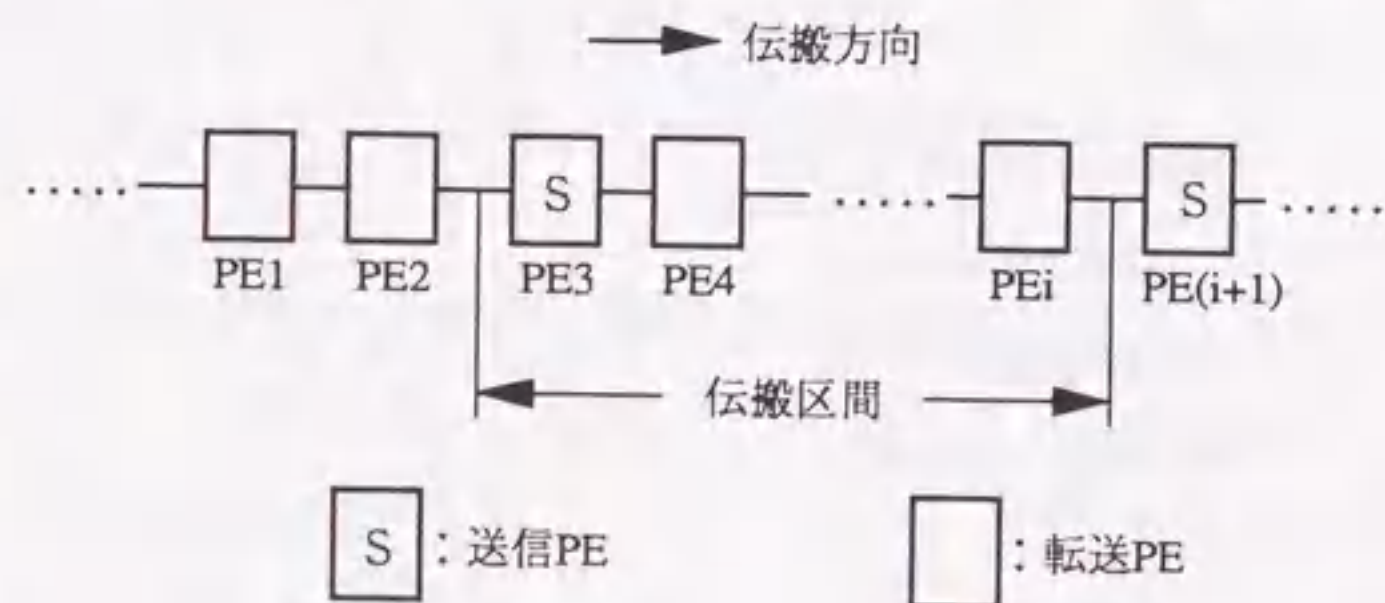
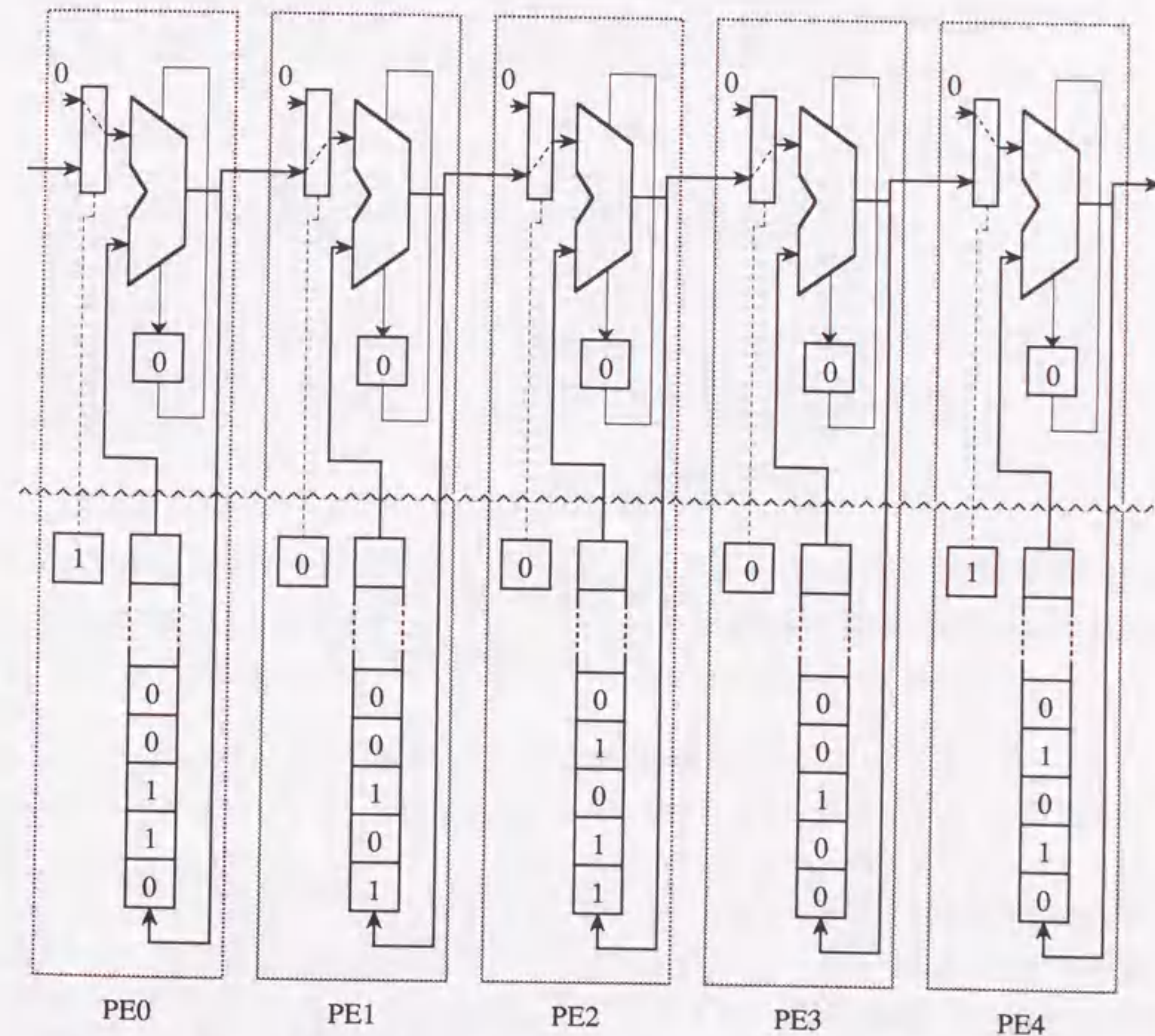


図2.3 PE間の隣接結合をそのまま利用する逐次伝搬演算方式

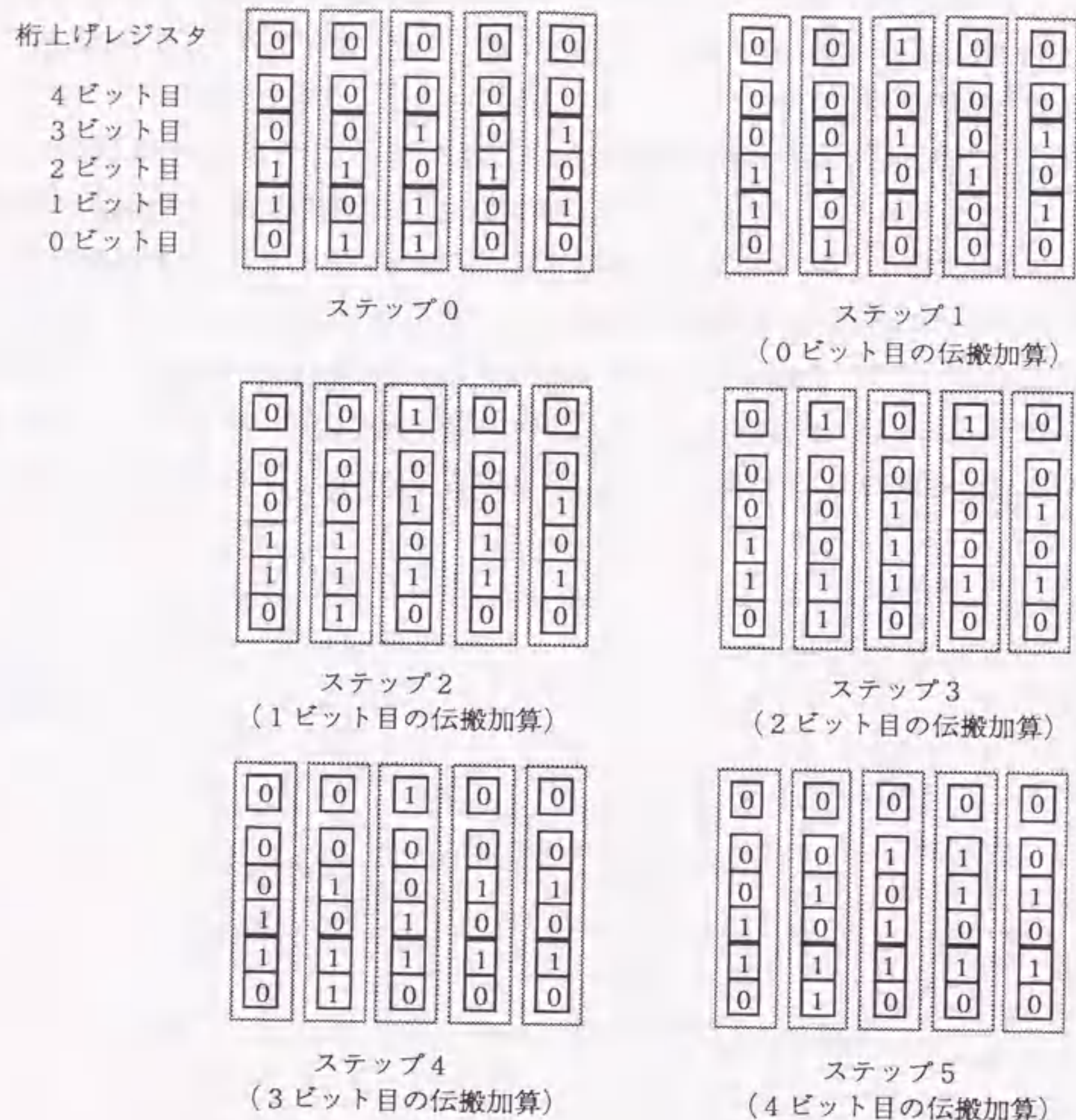
となく、所定の演算を施しながら順次転送すること(伝搬演算)によって部分配列ごとの伝搬演算を実現する。この動作から明らかなように、部分配列の区切りは送信PEの指定により自由に設定できる。

図2.4は、伝搬演算の中でも、セルラ配列型のSIMDプロセッサにおいて最も基本的かつ有用なビット直列型の伝搬加算を例に、逐次伝搬方式の伝搬演算を具体的に説明した図である。図の(a)は、ビット直列型の伝搬演算を行うための必要最小限のハードウェアである1ビット幅のALU(枠内右上)、レジスタファイル(枠内右下)、入力セレクタ(枠内左上)、桁上げレジスタ(枠内右ALU直下)、セグメントフラグレジスタ(枠内左中央)からなるPEの1次元の配列を示している。図の(b)は、最下位ビットから1ビットずつ順次伝搬加算を実行し、それぞれのPEのレジスタ



(a) 伝搬演算に必要な最小限のハードウェアからなるPE配列

図2.4 ビット直列型の伝搬加算



(b) 伝搬加算の演算手順
図2.4 ビット直列型の伝搬加算

ファイルの0から4ビット目に伝搬演算結果が得られる様子を示している。各ビットの伝搬演算は、各PEが入力セクタから入力される伝搬データとレジスタファイルの指定ビットのデータと桁上げレジスタに保持されているデータとの間で加算を実行し、和を右隣のPEに出力するとともに、桁上りを桁上げレジスタに格納することで、左から右方向に向かって進行する。この例では1ビットずつ伝搬演算を繰り返すビット直列型の伝搬加算を説明したが、2ビット以上を単位とする伝搬演算も可能である。この演算幅の拡張によれば、繰り返しの回数が低減されるので、伝搬演算の高速化が可能である。また、本方式は後述する他方式に比べるとこの演算幅の拡張が容易であるという特

徴を持っている。その理由は、PE間の結線を、隣接PE間のみ演算のビット幅分だけ設ければよく、他方式のように遠隔PE間に多重の結線を設ける必要がないからである。

伝搬演算に直接関わる波線より上の部分を、伝搬演算エレメント(POE)の配列として切り出すことにより構成した伝搬演算機構のブロック図を図2.5に示す。図2.4とは異なり、伝搬演算を実行する上で意味のない左端のPOEは省略している。各POE内のP1_iは伝搬演算器であり、入力セクタ、ALU、桁上げレジスタ等で構成され、左隣から太実線の経路を介して入力される伝搬データとPEから太実線の経路を介して供給される演算対象配列の要素データとの間で所定の演算を実行し、結果を右隣へやはり太実線の経路を介して出力する。太破線は演算結果をPEへ戻すための経路である。細破線はPEのセグメントフラグレジスタから入力セクタに対する制御信号を受け取る経路であり、この値が1の場合に、P1_iは始点の伝搬演算器として、PEからの要素データを直接右隣のPEに出力するようにしている。

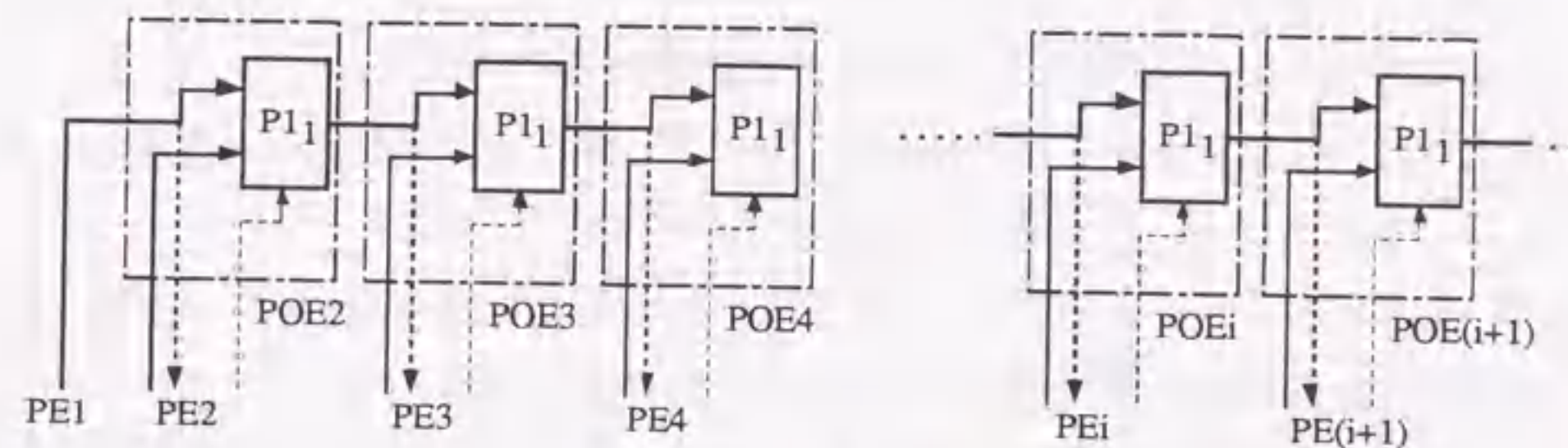


図2.5 逐次型伝搬演算機構

ここで伝搬演算時間 T_{OP} を評価する。伝搬演算時間 T_{OP} は、途中のPOEをすべて経由する必要があることを考慮すると、個々のPOEの伝搬時間が積算されるため、最悪条件(部分配列が配列全体に一致する場合)で、次式のように表される。

$$T_{OP} = (M - 1) \cdot t_{pd} \quad (2.1)$$

ここで、Mは伝搬演算対象の1次元データ配列全体に1対1で対応する1次元PE配列のサイズ、 t_{pd} は各P1_iの伝搬遅延時間である。この(2.1)式から明らかなように伝搬演算時間が経由するPE数に比例して増加するため、遠隔PE間の転送には全く対応できないように見える。しかし、1チップに複数PEを搭載するセルラ配列構成では、伝搬時間 t_{pd} が、POEの構成を最適化することでクロックのサイクルタイムの数

分の1以下にまで低減できることから、伝搬距離をPE通過段数で十数段以下に限定できるならばこの単純な構成でも転送サイクル数の多さが問題になることはない。なお、ハードウェア規模の指標であるPOE数 N_p は、図2.5から明らかなように、 $M-1$ である。

2.3.2 バイパス伝搬方式

2.3.1の単純な構成では、配列サイズ M が大きく、PE通過段数で数十段以上の距離の転送が必要になる場合には、当然伝搬時間が長くなりすぎる。これを改善する最も直感的な方法は、バイパスを付加することである。図2.6は、そのバイパスを一定の間隔ごとに付加するバイパス伝搬方式の概念図である。図2.7は、バイパスの付加単位であるPE配列（伝搬演算ユニット：POU）の構成を示している。POUは、伝搬演算器 P_{11} 、 P_{21} 、始点検出器DTからなる $N-1$ 個のPOEおよびユニット内の集約用伝搬演算器 P_{12} からなり、次の3つの演算を並列に行うことで、伝搬演算を実行する。

- ① POU内の P_{11} の並びで行なう局所的な伝搬演算、
- ② 局所的な伝搬演算結果をPOU間でバイパスと演算器 P_{12} を介して累積し、各POUのオフセット値を求める伝搬演算、
- ③ この累積によって得られたオフセット値と先の局所的な伝搬結果から全域的な伝搬演算結果を得る P_{21} での演算。

ここで、POUのオフセット値とは、そのPOUがカバーする部分伝搬領域の直前までの伝搬演算結果である。また、図2.7で、実線、破線等によって示される経路の役割は、図2.5の場合と同じである。

POU内に始点のPEがある場合、演算器 P_{12} は、バイパス側の入力は無視し、直前のPOENからの入力をそのまま出力する。また、ユニット内の始点以降の演算器 P_{21} は前段のユニットからの累積結果を無視し、 P_{11} で得られた結果を最終的な伝搬演

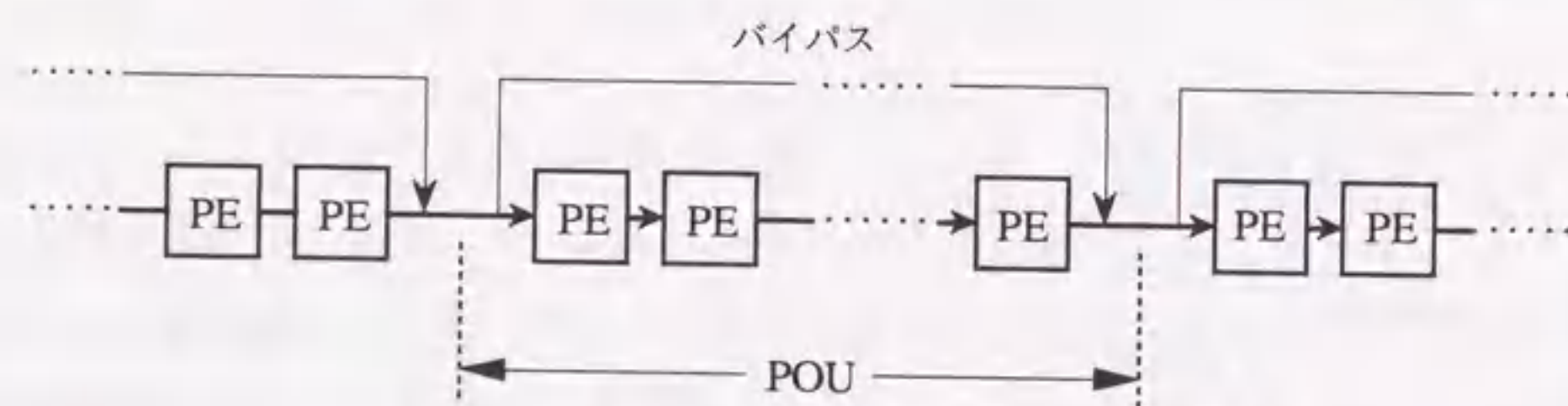


図2.6 バイパス型伝搬方式の概略

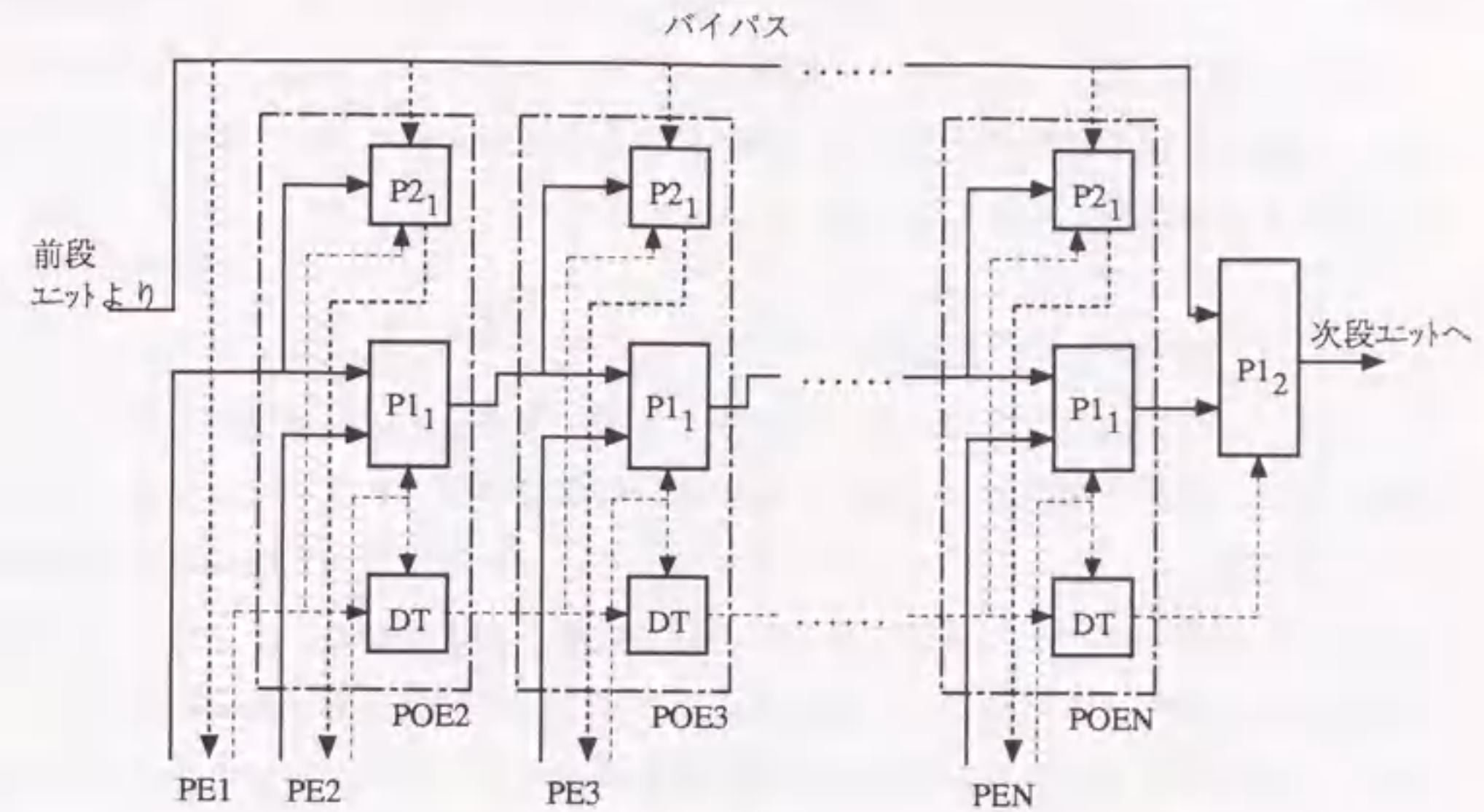


図2.7 バイパス付き伝搬演算ユニットPOUの構成

- P_{11} : 局所伝搬用演算器, P_{12} : 累算用演算器,
 P_{21} : 広域的な演算結果生成用演算器,
 DT: 送信PE検出器, POE $_i$: 伝搬演算エレメント

算結果とする。これらの始点の有無に関係する制御は、始点検出器DTの連なりから生成される信号によって行なう。なお、 P_{11} 、 P_{12} は同一構成の伝搬演算器である。

POUの構成・動作から明らかなように、POU内の伝搬は、POUごとに並列に実行され、POUの個数分だけの並列度が得られる。広域的な伝搬演算結果を得るための P_{21} の演算についてはPE間で独立に実行され、PE数分だけの並列度が得られる。

以上より伝搬演算の伝搬遅延時間 T_{OP} は、

$$T_{OP} = (N-1)t_{pd} + \{(M/N)-1\}t_{pd} + t_{pd} \quad (2.2)$$

となる。ここで、第1項はPOU内の伝搬遅延時間、第2項はPOU間の累積のための伝搬遅延時間、第3項は最終的な結果を得るための P_{21} の演算遅延時間である。Nは各々のPOUが分担する部分配列サイズである。 P_{11} 、 P_{12} 、 P_{21} の演算機能が同一であることから、伝搬遅延時間はいずれも t_{pd} としている。また、POUが、 $N-1$ 個のPOEと1個の P_{12} の計N個と、伝搬演算対象の部分配列の要素数と同数で構成

されることから、全体のPOEとP1₂の個数N_pはMとなる。なお、ビット幅の拡張は、ビット幅に比例して増加する伝搬演算系のハードウェア規模については問題ないものの、遠隔PE間の結線に相当するために、長くなりがちなバイパスがビット幅分必要になる点で、逐次型より難しくなる。

2. 3. 3 セレクティブ伝搬方式

以上述べてきた各方式では、その構成から明らかなように、高速化が必要ならば、演算器、POE間の結線のビット幅、すなわち伝搬演算の演算単位を大きくすることによって対応できる。しかし、バイパス伝搬方式では、長くなりがちなバイパス用結線の増加が、LSIやボードの端子数ネックの原因になる可能性がある。このため、小型化・経済化を優先する場合には、演算単位拡大による高速化はできれば避けたい。

逆に、演算単位をビット直列型の伝搬演算を前提に1ビットに限定するならば、加算器の高速桁上げ方式の一つであるセレクティブ伝搬方式²⁷⁾によって、一層の伝搬演算の高速化が可能である。

この方式は、前段ユニットからの入力される可能性のあるデータ（演算単位を1ビットに限定する場合には0, 1の2通りしかない）の両方に対し、ユニットごとに専用にした伝搬演算系で、先行して伝搬を行なっておき、その結果を実際の入力で選択することによってユニット間の伝搬の高速化をはかる。

図2. 8に演算単位が1ビットのセレクティブ構成のPOUの構成を示す。各POEの演算器P1_AとP1_Bは演算器の幅が1ビットに限定されることを除きバイパス伝搬方式のP1₁と同一機能の演算器である。それぞれはユニット内で伝搬系を形成し、ユニットに対する入力に先行して、0入力と1入力の2通りの伝搬演算を行なう。各POEの下側のセレクタSELは、POUへの実際の入力に応じて、2通りの先行演算結果を選択し、最終的な結果を得るためのものである。また、ここでは、理解の容易化のため図示していないが、加算や減算の伝搬演算を1ビット単位で逐次的に実行する際に必要となる桁上げに関しても、同様に先行生成結果を二者択一する構成によって求められる。

なお、この構成では、基本の伝搬方式、バイパス伝搬方式とは異なり、演算結果が伝搬方向にずれることがないので、後処理無しで最終的な伝搬演算結果が得られる。

伝搬演算時間T_{OP}は、以上の動作より、

$$T_{OP} = N t_{pd} + \{(M/N) - 1\} t_{pds} + t_{pds} \quad (2. 3)$$

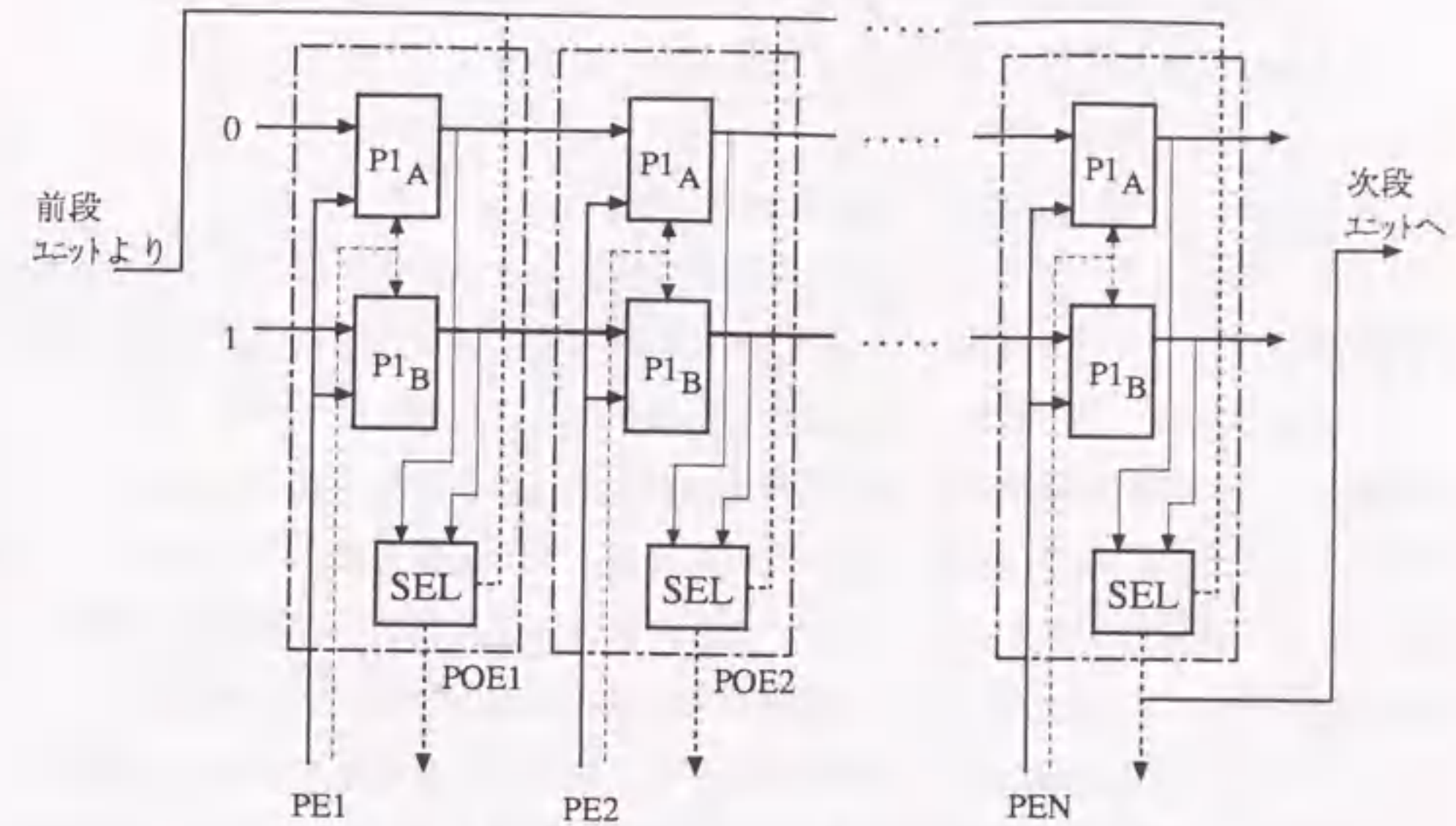


図2. 8 セレクティブ伝搬ユニットPOUの構成
P1_A, P1_B:演算器, SEL: 2-1 セレクタ

となる。ここで、第1項はPOUの内部遅延時間、第2項はPOU間の遅延時間、第3項は、先行演算結果の二者択一のための遅延時間である。また、t_{pd}はP1_AあるいはP1_Bに関わるPOEの伝搬遅延時間、t_{pds}はセレクタに関わる伝搬ユニットの伝搬遅延時間である。第1項の係数から-1がなくなっている点を除くと、(2. 2)式と同じ形をしている。ただし、伝搬演算器P1₂がセレクタに置換されて遅延時間がt_{pds} (t_{pd}の半分程度以下)になる分、演算時間は短縮される。また、POEの全個数N_pは、各POUが伝搬演算対象の部分配列の要素数と同数のN個のPOEで構成されることから、Mとなる。

なお、このセレクティブ伝搬方式でも、演算単位拡大による高速化は不可能ではない。しかし、拡大可能な演算単位は精々2ビット程度までである。伝搬ユニットごとで、入力される可能性のあるデータの種類(2^W通り。ここでWは、演算単位である。)だけ、伝搬演算系を設けねばならず、これ以上では、ハードウェア規模が大きくなり過ぎるからである。

2. 4 N進ツリー構成の伝搬演算機構

本節では、バイパス方式とセレクティブ伝搬方式の両方に対して一層の高速化をはか

るために、多段の階層化について検討する。ただし、演算単位については、セレクトイブ伝搬方式に合わせてすべて1ビットとして議論を進める。

2. 4. 1 バイパスN進ツリー構成導出

バイパス伝搬方式では、POE間の伝搬遅延低減のために、POE間の並びに対して、POU内にバイパスを設けている。これと同様に、POUの並びに対して同様にバイパスを設ければ、POU間の伝搬遅延も低減される。高速化のために、さらに、その上の階層にもバイパスを設けることが考えられる。このようなバイパス階層の多段化に対応できるように、階層2段に対応している図2. 7の伝搬演算ユニットから、1階層分を切出した伝搬演算ユニットPOUの構成を図2. 9に示す。右端のP1_iの出力がこのPOUのカバーする局所領域の伝搬演算結果、右端のDT_iからの出力がこのPOUのカバーする領域に伝搬演算の始点があるか否かの判定結果であり、かつ各P2_iへ向かう入力にこの局所領域までの伝搬演算結果を入力すれば、各POEまでの伝搬演算結果が求まることから、結局、このPOUをツリー状に階層的に並べることによって、図2. 10に示すN進ツリー構成の伝搬演算機構が構成される。

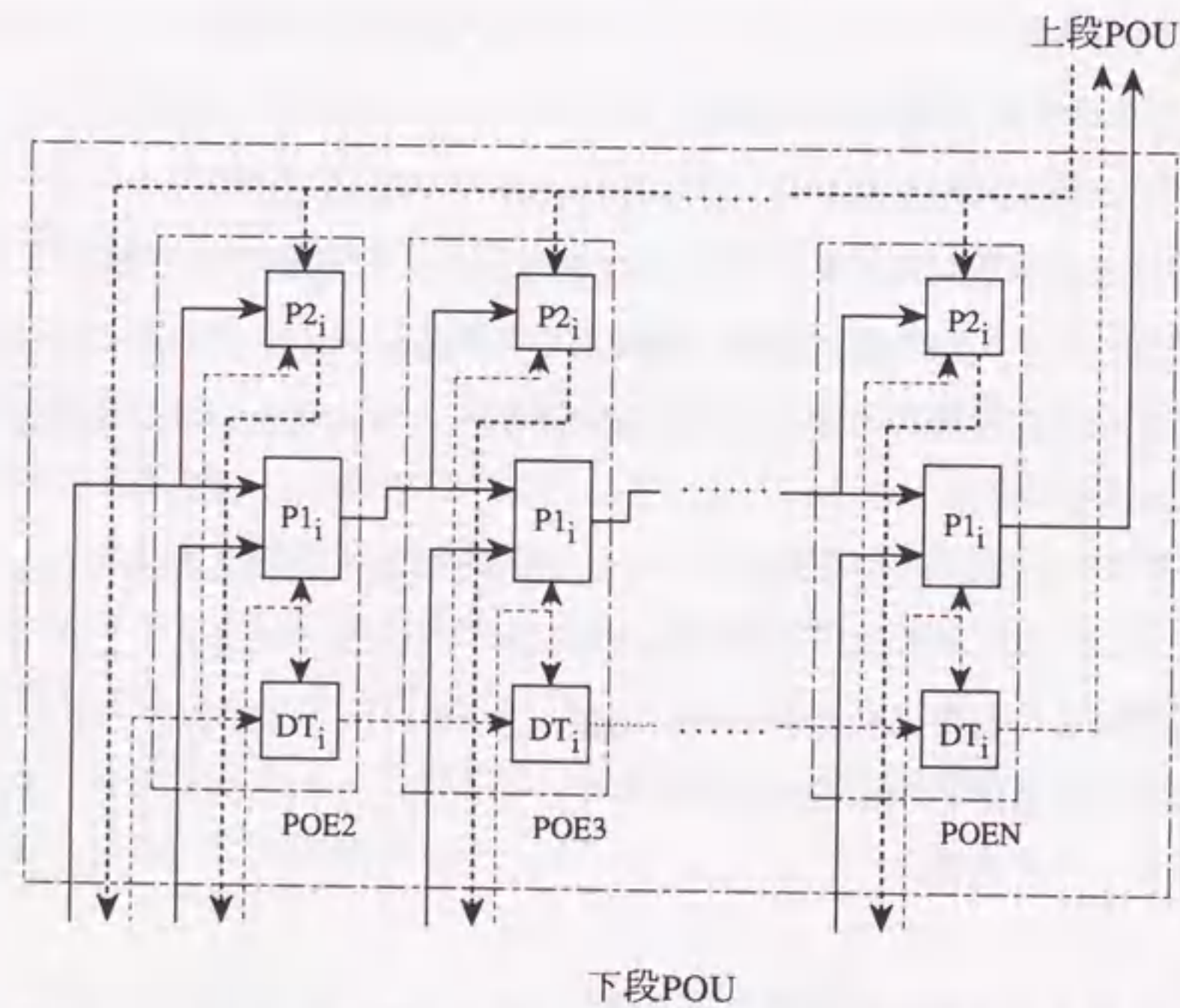


図2. 9 N進ツリー構成用バイパス伝搬演算ユニットPOUの構成

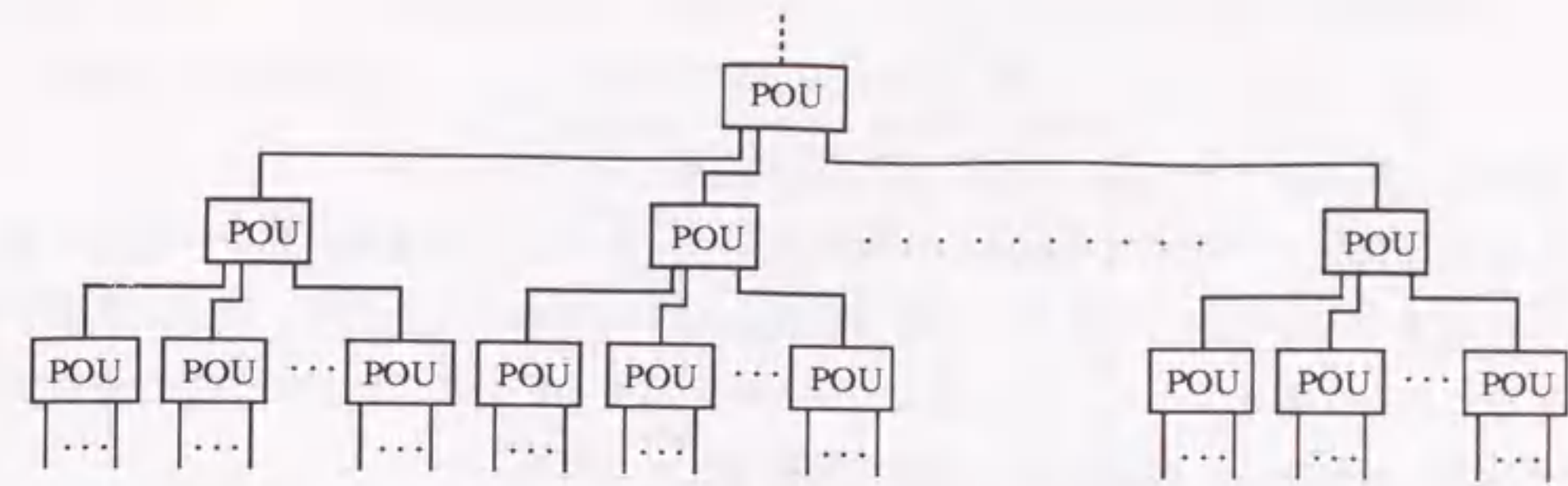


図2. 10 N進ツリー型走査演算機構の構成

このN進ツリーの動作は、

- ①各POEの局所的な伝搬による伝搬演算結果を、ツリーの下段から上段に向かって累積する、
- ②累積結果を元に、上段から下段に向かって各POEのオフセット値を順次求めるの2つからなる。この場合、最下段のPOEからのPEに対する出力は、やはり、オフセット値であり、各PEは、自身の保持データとの間で最終的な伝搬演算結果を得るための後処理を行なう必要がある。

ところで、図2. 10のツリー構成は以下のように、(2. 4)、(2. 5)の式の変形を再起的に繰り返すことによっても、導かれる。

$$\sum_{i=S}^E a_i = \left(\sum_{i=S}^{N[S/N]+N} a_i \right) @ \left(\sum_{k=[S/N]+1}^{[E/N]} \sum_{j=(k-1)N+1}^{kN} a_j \right) @ \left(\sum_{i=N[E/N]+1}^E a_i \right) \quad (2. 4)$$

$$= \left(\sum_{i=S}^{N[S/N]+N} a_i \right) @ \left(\sum_{k=[S/N]+1}^{[E/N]} b_k \right) @ \left(\sum_{i=N[E/N]+1}^E a_i \right) \quad (2. 5)$$

ここで、 $a_1, a_2, a_3, \dots, a_M$ は演算対象のデータ配列を、 $a_s, a_{s+1}, a_{s+2}, \dots, a_{E-1}, a_E$ は、その部分配列を、 Σ は伝搬演算の累積を、 $[]$ はガウス記号を表している。また、 b_k は次式で表される。

$$b_k = \sum_{j=(k-1)N+1}^{kN} a_j$$

(2. 4)式は部分配列全体に対する走査演算が、N項単位の伝搬演算にくくれることを意味している。(2. 5)式は、 Σb_k が上位階層の伝搬演算と考えることがで

き、伝搬演算が2段に階層化されることを意味している。また、(2.4)と同様の変形を Σb_k に対しても加えれば、さらに2段に階層化されることがわかる。結局、この階層化を順次繰り返せば、N進ツリー状の演算手順が導かれる。

このN進ツリーの演算手順から伝搬演算ユニットPOUの機能が導かれる。すなわち、(2.5)式が、あるPOUの下位出力の1つを表す式であり、第1, 2項がそのPOUに対するオフセット値、第3項がそのPOUの出力までの局所的な走査演算結果であること等から、図2.9のPOU機能が求められる。

2.4.2 セレクティブN進ツリー構成導出

セレクティブ伝搬方式もバイパスと同様に階層の多段化が可能であり、図2.11に示すPOUを積み重ねることによって、N進ツリー構成が実現される。このPOUの構成とN進ツリー構成は、セレクティブ伝搬のための伝搬系の二重化を、POE間の伝搬経路から、バイパスに相当するセクタ制御経路の二重化、さらにそれによって新たに必要となる上位のセクタ制御経路と、順次拡張することから導かれる。

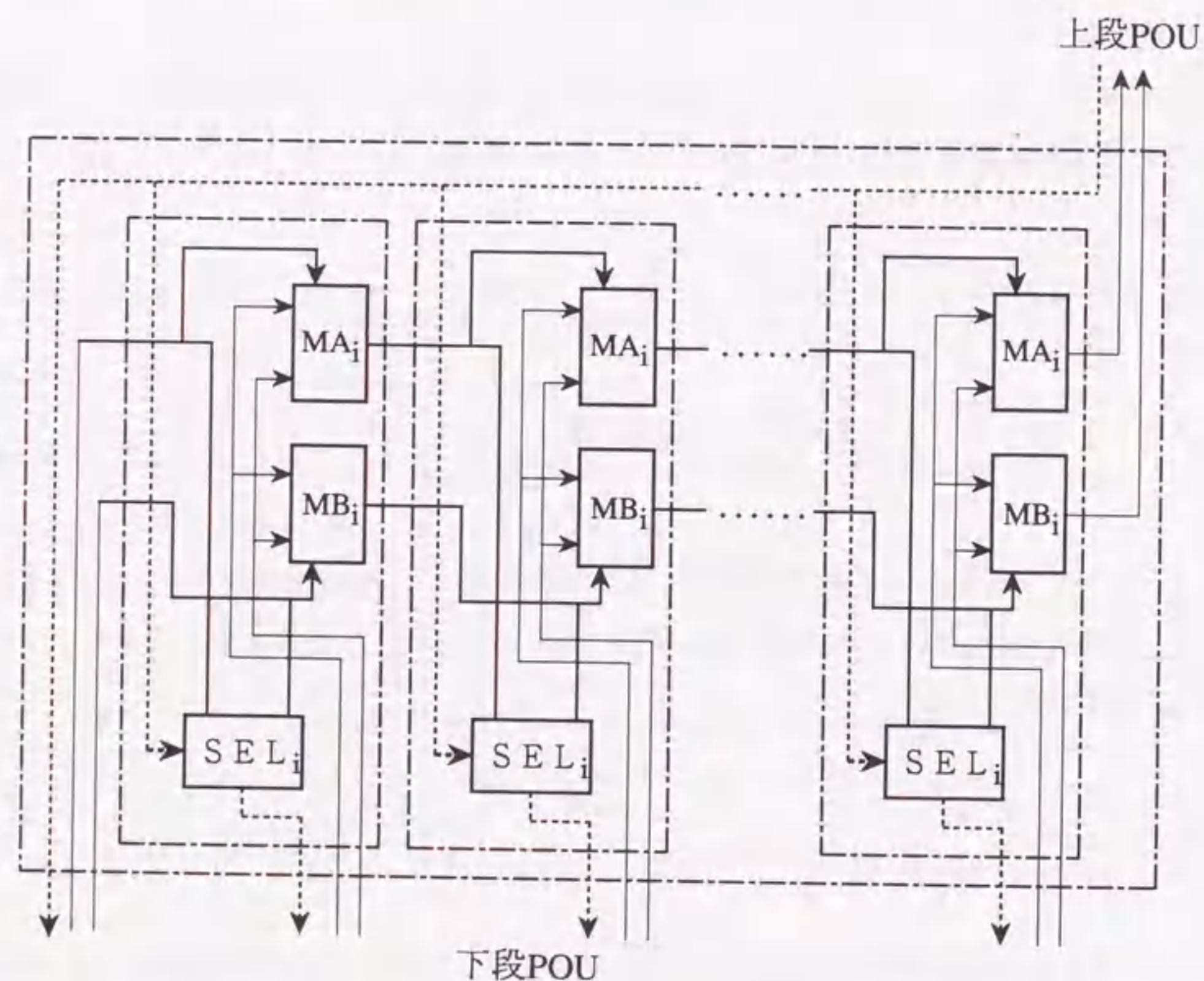


図2.11 N進ツリー構成用セレクティブ伝搬演算ユニットPOUの構成

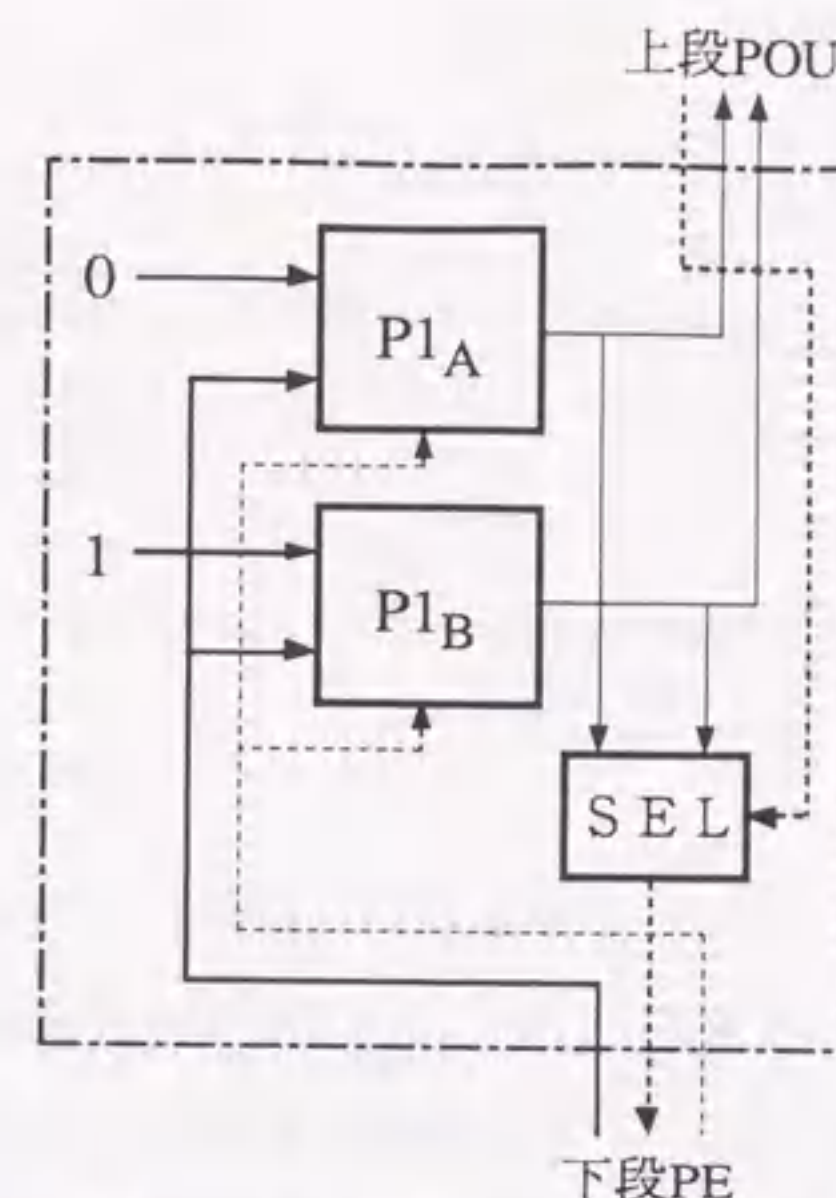


図2.12 インタフェースエレメントIEの構成

ここで、POU内の左右間の太実線は、二重化したセクタ制御信号の伝搬経路である。また、各POEに対する2本の細実線は下位のPOUからの2通りの演算結果の入力経路であり、太破線は上位から下位に向かう2通りの演算結果のいずれを選択するかを制御するための信号用経路である。

最下位のPOEと、PE間の接続は、直接は無理で、2通りの演算結果を生成するために、図2.8のPOEと基本的に同一の構成の図2.12のインタフェースエレメントIEを介して行なう。これによって、演算器の遅延がすべてセクタの遅延に置き換えられることになる。ここで、PEに向かう太実線、太破線、細破線の経路の役割は2.4.1の伝搬方式のそれと同じである。

セレクティブN進ツリーで注目すべきは、POEが、3つの2-1セクタのみで構成される点である。伝搬演算のためにP1, P2の演算器を必要とする図2.9のバイパス構成のPOEと比べると、はるかに単純である。

2.4.3 N進ツリー構成の演算時間とPOE数

N進ツリー構成の伝搬遅延時間は、下位から上位に向かう演算におけるPOU内の伝搬段数がN-1あるいはN段であること、上位から下位に向かう演算における伝搬段数がP2あるいはSELの系の1段のみであること等から、バイパスツリー構成では、

$$T_{OP} = (N-1)(\log_N M) t_{pd} + ((\log_N M) - 1) t_{pd}$$

$$= (N \log_N M) t_{pd} - t_{pd} \quad (2.6)$$

となる。セレクトティブツリー構成では、さらにインタフェースエレメント I E による遅延を考慮すると

$$T_{OP} = (N-1)(\log_N M) t_{pdS} + ((\log_N M) - 1) t_{pdS} + (t_{pd} + t_{pdS})$$

$$= (N \log_N M) t_{pdS} + t_{pd} \quad (2.7)$$

となる。ここで、P1, P2, DTの遅延をいずれも t_{pd} 、各セクタの遅延をいずれも t_{pdS} としている。これらの式より、 t_{pdS} は t_{pd} の半分程度以下にできることから、セレクトティブツリー構成がバイパスツリー構成の2倍程度以上高速であること、N値が自然対数の底 e で最も高速化されること等がわかる。

また、POE数は、いずれの方式もPOUがN-1個のPOEからなることから、各階層の総和より、

$$N_p = (N-1) + N^1(N-1) + \dots + N^{(\log_N M)-1}(N-1)$$

$$= M-1 \quad (2.8)$$

となり、Nによらない。

2.5 性能、ハードウェア規模の評価

2.5.1 速度性能

式(2.2)、(2.3)、(2.6)、(2.7)より、各方式における配列サイズMと数式上の最適条件(階層2段の方式では、 $N = \sqrt{M}$ 、ツリーの方式では、 $N = e$)での等価伝搬段数 T_{op}/t_{pd} との関係を図2.13に示す。ただし、 t_{pdS} は t_{pd} の1/2と仮定している。これより、セレクトティブN進ツリー構成が、従来型の2進ツリー構成、バイパスN進ツリー構成と比べ2倍程度高速であることがわかる。

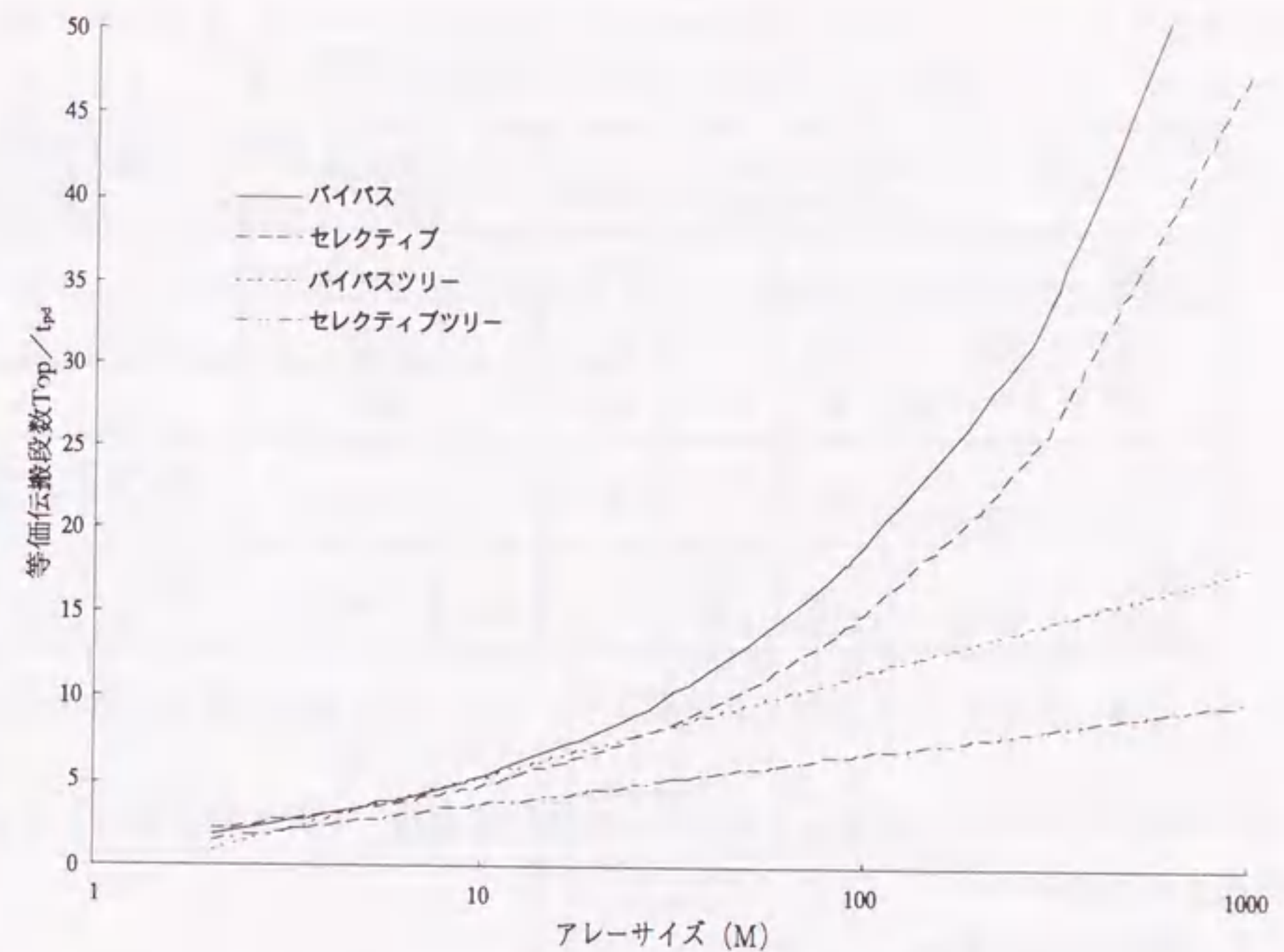


図2.13 配列サイズMと伝搬段数の関係

2.5.2 ハードウェア規模

ハードウェア規模は、その1つの指標であるPOE数からいえば、いずれの方式もM程度で差がない。しかし、実際には差が生じる。これは、①POE単体の規模、②規則性、局所性の低いPOU間の結線量、③PEに一体化が容易な最下層のPOE数等が方式間でかなり異なることによる。これらについて後述する回路構成を前提に各方式を比較した結果を表2.1に示す。この表より、N進ツリー構成について、以下の2点が明らかとなる。

- ①PE配列に一体化困難でPE配列の外に出るゲート数(伝搬演算機構の実効的なハードウェア規模とみなせる)とPOU間結線数は、ともにNの増加にほぼ反比例して減少する。
- ②セレクトティブ構成は、総ゲート数が、バイパス構成に比べ25%程度大きい。しかし、PEに一体化困難な部分のゲート数については、バイパス構成の1/4程度と小さいことから、チップレイアウト時の実効的なハードウェア規模では、バイパス型と同程度かそれ以下にできる。

表2. 1 ハードウェア規模決定要因の比較

方式	POE規模*	IE規模*	POE数	IE数	演算機構 総ゲート数	PE配列以外 のゲート数	POU間 結線数
伝搬	19	—	M-1	—	19(M-1)	0	—
バイパス伝搬	40	—	M	—	40M	19M/N	—
セレクトティブ伝搬	40	—	M	—	40M	0	—
N進ツリー (バイパス構成)	40	—	M-1	—	40(M-1)	$\frac{40(M-N)}{N}$	$\frac{3(M-1)}{N-1}$
N進ツリー (セレクトティブ構成)	10	40	M-1	M	50M-10	$\frac{10(M-N)}{N}$	$\frac{3(M-1)}{N-1}$

*: NTT LSI研究所開発のゲートアレイで構成する場合のセル数

以下に表2. 1のハードウェア規模算出において前提とした伝搬演算器、POE等の回路構成を説明する。

(1) 伝搬演算器の構成

伝搬演算器P1_i, P2_iの回路構成を図2. 14に示す。ゲート数は19である。

ALUのファンクション生成部は、4-1セクタで構成している。マスタースレイブ型のDフリップフロップ(MS-DFE)は、逐次的に算術演算を行なう際に用いる桁上げレジスタである。

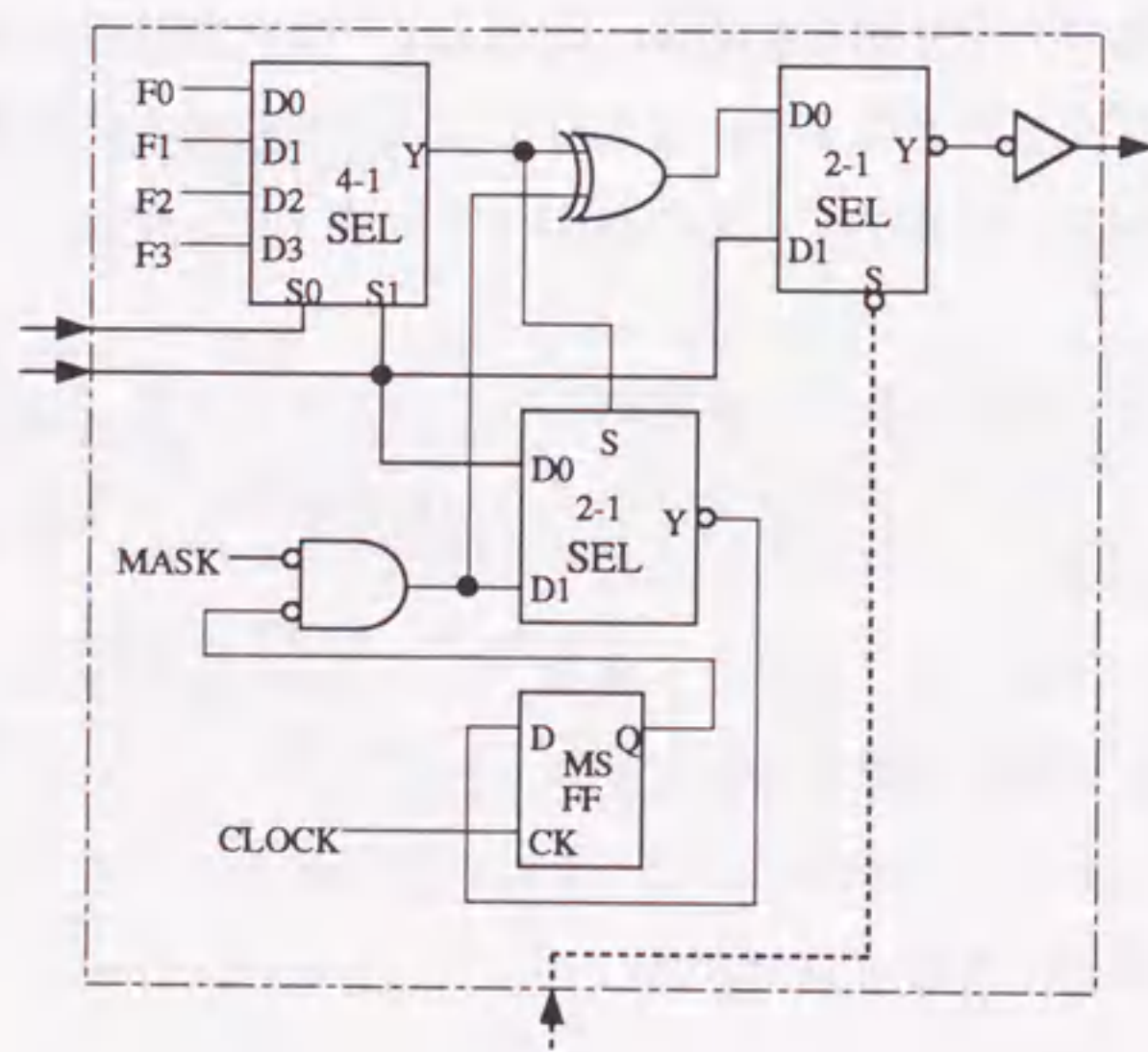


図2. 14 伝搬演算器P1_i, P2_iの構成

伝搬演算器P1A, P1Bの回路構成は、基本的には、P1_i, P2_iと同一である。ただし、桁上げレジスタについては、A, Bで共通に一個設け、セレクトティブ伝搬終了後の確定値を格納する構成としている。P1A, P1B合わせたゲート数は36である。

(2) セレクトティブN進ツリー用POE

2-1セクタをOR-NANDの複合ゲートで構成する場合の回路構成を図2. 15に示す。ゲート数は10である。

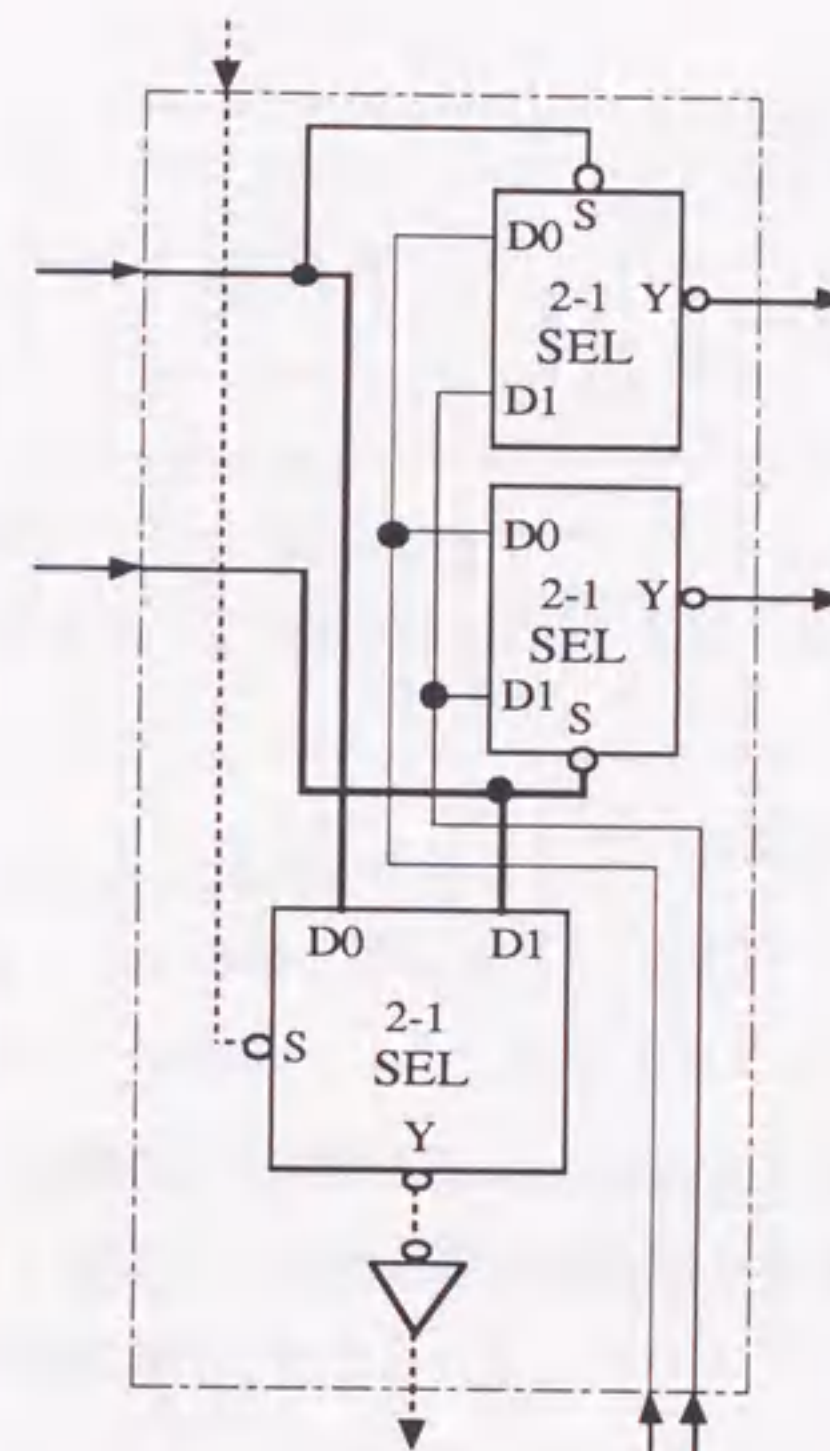


図2. 15 セレクトティブN進ツリーのPOEの回路構成

2. 6 N進ツリー構成伝搬演算機構を付加した1次元PE配列のチップレイアウト

2進ツリーのレイアウト方式²⁸⁾を参考にレイアウトの基本的な検討を行なった結果を述べる。

2. 6. 1 N値の選択

ハードウェア規模低減の観点からは先に述べた通り、Nは大きいほどよい。しかし、高速化の観点からは、2. 4. 3で述べたように自然対数の底e前後が望ましい。そこで、ここでは演算速度を重視し、N値を4に選ぶ。従来の2進ツリー構成と比べると、演算時間、POU間の結線数、PEに一体化困難なゲート規模等がほぼ半分になる。

2. 6. 2 4^2 進POUの構成

4進POUの2段のツリー結合で、 4^2 進のPOUを構成する。レイアウト構成の単純化をはかるためである。そのPOU構成を図2. 16に示す。ここで、上段のPOUは、POE単位に分解し、下段のPOU間に分散配置することで、等価的に1段構成のレイアウトとし、集積度の向上をはかっている。

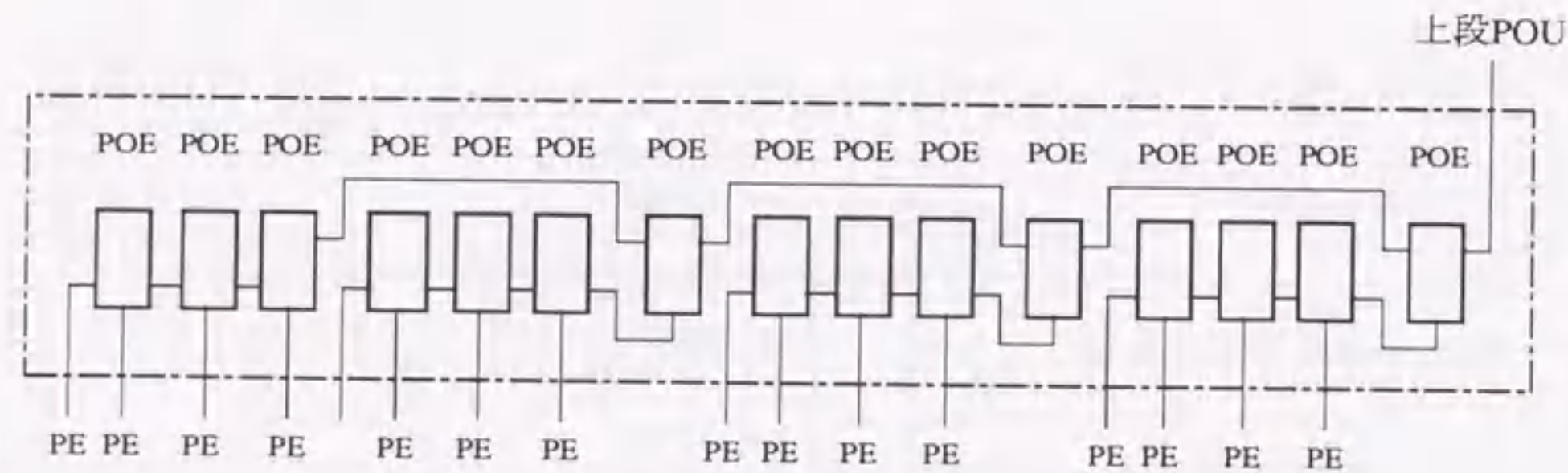


図2. 16 4^2 進伝搬演算ユニットPOUの構成

2. 6. 3 4^2 進ツリー構成のチップレイアウト

1段目の 4^2 進POUを、1次元のPE配列と組み合わせることによって構成した2段構成の 4^2 進ツリーのレイアウトを図2. 17に示す。この図から明らかなようにPOUを構成する15個のPOEは、その下側のPE配列に一体化することで、集積度の向上がはかれる。レイアウトは単体では、長さ256のPE配列に対応するツリーであるが、複数結合することによって、2~3千までの配列サイズに対応可能である。

2. 6. 4 性能推定

このようなレイアウトによって、長さが256の1次元PE配列と伝搬演算機構が5~6年前より利用可能になっている $0.8\mu\text{m}$ 程度の製造技術のCMOS LSIチップに搭載されるならば、POE間の伝搬経路の配線長は数mm以下となり、セレクタの平均の伝搬遅延 t_{pdS} は1ns程度になる。従って、セレクトティブ伝搬方式の場合、式

(2. 5) より、 T_{op} は、20数nsとなる。この値は、実現可能なPEの最小クロックサイクルと同程度であり、長さが256の配列に対する伝搬演算が、1~2クロックサイクルで実現されることになる。

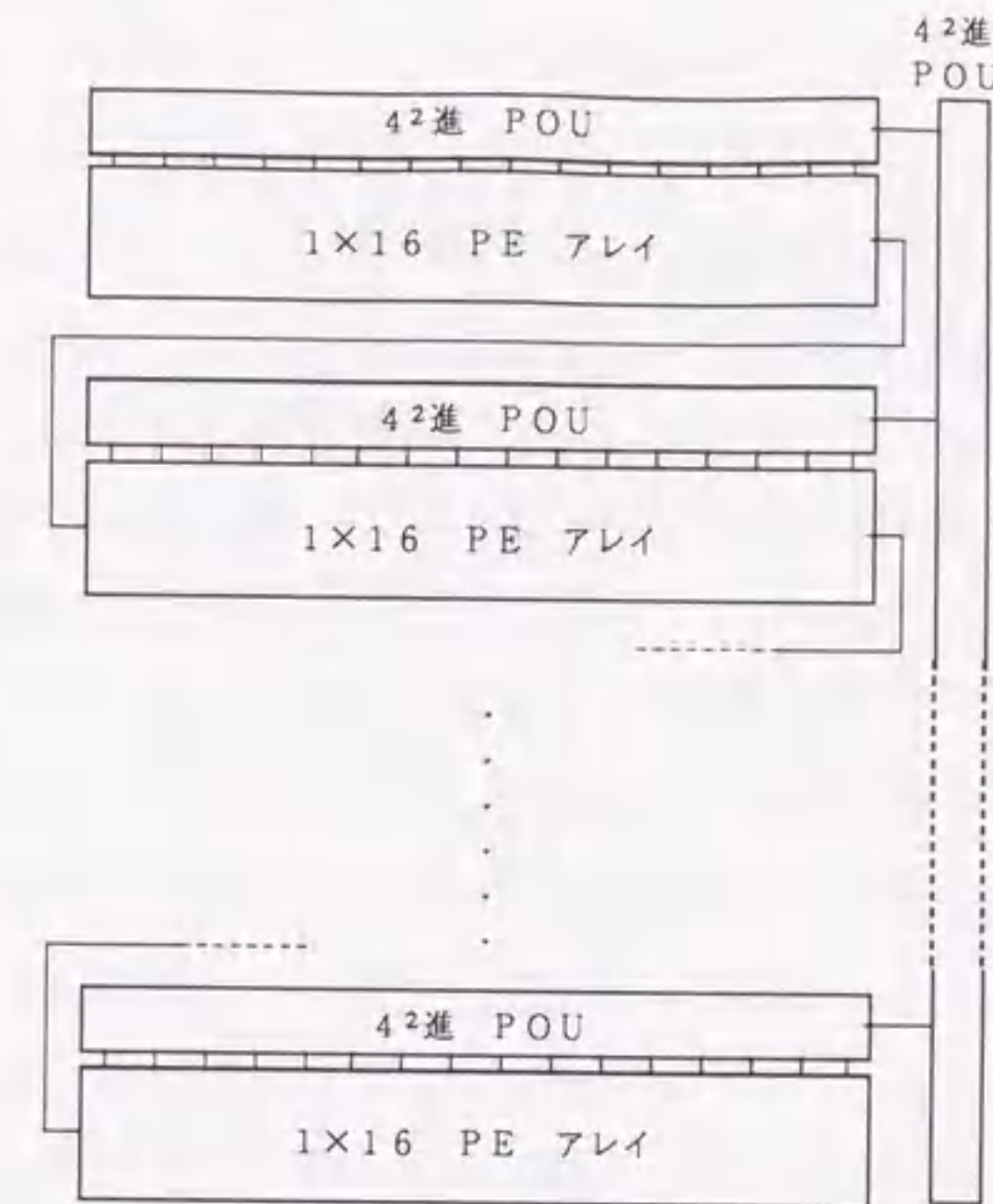


図2. 17 ツリー構成走査演算機構を備えた1次元PE配列のレイアウト構成

2. 7 むすび

隣接結合ネットワークの遠隔PE間の演算あるいは転送を補強する伝搬演算機構の構成方式を検討し、N進ツリー構成の高速の伝搬演算機構を提案した。

はじめに、基本的な伝搬演算方式として、

①演算結果をM個のPE間で直列に伝搬させる逐次伝搬方式
その高速化方式である

②N個のPEからなる局所ごとに並列に行う伝搬演算結果をバイパスを介して累算するバイパス方式

③演算単位を1ビットに限定し局所ごとに0,1の両方の入力に対する演算を先行的に行い、前段からの真の入力値により先行演算結果を選択するセレクトティブ方式のそれぞれについて、構成法を明らかにするとともに、伝搬演算時間を評価し、①の逐次伝搬方式ではM段分のPEの演算時間がそのまま積み重なるのに対し、②、③の高速化方式では、各々N段からなる局所ごとの演算時間と局所間のM/N段の演算時間との和にまで短縮されることを示した。

次いで、②、③のバイパス方式、セレクトティブ方式について、階層化による一層の高速化を試み、ともにN進ツリーの階層的構成になること、特にセレクトティブ方式では、

3個の2-1セレクタからなる伝搬演算エレメント(POE)のみでN進ツリーが構成されることを導いた。また、その伝搬演算時間についても評価し、ともに $N \log_2 M$ 段分にまで低減されること、セレクトティブ方式では演算がビット単位に制限されるものの途中セレクタのみを経由するために、演算器を経由するバイパス方式に比べ、さらに2倍近く高速になることを明らかにした。

N進ツリー構成については、ハードウェア規模の評価も行い、伝搬演算を1ビット単位で行なう場合には、4進のツリーにより、セレクトティブ方式が従来のバイパス2進ツリーに比べ、半分以下の実効的なハードウェア規模で実現されることを示した。また、LSIチップレイアウト構成も検討し、このセレクトティブ4進ツリー構成の伝搬演算機構が、稠密な構造で比較的容易にチップレイアウトできること、長さ256の配列データに対する伝搬演算を1ないし2クロックサイクルで実行できることを示した。

第3章 可変構造と演算アルゴリズム

3.1 はじめに

SIMDプロセッサは、定型的な処理を得意とするといわれている。しかし一般的には、定型的な処理といっても、常に同一サイズの配列データのみを処理すればよいわけではない。PE配列のサイズをはるかに越えるような大規模な配列データの処理から、PE配列のサイズをかなり下回る配列データの処理まで、ダイナミックに変化する場合が大半である。一般にPE配列のサイズを越えるような大規模な配列データの処理は、その配列をPE配列のサイズに収めるように分割し、それぞれを個別に処理することで容易に50%以上の高いPE配列の稼働率を得ることができる。これに対して、PE配列のサイズを大きく下回るような配列データの処理は、そのまま実行しようとする、PE配列の大半が遊ぶことになり、PE配列の稼働率はかなり低下することになる。これでは、せっかくの並列処理による高い性能が生かせない。このサイズの小さな配列データに対する処理は、別に設けるスカラ処理プロセッサによって実行することも考えられる。しかし、その配列データが一連の配列データの処理の途中結果である場合には、それを外部のスカラプロセッサにいちいち転送しなければならないのに加え、場合によっては、スカラプロセッサの処理結果を再度PE配列にもどすことが必要になる。従って、スカラプロセッサにまかせる方法では、並列処理による高速化が期待できなくなるばかりか、データ転送のオーバーヘッドも加わり、処理性能を大きく低下させることになる。

このような被処理データの構造のダイナミックな変化に対応するために、代表的なセルラ配列型SIMDプロセッサの1つとして知られているICLのDAP¹²⁾では、PE配列の各行を1,2,4,8,16,32のいずれかの長さのPE連結(以下PU

[Processing Unit: 処理単位]と呼ぶ)を、より高速な処理単位として動作させる機構を付加している。これによって、PE配列のサイズ(厳密にはPU配列のサイズ)が可変となり、小さな配列データを処理する際のPE配列の稼働率を改善している。

本章では、このようなセルラ配列型SIMDプロセッサにおけるPE配列の可変構造²⁹⁾を追究し、PE配列の一層の稼働率向上をはかる。また、この可変構造を利用することによって、汎用化に有効なPE配列内の矩形の局所領域ごとの総和演算、LSI-CADの論理シミュレーションの基本処理等が効率的に実行できることを示す。

3.2 修飾型SIMD制御による適応的な可変構造の実現

PE配列を可変構造とするには、所定の命令に対し、各々のPEの動作を個別に設定

できることが必要である。このPE毎の個別の動作制御の機能は、各PEに制御レジスタを設け、これによって制御部から受ける全PE共通の命令を個別に修飾すること（修飾型SIMD制御）で実現される。

本節では、PE毎の個別の動作制御により得られる可変構造の基本機能のうち、セルラ配列型SIMDプロセッサの性能向上に有効と考えられるものを示す。

3. 2. 1 PE配列への適切的なPU割付け

PUを構成するためのPE間の連結は、各PEを適当なサイズのビット幅で分割したレジスタ付きALUとして動作させることによって実現される。すなわち、算術演算における桁上げを、PE間で下位桁のPEから上位桁のPEに向かって伝搬させることによって実現される。ただし、この場合、最下位桁の(LSD)PEは、桁上げを下位側

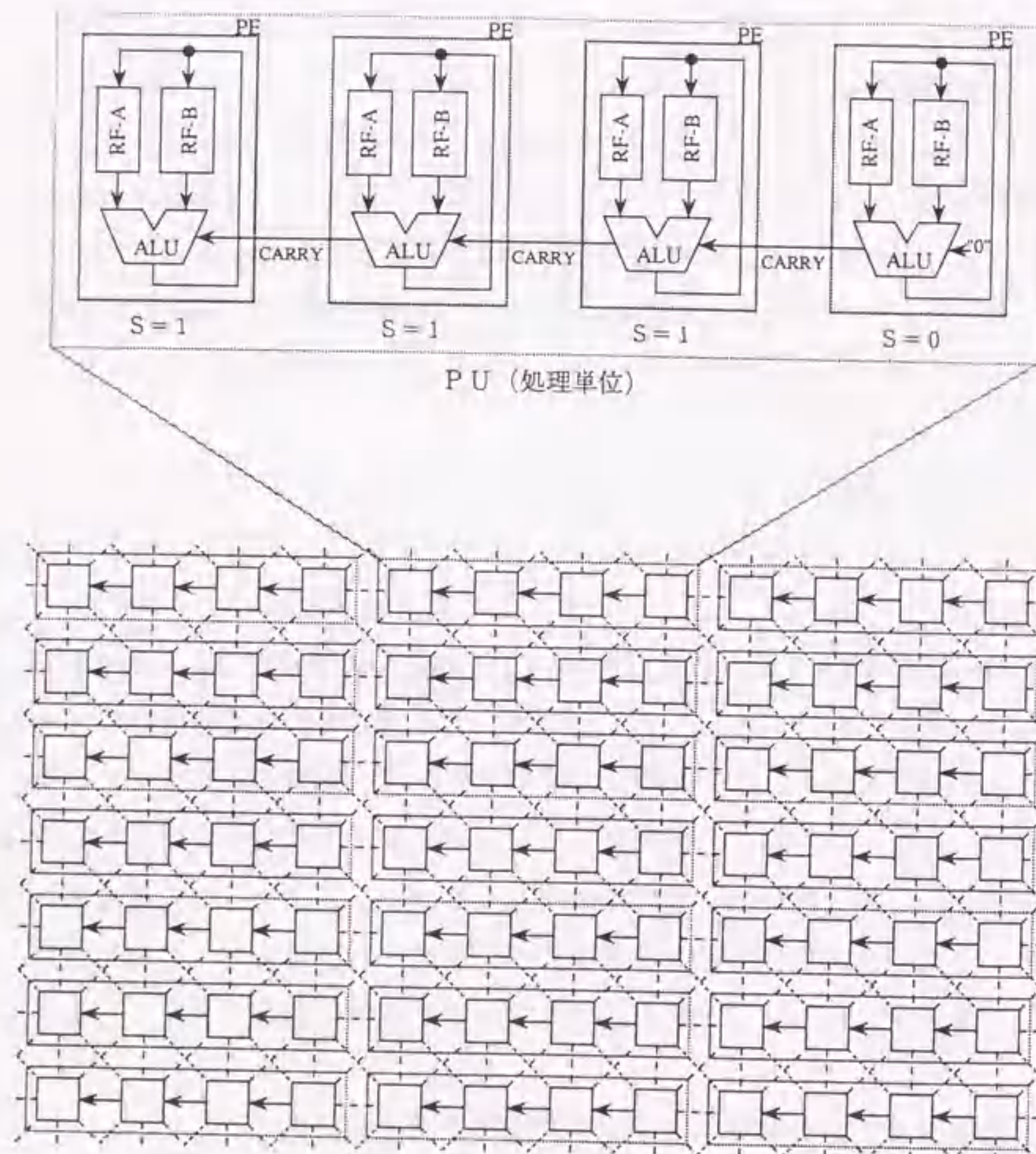


図3. 1 PUの構成

のPE（存在しない）から受け取るのではなく、自身で加算の場合には0、減算の場合には1の固定値桁上げを発生できる必要がある。

この場合、LSDかそうでないかのPE機能の違いをPE毎に個別に設ける制御レジスタで切り換えることによって、任意の位置に任意のビット幅のPUを構成することができる²⁶⁾（図3. 1参照。ここで、Sは制御レジスタの内容を示している。また前章の伝搬演算の場合とは矢印の方向をわざわざ逆に選んでいるのは、LSD側を右、MSD側を左側に配置して桁上げを左方向に伝搬させる加算器の記述慣習に従ったためである。）。セルラ配列型のSIMDプロセッサでは、PEの構成がもともとビットスライスのレジスタ付きALUに近い構成を採っているため、PE連結により無駄なくビット幅の大きいPUを構成することができる。しかし、桁上げをPE間で伝搬させるために、ビット幅の大きいPUを構成すると、演算時間がその分増加する。桁上げの伝搬経路を最適化してもクロック当たりPE通過段数で5,6段程度が限界と考えられ、これ以上のビット幅のPUを構成しようとする、処理語長と単位時間当たりの実行可能総演算量との積で表されるPE配列全体の性能は、PUのビット幅を大きくした分だけ低下することになる³⁰⁾。従って、処理データの語長が大きくとも、配列のサイズがPE配列のサイズと同程度かそれ以上の場合には、PE毎に多倍長の演算（PEが1ビット構成の場合にはビット直列演算）を繰り返すのが有利となる。

3. 2. 2 指定PEでの直並列変換

各PEは処理データを、通常、内蔵する演算系の演算単位に適合するように折り畳んで局所メモリ内に格納している。このため、特定のPEの保持データを演算幅の広いPUで処理するには、そのPEから、一旦、被処理データを取りだし、PU上に展開する必要がある。PEが1ビット構成で、かつデータがビット直列形式で格納されている場合には、PUで処理するデータをビット並列形式に変換することになることから、直並列変換と呼ぶ。

直並列変換は、図3. 2に示すように各PEの制御レジスタによる個別制御により送信PEとシフトPEを設定し、送信PEから取り出すビット直列データをPE間の隣接結合で形成されるデータ転送パスに乗せて順次シフトすることによって実現される。ここで、送信PEとは保持するビット直列データを順次シフトすることにより出力するPEであり、シフトPEとは、隣接PEから受け取ったビット直列データをそのまま転送方向にシフト出力するPEである。

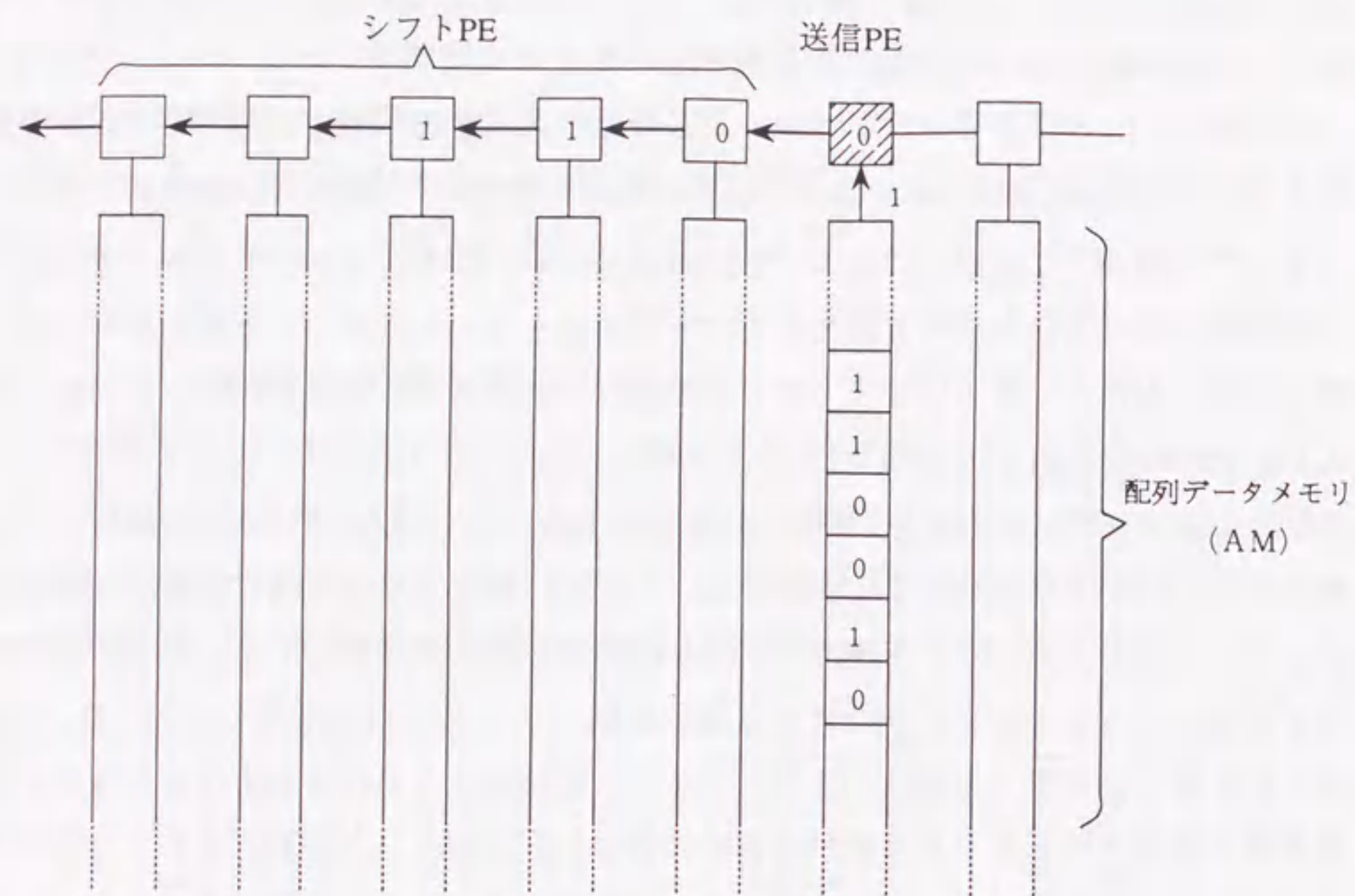


図3. 2 1次元PE配列上での直並列変換

3. 2. 3 指定PEでの転送パス間の乗り換え

並列に動作可能なPE間転送パスを複数設け、その間でデータを往来させることによって、配列データのPE配列上での90度回転やミラー反転を効率良く実行することができる。

図3. 3は、並列に動作する上から下方向と右から左方向に向かう転送パス間を左上角から右下角に向かう対角線上で乗り換えることによりミラー反転がかかった形の90度回転が実現されることを示している。この転送と乗り換えの制御は、やはり制御レジスタによる個別の制御で、対角線上のPEのみを乗り換えPEに、それ以外のPEをパスPEに設定することによって実現される。この方法によれば、上下の辺間と左右の辺間をそれぞれ結ぶ標準的なトラス接続だけで、N（配列の一边の長さ）回の上から下と右から左への同時シフトにより90度回転を実行することができる。もちろん、図3. 4に示すように上下の辺のいずれかと左右の辺のいずれかを結ぶパスを新たに設け、それを介して転送パス間の乗り換えを行えば、同じシフト回数で90度回転を実行することができる。しかし、この方法では、配列サイズが大きくなるに従って、プロセッサ配列の辺間を結ぶ結線の負担が重くなる欠点がある。

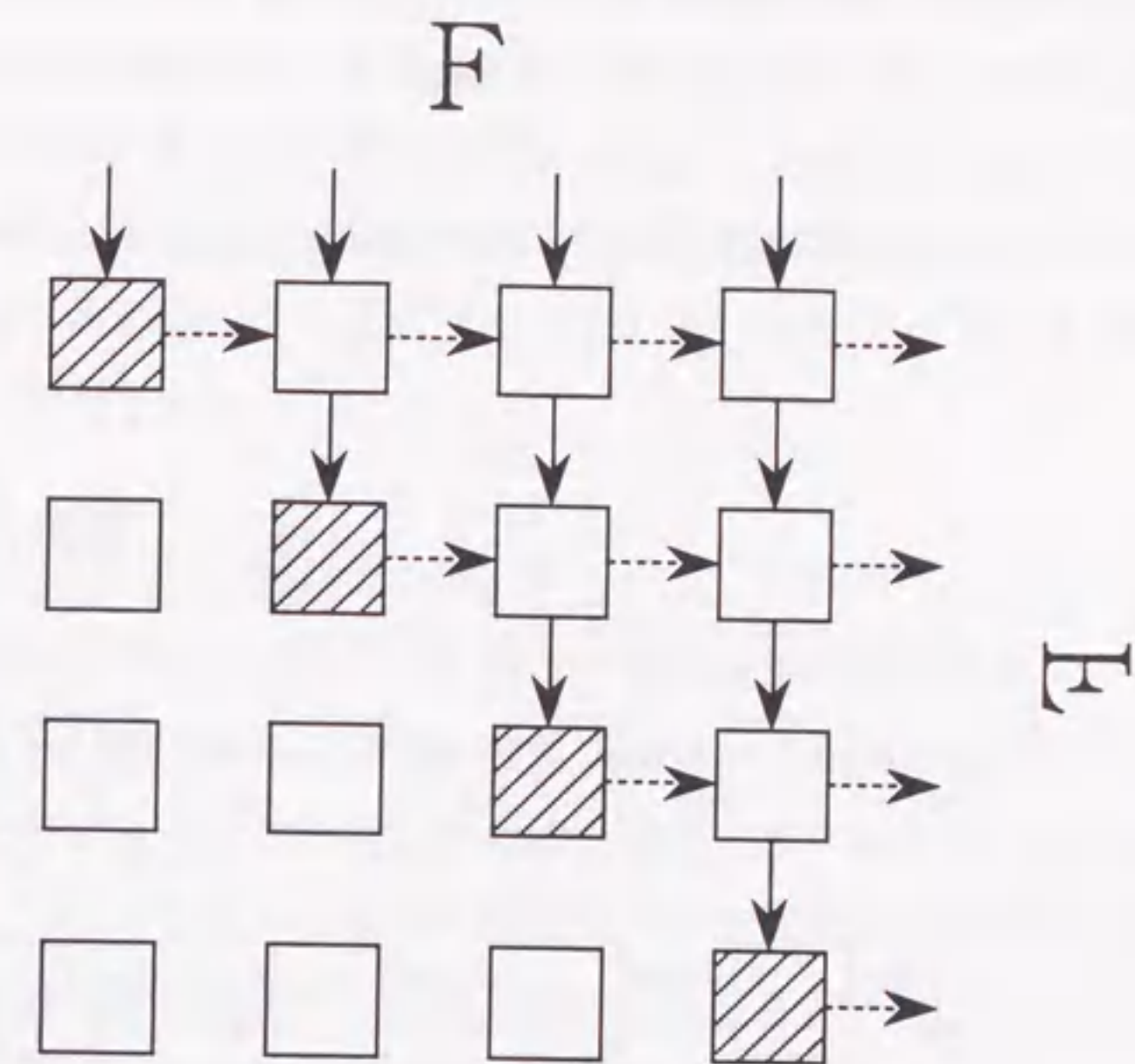


図3. 3 対角線上でのパス乗換え機能を利用した90度回転

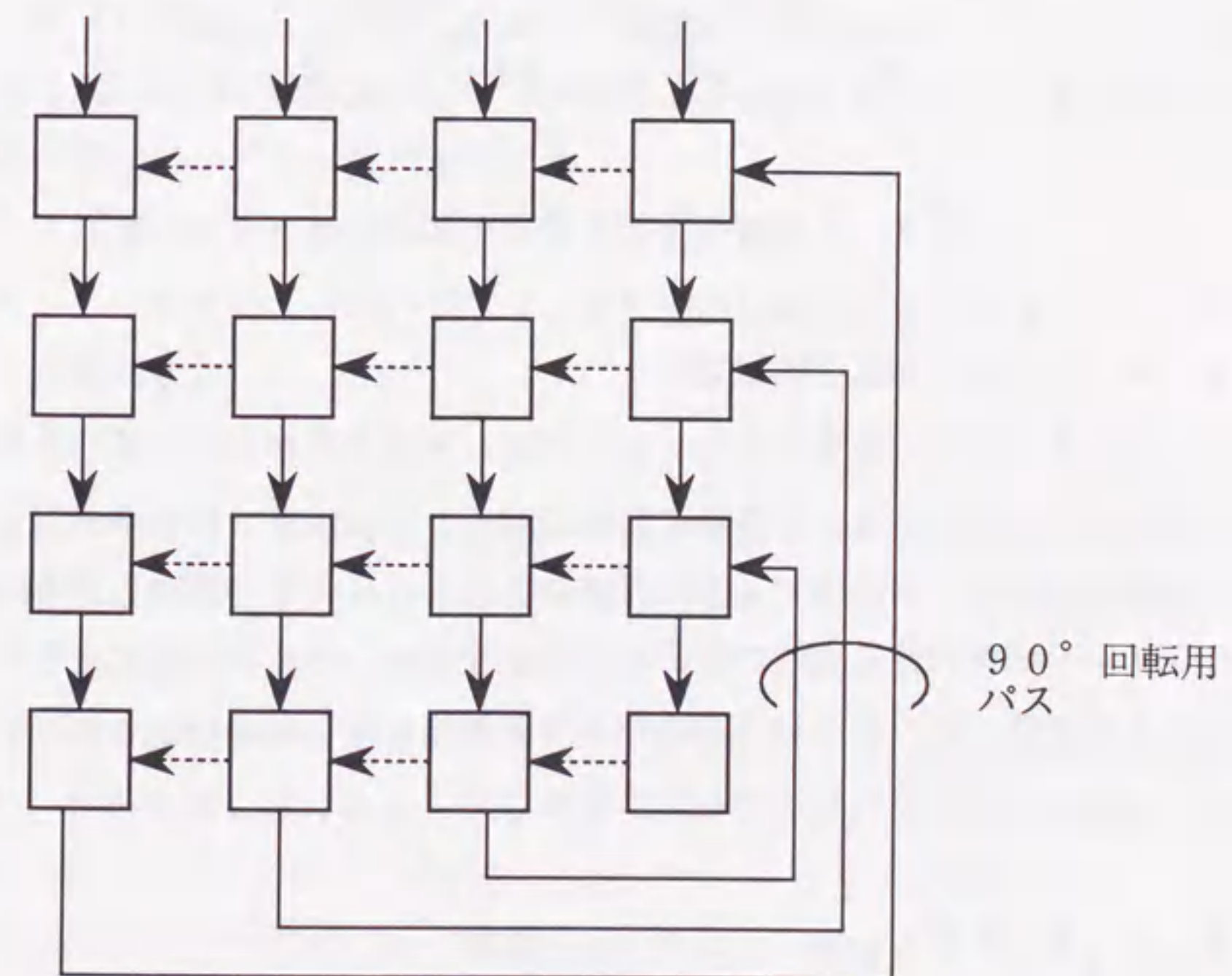


図3. 4 下辺と右辺を結ぶパスを用いる90度回転

図3. 5はパス間の乗り換え機能によるミラー反転のイメージを示している。左右方向のPE間結合を2重に設けて転送パスを2系統とし、それぞれの転送方向を逆に設定する。この状態で、配列の右端のPEのみ乗り換えPEに、それ以外をシフトPEに設定することによって、配列の右端でのパスの乗り換えが行なわれミラー反転が実現される。

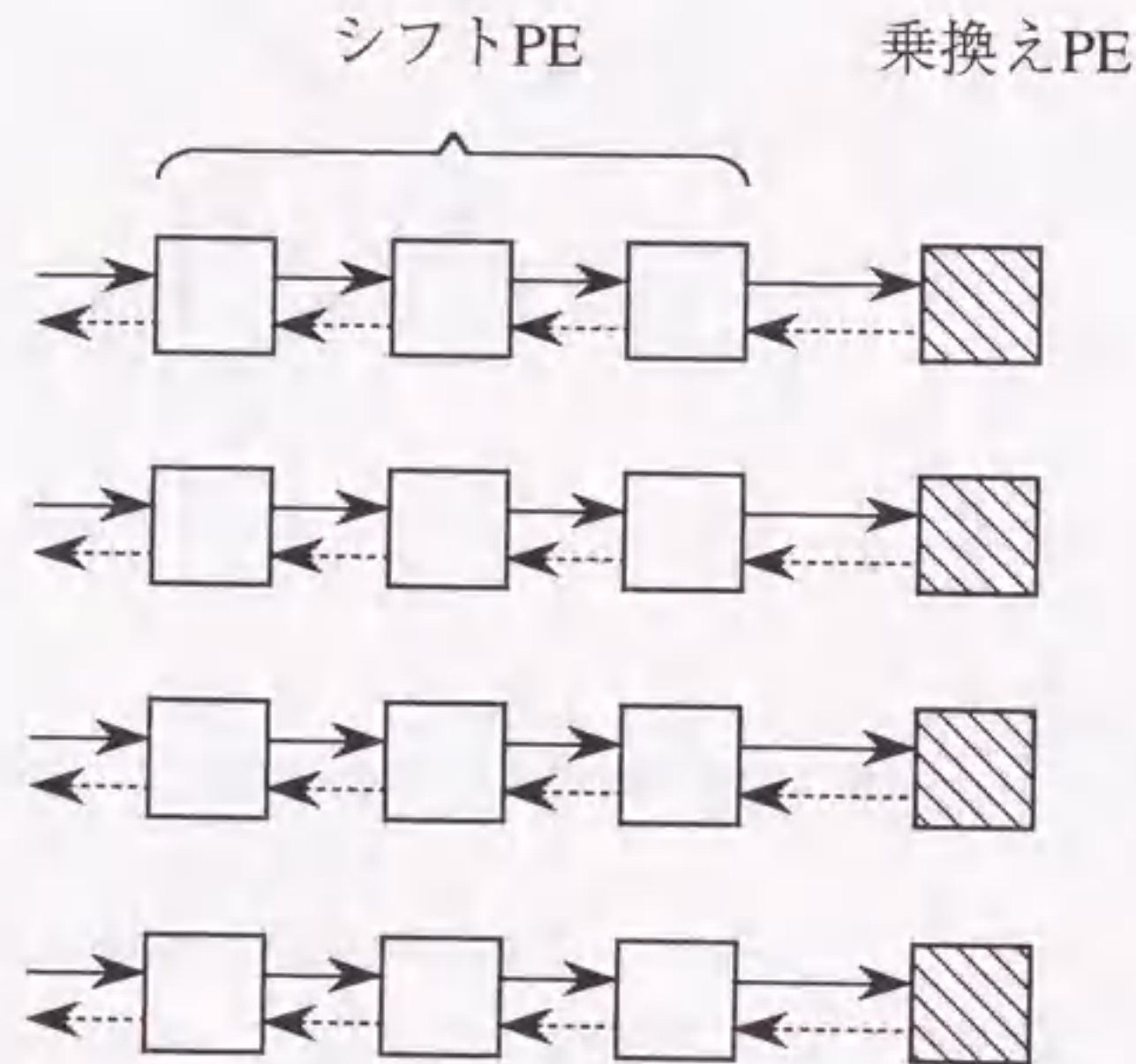


図3. 5 パス乗換え機能を利用したミラー反転

3. 2. 4 PE毎の転送方向設定

バイトニックソート、論理シミュレーション、シストリックアルゴリズムを前提とした線形計算等³¹⁾の応用では、PE毎に個別に転送方向を設定できるのが望ましい。この転送方向設定機能は、PE内で4方向の隣接PEからの入力を選択して内部に取り込む入力セレクタの選択制御を個別の制御レジスタで行うことによって実現できる。2系統の転送パスを設け、先の転送パス間の乗り換え機能と組み合わせれば、PE配列上でかなり自由度の高い転送経路の設定が可能になる。

3. 2. 5 PE毎のALUファンクション設定

論理シミュレーション、シストリック処理の実現には、局所的にPEの演算機能を変

更できることが必要である。特に論理シミュレーションにおいては、個別に演算機能を指定できれば、各PEで割付けられた論理ゲートを直接エミュレートできるようになるので、高速化に大きく寄与するものと考えられる。

この演算機能の個別の指定は、従来の全PE共通の制御信号線による一括制御とは別に、各PEに内蔵する制御レジスタによりALUのファンクションを個別に制御できるようにすることで実現される。

3. 3 可変構造を活用する効率的な演算法

3. 3. 1 伝搬加算を巧みに利用する矩形の局所領域の総和演算

前章では、1次元の配列データが複数の領域に区分けされた場合、その各領域に対する並列の総和演算が1次元のPE配列上の伝搬演算により効率良く実行されることを示した。これに対し、ここでは、2次元の配列データが複数の矩形領域に区分けされた場合、その矩形領域毎の総和演算が、2次元構成のPE配列上で、伝搬加算と可変構造によるPU割付けを巧みに利用する階層的な演算により効率良く実行されることを示す。

図3. 6はその階層的な総和演算の例である。1ビット構成のPEの16x16の部分配列内で、1PEに1要素(ビット直列形式)が割付けられる形で格納された同一形状の区分領域の総和をとる場合を示している。この概略の演算手順は、行方向にビット直列データに対する伝搬加算を実行したあと、直並列変換を行ない、得られたビット並列データに対して列方向の伝搬加算を行ない最終的な結果を得るというものである。以下、この演算手順について具体的に説明する。

ステップ1: 1回目のビット直列データに対する伝搬加算

図3. 6(a)に示すように各行でビット直列型の伝搬加算を行なう。

ステップ2: 直並列変換

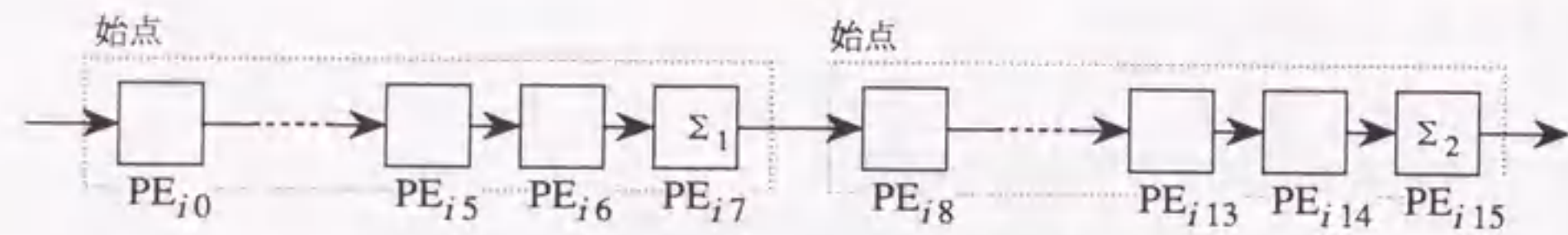
ステップ1で得られた各行の右端のPEに格納されている結果のビット直列データを、右端のPEのみ送信PEに設定し、MSDの側から順に送出し、その送出ビットデータを各行で左方向に順次シフトすることによって、ビット並列データに変換する。

ステップ3: ビット並列データに対する伝搬加算

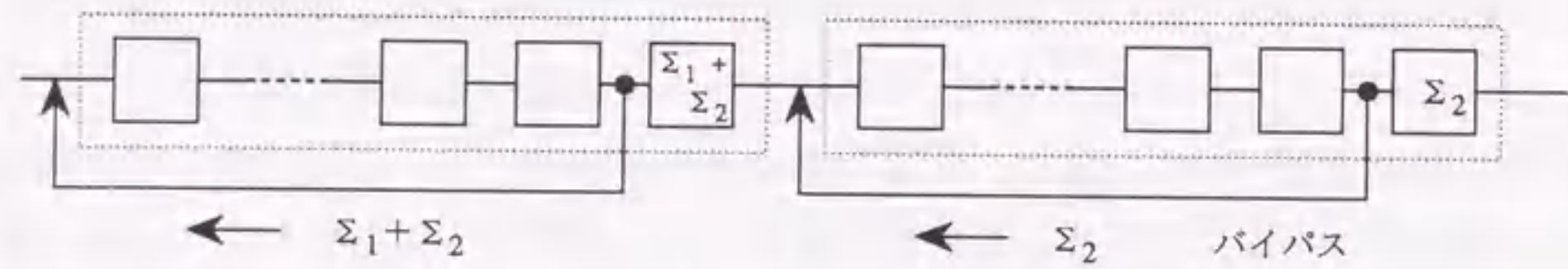
ステップ3で得られた各行のビット並列データ全体に対して、列方向の伝搬加算により総和をとり、行列要素全体の総和を得る(図3. 6(c))。ビット並列データに対する伝搬加算では、16x16のPE配列全体を一つのPUとして動作させる。各PEは左方向に桁上げを下方向に和を出力する全加算器として機能させ、16PEからなる各行を16ビット幅の加算器として動作させる。この結果、ビット並列データに対する

られる。ただし、この場合には、各行の最下位桁の指定の他に、上辺に伝搬の始点であることを指定することが必要で、これも各PEの制御レジスタに所定の値を設定することにより実現する。

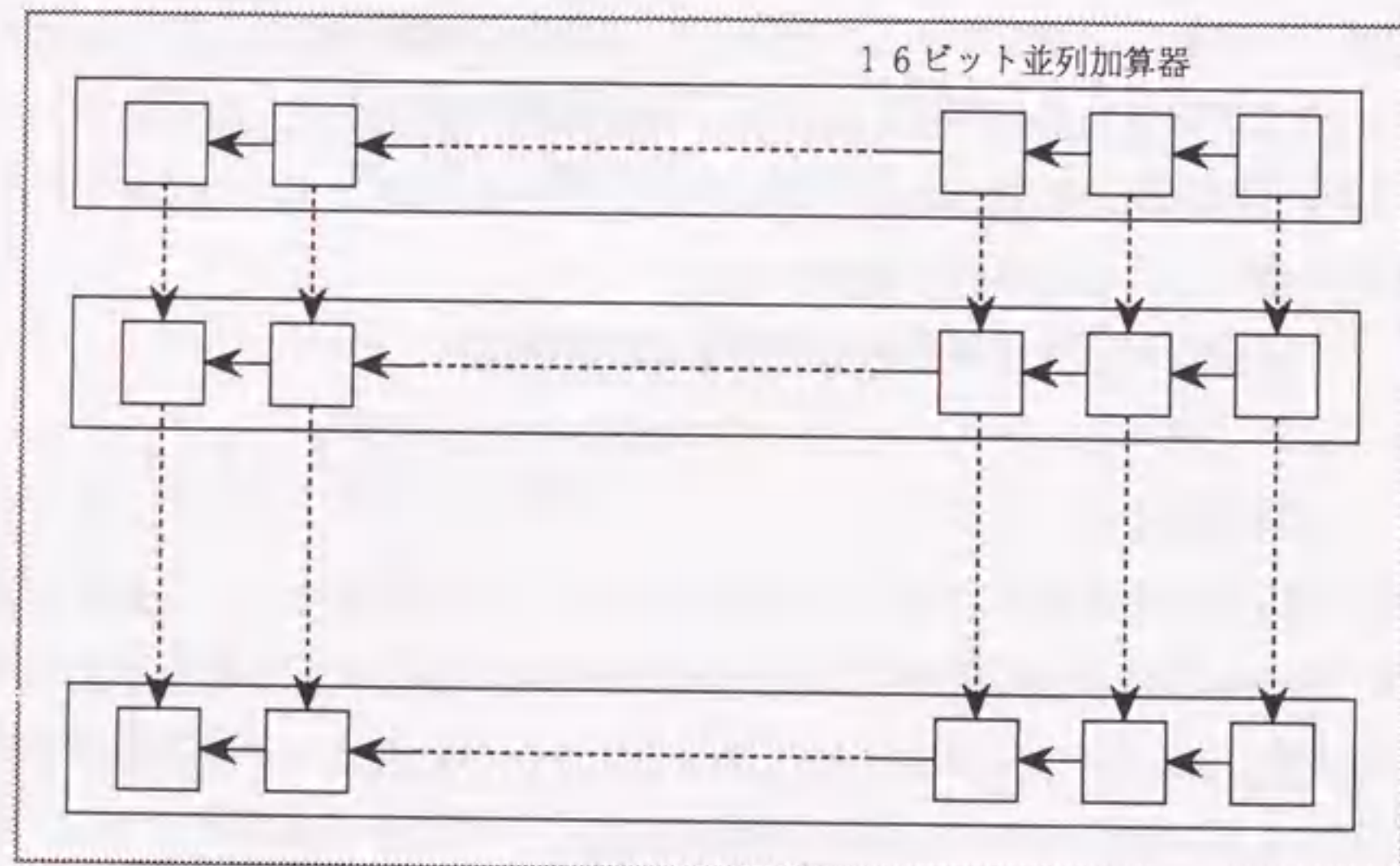
ステップ3でビット並列データに変換した後に伝搬加算を行なっているのは、ビット



(a) i行目における1回目のビットシリアル伝搬加算



(b) i行目におけるバイパスを用いたビットシリアル伝搬加算



(c) ビット並列データに対する下方向の伝搬加算

図3.6 矩形局所領域の総和演算法

直列データのままでビット単位で語長(W)の分だけ繰り返す行わなければならない列方向の伝搬加算が、パラレル形式で並列に実行することによって1回に低減されるからである。これにより短縮された演算時間とともとのビット直列の伝搬加算を繰り返す場合の演算時間との比率Rは、制御のオーバーヘッドを無視すると、

$$R = \frac{(t_{sp} + t_c) \times W + t_{pp}}{(2t_{sp} + t_c) \times W}$$

となる。ここで、右辺の分子側は可変構造を利用する総和演算の演算時間で、分母側はビット直列の伝搬演算のみで得た総和結果に直並列変換を行い、最終的なパラレル形式の総和結果を得るまでの演算時間を表している。また、 t_{sp} はビット直列データのビット単位の伝搬加算時間、 t_{pp} は、ビット並列データに対する列方向の伝搬加算時間、 t_c はビット単位の直並列変換時間である。この式から明らかなように、伝搬演算機構の構成が単純で、 16×16 のPE部分配列に対し、 $t_{sp}=4$ クロックサイクル、 $t_c=1$ クロックサイクル、 $t_{pp}=8$ クロックサイクル、 $W=16$ とすると、Rは0.61程度と、もともとの演算時間の2倍近くまで高速化されることがわかる。ただし、ツリー構成の高度の伝搬演算機構が組込まれている場合には、効果は小さくなる。例えば、セレクトティブツリー構成の伝搬演算機構により、ビット単位の伝搬加算が、 $t_{sp}=1$ クロックサイクルまで高速化されている場合には、Rは0.83となる。また、当然ながらPE間で直接ビット並列データに対する伝搬加算が実行できる機構が組込まれている場合には、効果はない。

3.3.2 論理回路シミュレーション

L S Iの集積規模の拡大にともなう設計上の切実な要求の一つに、論理シミュレーションの高速化がある。セルラ配列型S I M Dプロセッサはこの論理シミュレーションの基本処理が、可変構造を利用することによって、エミュレーションの形で高速に実行できる。図3.7は、エミュレートに必要な構成要素のみを記述したPEのブロック構成である。図3.8(a)はこのブロック図を用いてPE1個に論理ゲート1個を割り付けて2-1セレクタをエミュレートする場合の各論理ゲートのPE配列への割付け方を示している。図(b)はその論理ゲートを実現するためのPE配列の内部接続構成を示している。

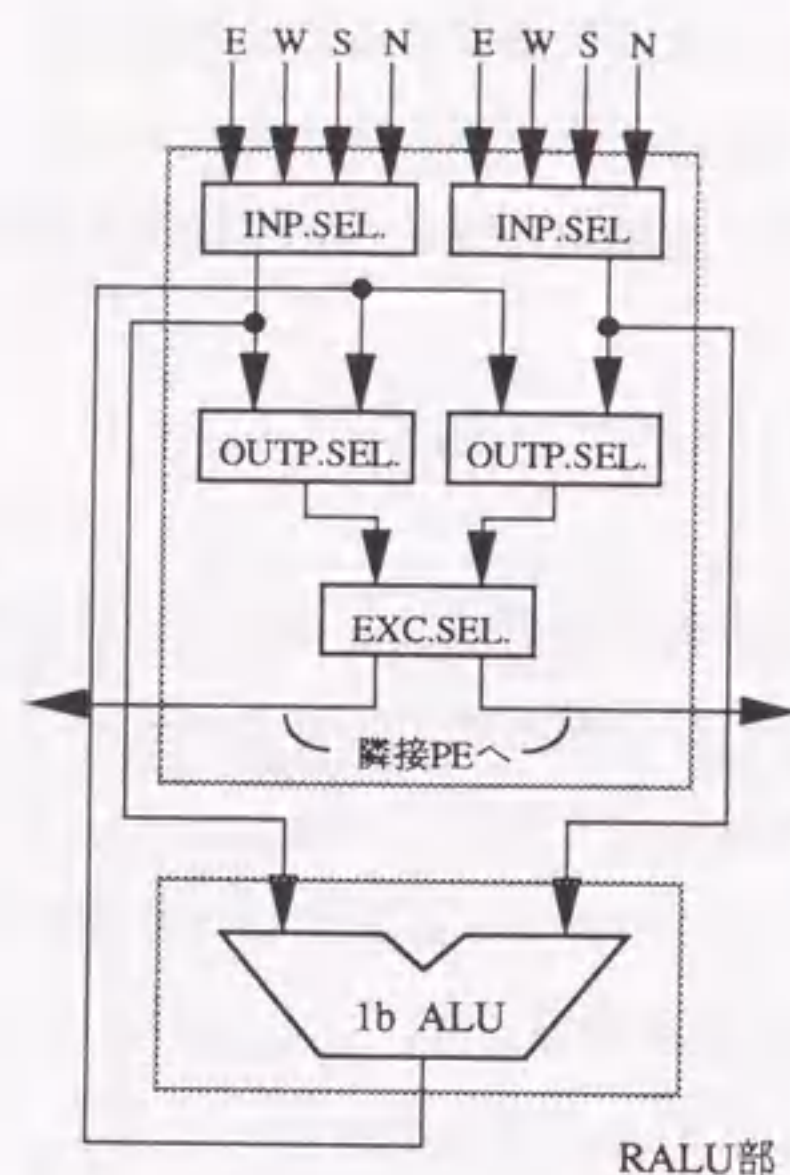


図3.7 PEのブロック構成

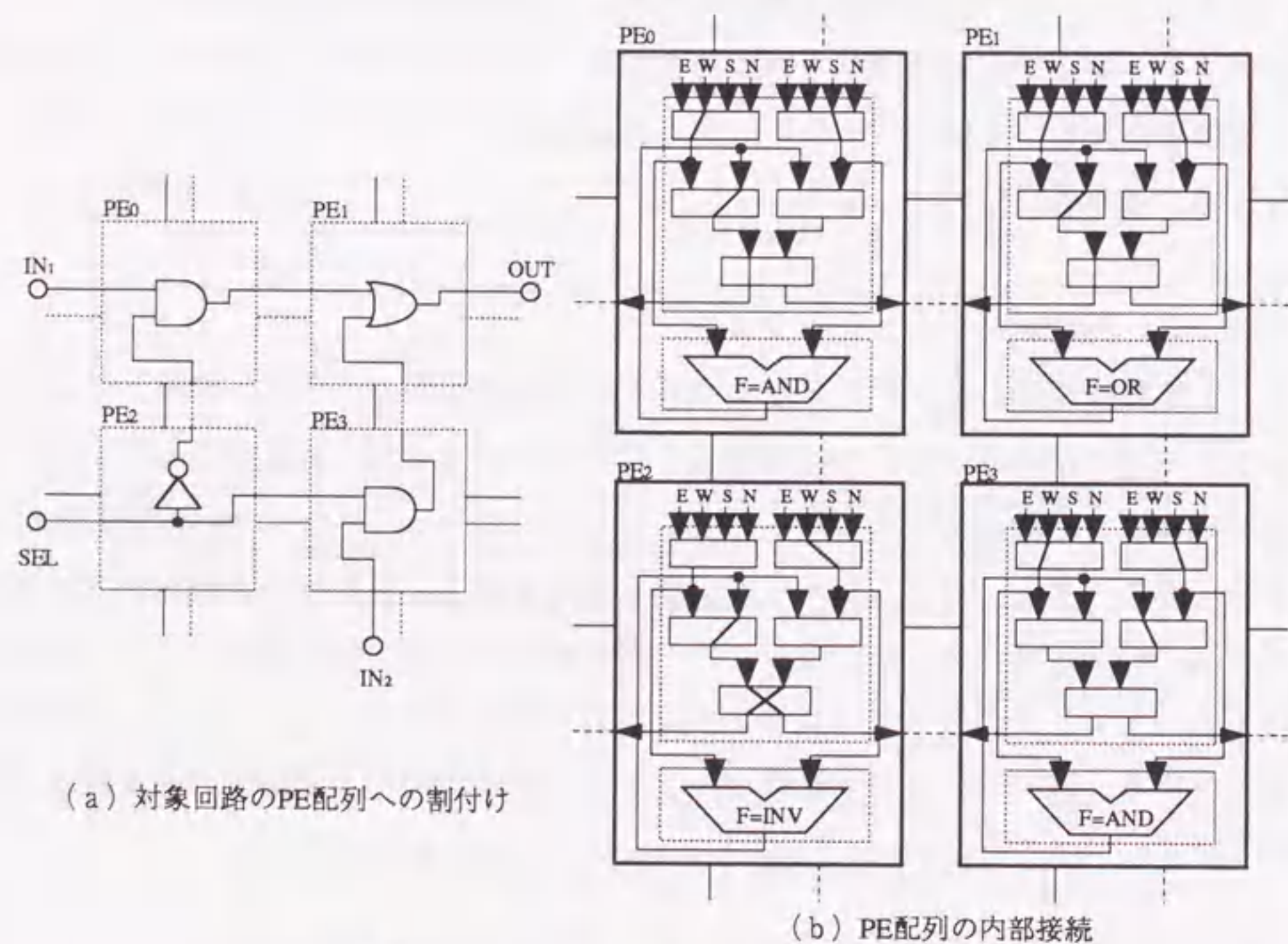


図3.8 論理回路のエミュレーション

ここで、ALUへの入力経路とPE間の信号経路となる転送パスの設定は、転送方向設定機能すなわち制御レジスタによりデータ転送ユニット内の入力セクタ (INP. SEL. [input selector]) を適当に選択することにより実現される。これを図ではセクタの箱の入力と出力を直線で結ぶことにより示している。また、PE2ではパス間の乗り換え機能 (EXC. SEL. [exchange selector]) により、左側のパス1から入力された信号をそのまま右側のパス2に出力している。

以上説明した動作から明らかなように、この演算は論理回路をソフトウェアでシミュレートするのではなく、ハードウェアで直接的にエミュレートする。従って、配列の規模を大きくして、配列上に実際のハードウェアと等価な論理回路を実現すれば、汎用コンピュータ上のソフトウェアに比べ、2桁から3桁以上の速度の高速化が可能となる³²⁾。

3.4 むすび

全PE共通の命令を、PE毎に付加する制御レジスタで個別に修飾することによって得られる可変構造的なPE配列の各種の機能について述べた。

はじめに、基本的な機能として

- ① PEを連結して並列の加算単位を構成するモードで、固定値桁上げ入力の生成と下位桁PEからの桁上げ入力を制御レジスタを用いてPE単位で個別に制御することによって、任意のビット幅の加算単位を任意の位置に構成できること、
- ② 送信PEとシフトPEを制御レジスタにより指定し、送信PEの保持するシリアルデータをシフトPEに送りだすことによって直並列変換が実現できること、
- ③ PE間シフト転送パスを2系統設け、その間の乗り換えの制御をPE単位で個別に制御することによって、90度回転、ミラー反転が実現できること、
- ④ PE間の転送方向、ALUファンクションを制御レジスタにより個別に設定できること

を示した。

次いで、この①、②の機能を巧みに組合わせて階層化する伝搬加算法により、PE配列内に設定する任意の矩形領域内のビット直列データ配列の総和が、階層化しない場合に比べ2倍近い速度で実行できることを示した。また、④の機能を用い1個のPEに1ゲートを割付け、ゲート間の結線を伝搬転送と③、④の機能を利用してエミュレートすることによって、論理シミュレーションを高速に実行できることを明らかにした。

第4章 データの転送、回転、およびメモリアクセス機構

4.1 はじめに

近年に商用化されたセルラ配列型SIMDプロセッサでは、ハードウェア量の大半を占めるプロセッシングエレメント (PE) 配列部を、1チップに16個から32個程度のPEを搭載したプロセッサ配列型のLSIで構成し、装置の小形化、経済化がはかられている。当然ながら、一層の性能/コスト向上には、このプロセッサ配列型LSIのチップ当たりの搭載PE数の拡大が望ましい。しかし、2次元あるいは3次元以上のPE間ネットワークとPE配列と一体化するLSI構成では、搭載PE数をこれ以上大きくすることは期待できない。これは、転送性能、メモリアクセス性能等を維持するのに必要となるPE間バス用、局所メモリアクセス用等に必要信号数が、パッケージの許容端子数を越えてしまうからである。

これに対し、1次元の接続構成のネットワークと一体化する構成では、面あるいは立体を単位とする高並列処理は困難になるものの、隣接PE間バス用に必要な端子が、搭載PE数によらず、左右の隣接LSIとの間を結ぶ2組分のバス用だけと少ない。従って、この分、搭載PE数が大きくでき、2次元構成のSIMDプロセッサに比べると、小形化、経済化の点で有利となる。この点が評価され、最近では、大容量メモリに128個のPEを一体化したLSIの提案³³⁾、64ないし1024個ものPEを1チップに搭載したLSIの開発^{34,35)}等が行われている。しかし、これらのLSIでは、汎用性よりも、搭載PE数の拡大を優先しており、

- ①間接アドレッシングによるデータアクセス、
- ②2次元アクセス (2次元配列データから、行単位の1次元配列データ [以下、単純に行と呼ぶ] あるいは列単位の1次元配列データ [以下、単純に列と呼ぶ] のいずれでもアクセスできる機能³⁶⁻³⁹⁾、
- ③遠隔PE間的高速転送

等の1次元SIMDプロセッサの汎用化に必須とも言える機能が、全く組込まれていないか、組込まれていても同一チップ内に閉じる機能に限られていた。

さらに、本論文の主な研究対象である2次元構成のSIMDプロセッサにおいても、①から③の機能の強化は性能向上に大きく寄与する。これは、8章以降の文字認識処理への応用の経験から、並列度の低い小型の2次元SIMDプロセッサでは、2次元の配列構成を活かした面単位の演算をベースに処理するよりも、PE間の転送バスを強化した1次元構成のSIMDプロセッサとみなし、各PEが同時に動作することによって実現されるライン単位の演算をベースに処理を進める方が、プログラムの記述性に優れる

ばかりか、処理の効率も高いことが明らかになってきたからである。

そこで、本章では①、②の機能を強化するために、はじめに従来1次元構成のSIMDプロセッサに用いられていたメモリアクセス機構の問題点を明らかにする。次いで、この問題を解決するメモリアクセス機構として、基本的な1次元SIMDプロセッサに組込まれるルーティングレジスタ (RTRG) 配列を、W個のPEからなる部分配列単位に、直並列変換型90度回転器としても、外部メモリ用のアドレス生成器としても利用する構成 (シフト転送と回転の機能を合せ持つことから転送回転型メモリアクセス機構と呼ぶ) を提案する。さらに、その性能、ハードウェア規模などを従来の行、列いずれのアクセスも1サイクルで実行する速度性能の点で理想的な構成の2次元アクセス機構 (以下、従来型2次元アクセス機構と呼ぶ。) ³⁹⁾と比較評価する。なお、③の問題を解決する構成については、2章で論じた伝搬演算機構が該当するので、ここでは触れない。

4.2 従来のメモリアクセス機構の問題点

4.2.1 メモリアクセス方式とその有用性

最も基本的なものは、各PEが各々の局所メモリを、全PE共通のアドレスで並列にアクセスするメモリアクセス方式である。いわゆるSIMDプロセッサ向きといわれる規則性の高い処理のかなりの部分がこのメモリアクセス方式のみで実行できる。このため、このメモリアクセス方式のみをサポートするSIMDプロセッサも数多く開発されている¹¹⁻¹⁵⁾。

しかし、このメモリアクセス方式だけでは、一見SIMD向きと思える処理でも、PEごとでの関数計算のようにうまく処理できない場合がある。その例としては、ニューラルネットワークシミュレーションにおけるシグモイド関数、三角測量処理における三角関数、画像処理におけるヒストグラム解析等があげられる。この種の計算は、各PEが局所メモリに格納される配列データの異なるアドレスを並列にアクセスすることで効率良く処理できるが、そのためには、各々のPEで異なるアドレスを生成するための間接アドレッシングによるデータアクセス機能がPEごとに必要になる。

さらに、1次元SIMDプロセッサには、非常に有用性の高いメモリアクセス方式として、2次元配列データに対し、行、列のいずれに対してもアクセス可能とする方式 (2次元アクセス方式) がある。行あるいは列の一方のみの1次元配列データのアクセスでは、効率良く処理できない場合が少なくないからである。例えば、斜交軸変換による画像のアフィン変換⁴⁰⁾、行、列両方向の密度分布を求める文字の特徴抽出処理 (8章

参照), 行, 列両方向の1次元FFTを繰り返す2次元FFT³⁹⁾等では, 一連の行単位の処理に続き, やはり一連の列単位の処理を行なう必要がある。

4. 2. 2 実現上の問題点

より多くのPEを1チップに搭載しようとする, 集積度の制約から各PEの局所メモリは, 同一チップ内に全量を搭載するわけには行かない。この場合, チップ内のPEから外部に置く局所メモリを高速にアクセスする必要があり, そのためには多数のデータアクセス用の端子を設けねばならない。また, PEの機能の向上をはかろうとすると, PE配列部の複数のLSIによる分割構成が避けられず, そのために隣接LSI間の接続に用いる信号端子も必要になる。これらの端子はもちろん, チップ全体の端子数の制約から無制限に設けるわけには行かない。従って, 外部の局所メモリに対して, 分割単位のLSIごとに, より少ない端子数で各種メモリアccessを効率良く行う仕掛けが必要になる。

実際, 間接アドレッシングによるデータアクセスについては, 各PEで完全に並列に行なう代りに, 同一LSI内のPE同士が, 一組の間接アドレス生成機構を時分割的に使用することによって, LSIの端子数とハードウェア規模の低減をはかる方法がある¹⁷⁾。

しかし, 2次元アクセスについては古くから検討されてはいるものの, ハードウェア規模低減の観点から理想的といえる方式, すなわちPE部分配列を内蔵する同一構成の大規模LSIとそのPEごとの局所メモリの配列(メモリユニット)だけで分割単位を構成するような方式は見当たらない。例えば, 従来型2次元アクセス機構³⁶⁻³⁹⁾では, PE部分配列の1チップ化は現在の大規模LSIを用いたとしても, 端子数の制限から実現は容易でない(表4. 1参照)。これは, PEごとに個別のアドレス端子を必要としたり, バレルシフタ, フリップネットワークのような分割構成時の相互の結線数の大きい高速ルーティングネットワークが必要となるからである。

4. 3 提案する転送回転型メモリアccess機構の構成

転送回転型メモリアccess機構は, 基本的な構成のSIMDプロセッサがもともと備えている隣接PE間の転送パスを活用することによって, 各種のメモリアccessを効率的に実行することをめざす。従来型2次元アクセス機構のように高度なネットワークは用いないので, 端子数, ハードウェア規模の増加は小さく, 基本的な1次元SIMDプロセッサの備えるLSI化時の搭載PE数の拡大容易性はほとんど低下しない。

転送回転型メモリアccess機構を組込んだ1次元セルラ配列型SIMDプロセッサの

並列演算部の構成を図4. 1に示す。この並列演算部は, 一点鎖線で囲まれるPE部分配列とRTRG部分配列にメモリユニット(局所メモリの部分配列)を外付けして1つの単位とし, これを規則的に並べることで構成される。PE部分配列とRTRG部分配列は, 図4. 2に示すように, それぞれW個のPEとRTRGで構成され, 幅Wのバスでメモリユニットと接続されている。ここで, RTRG部分配列は, PE部分配列の各PEが縦続接続されることによってその内部に形成されるRTRGの配列を, 抜き出して示したものである。従って, 図4. 2に示すようにレジスタ要素(RE)のW×W個の行列状の2次元配列が, 左右の隣接レジスタ要素間のバスと列ごとのバスで結合される構成となっている。

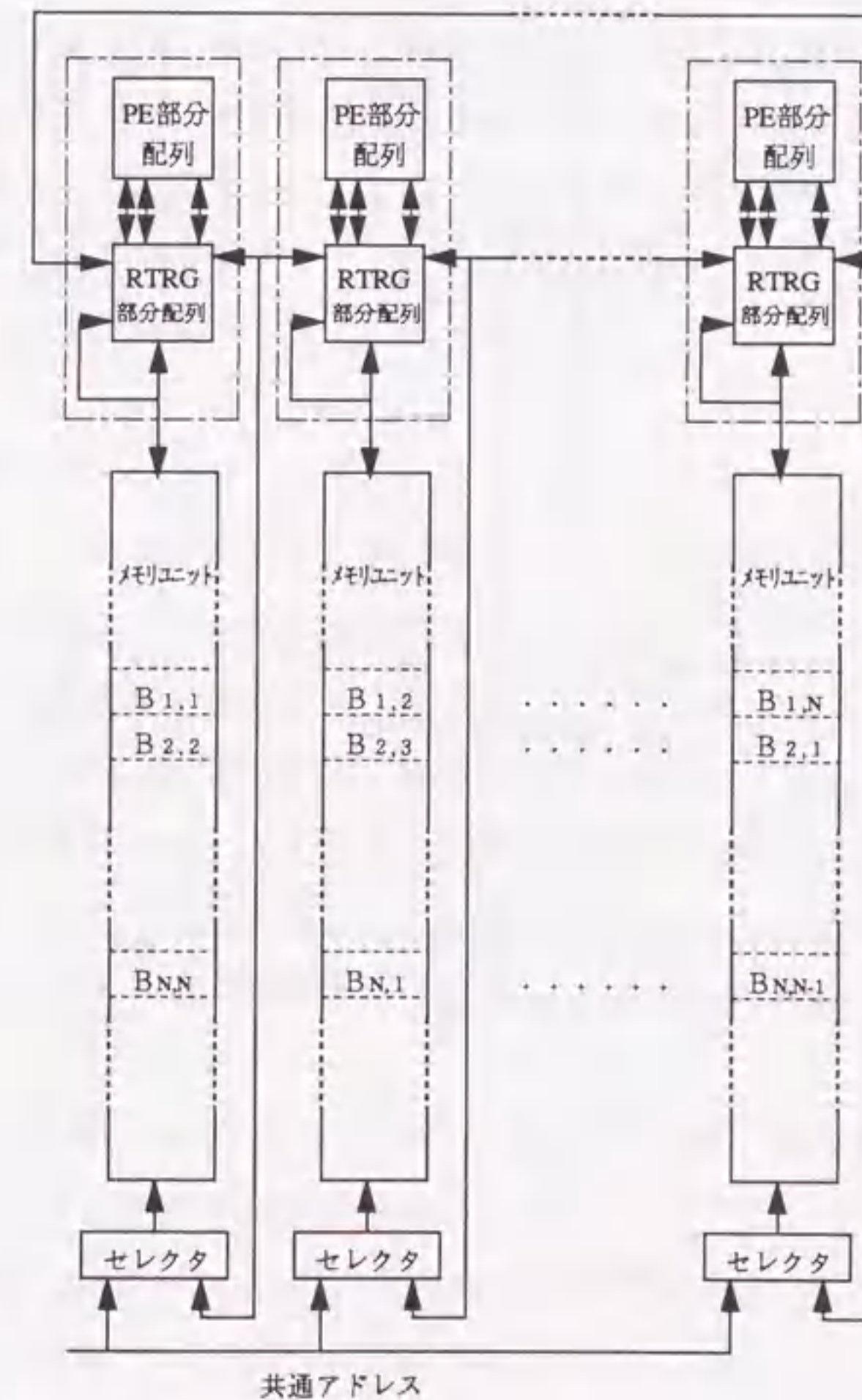


図4. 1 転送回転型メモリアccess機構を組込んだ1次元SIMDプロセッサのPE配列部の構成

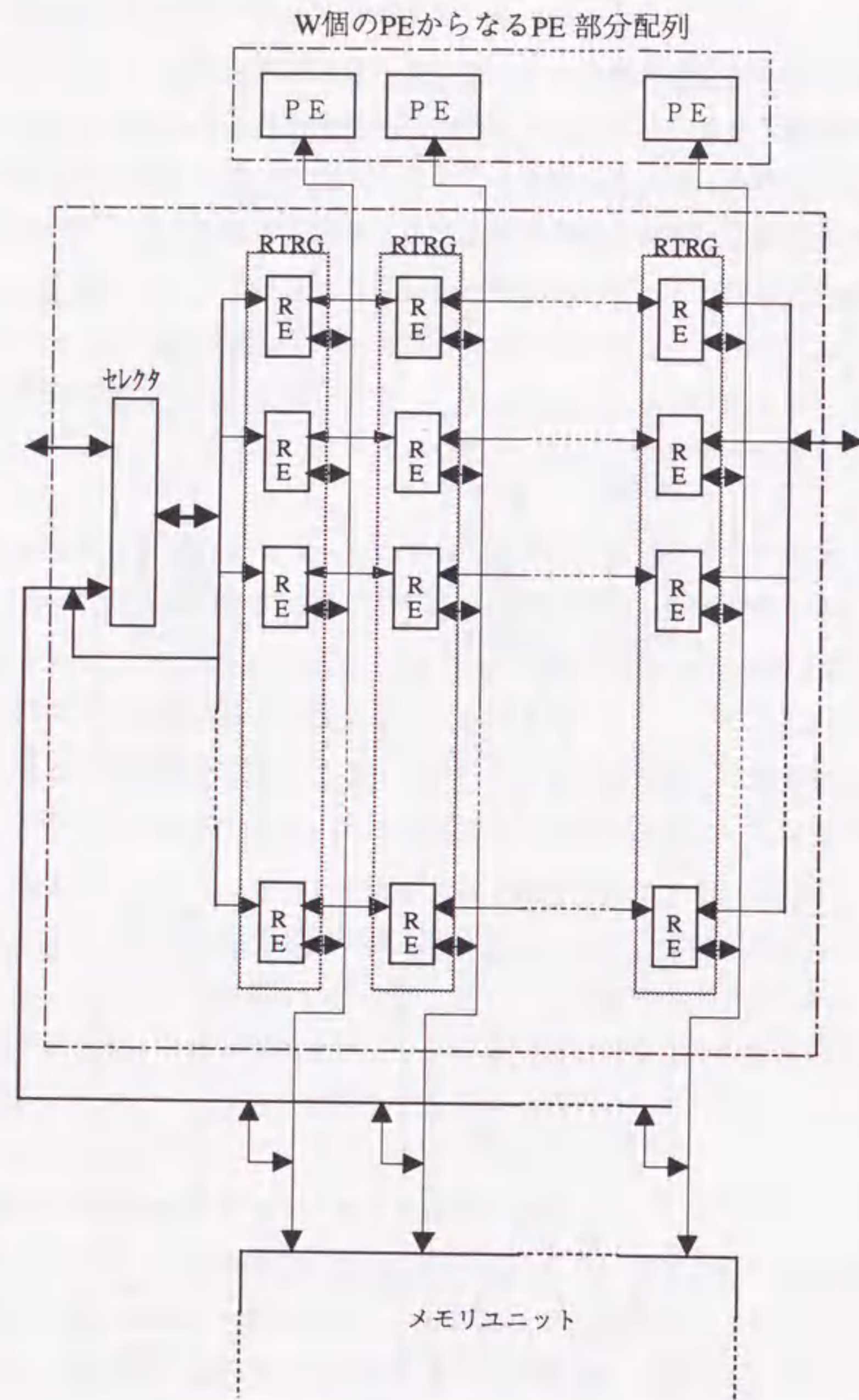


図4.2 ルーティングレジスタ (RTRG) 部分配列の構成

転送回転型メモリアクセス機構は、メモリユニットに対して、

- ① RTRG部分配列右端からの出力を下側のセレクタを介して、メモリユニットのアドレスとして加えられるようにしたこと (図4.1)、
- ② RTRG部分配列の列ごとのバスを介する上下方向の入出力バスを設けると共に、RTRG部分配列の2次元配列の左端にセレクタを介した左右方向の入出力バスを設けたこと (図4.2)

に特徴がある。これらによって、部分配列ごとに、RTRG部分配列からの出力によるメモリユニットのアドレッシング、メモリユニット内の $W \times W$ のブロック単位の配列データに対する90度回転等が可能となる。

4.4 メモリアクセス法

転送回転型メモリアクセス機構では、並列演算部のRTRG間の本来のデータ転送機能に、RTRG部分配列ごとの回転機能、メモリユニットに対するアドレッシング機能を組み合わせることによって、間接アドレッシングによるデータアクセスと、従来型、疑似型、間接型の3タイプの2次元アクセス方式が実現される。以下、これらのアクセス法について所要サイクル数の算出も含めて説明する。

4.4.1 間接アドレッシングによるデータアクセス

間接アドレッシングによるデータアクセスは、テーブル参照だけでなく、次節以降の各種2次元アクセスを実現するための基本機能として重要である⁴¹⁾。転送回転型メモリアクセス機構では、この間接アドレッシングによるデータアクセスがRTRG部分配列の出力によるメモリユニットに対するアドレッシング機能とブロック単位の配列データに対する90度回転機能を組み合わせることによって実現される。具体的には、図4.3に示すように、以下の①、②を並行して行うことで実現される。

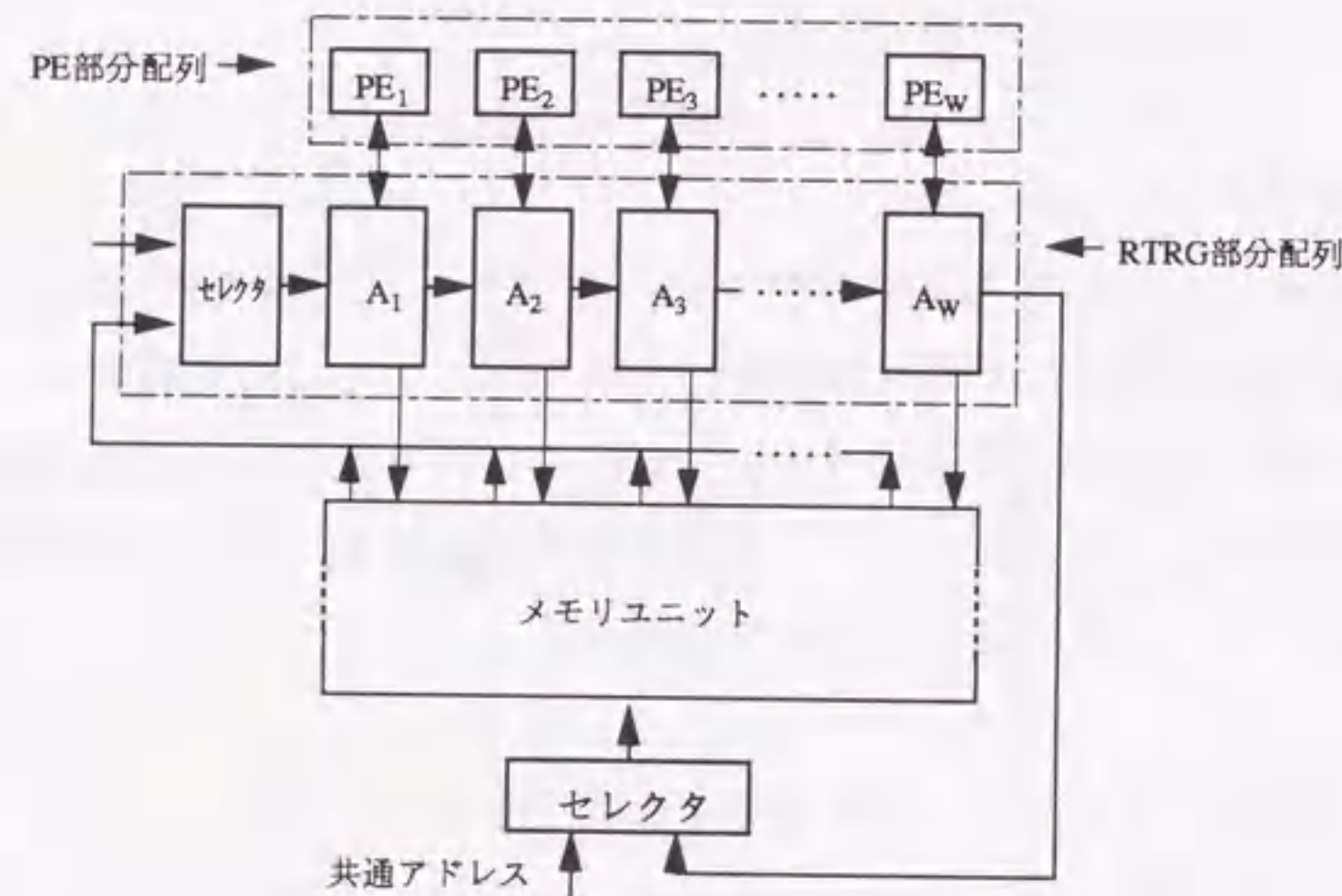
- ① シフト転送によりRTRG部分配列から出力される A_w, A_{w-1}, \dots, A_1 でメモリユニットをアドレッシングする。
- ② ①のアクセスデータ列 $[A_w], [A_{w-1}], \dots, [A_1]$ を、左端からシフト転送により90度回転する形でRTRG部分配列に戻す。

ここで、 $[A_i]$ はメモリユニットの A_i 番地のワードデータであること、すなわちこの間接アドレッシングによるデータアクセスがメモリユニットのワードデータの配列に対するものであることに留意されたい。

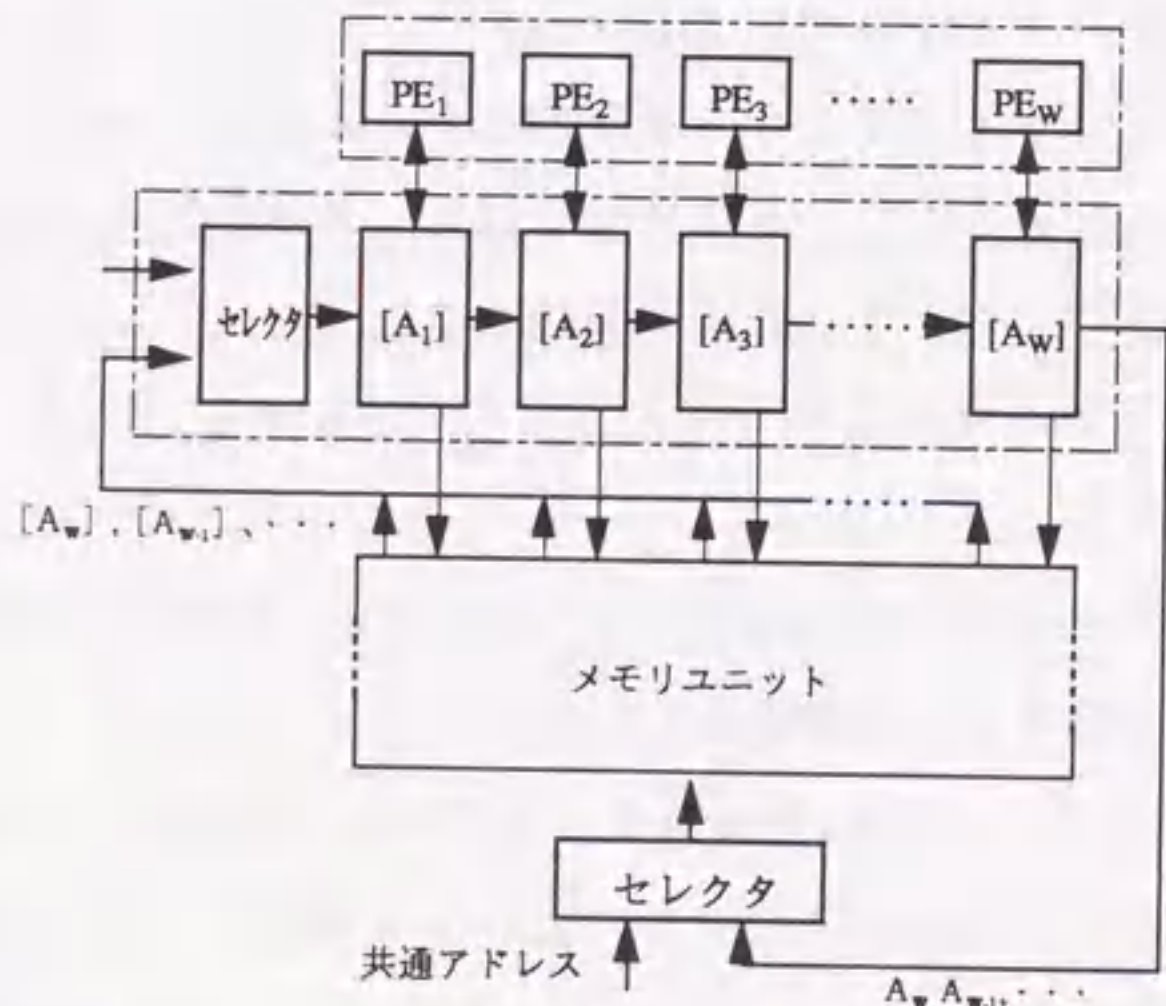
この間接アドレッシングによるデータアクセス方式には、ハードウェア規模を低減する

上で、2つの利点がある。1つはテーブル参照処理において、テーブルの格納領域を $1/W$ にまで低減できることである。これは、間接アドレッシングによるデータアクセスの対象となるワード単位のデータ配列が、PE部分配列単位で共有されるからである。もう1つは、従来方式⁽¹⁷⁾で時分割使用により間接アドレッシングによるデータアクセスを実現していたPE部分配列単位のデータバスとアドレスバスが不要になることである。

ところで、このアクセス法では、 A_1, A_2, \dots, A_w のアドレス値の算出は、



(a) 読み出し前の状態



(b) 読み出し後の状態

図4.3 転送回転型メモリアクセス機構での間接アドレッシングによるデータアクセス

各PEの並列処理により効率良く計算される。しかし、メモリユニットからのアクセスは1ワード(1PE分)単位と逐次的であり、この分についてのアクセス速度の向上はない。ただし、セルラ配列構成のチップのように、搭載PE数が大きく、端子数が制限されて、アクセス速度が端子の稼働率だけで決まる場合には、他の並列化可能な方式を採用したとしても意味はない。すでに、この方式でメモリユニットアクセス用の端子を、アドレス供給とデータのアクセスのために100%稼働させており、並列化の効果を活かせないからである。

4.4.2 従来型2次元アクセス

従来型2次元アクセスは、従来型の2次元アクセス機構で用いているアクセス方法である³⁶⁻³⁹⁾。転送回転型メモリアクセス機構では、このアクセス法をRTRG配列によるPE間シフト転送機能に前節の間接アドレッシングによるデータアクセス機能を組み合わせることによって実現できる。

ただし、アクセス対象は、行ごとに要素単位でずれる形式(スキュード配列 [Skewed Array] 形式)³⁷⁾の2次元配列データであることが必要となる。このため、メモリユニットの配列からアクセスされる各行がPE配列側でねじれた形式となるように、メモリユニットに対してPE部分配列ごとに回転をかけながら書込みあるいは読み出しを行う⁴²⁾。これによって、各メモリユニットに行の1要素が1ワードに入る形式で格納され、間接アドレッシングによるデータアクセス機能により、任意の列を行同様に効率良くアクセスできるようになるからである。

このアクセス法での1行および1列のライン当たりの所要サイクル数 C_{rc} 、 C_{cc} は、並列演算部の分割数を N 、間接アドレッシングによるデータアクセスのアドレス生成の所要サイクル数を A 、サイクル当たりのシフト量を S とすると、要素の語長が W に等しい場合で、

$$C_{rc} = W + (W \times N) / (4S) \quad (4.1)$$

$$C_{cc} = W + (W \times N) / (4S) + A \quad (4.2)$$

となる。これらの式で、右辺第1項はメモリユニットからねじれた形式の行あるいは列をそのままアクセスするのに必要なサイクル数である。また、右辺第2項は、そのねじれをPE間のシフト転送機能を利用して補正するのに必要なサイクル数であり、 N が偶数で逆方向のシフト転送も利用する条件でのずれ補正のための平均シフト量

$(W \times N) / 4$ に、サイクル当たりのシフト量 S を考慮して求められる。

4. 4. 3 疑似2次元アクセス

本アクセス方法は、従来型2次元アクセスのように行あるいは列を直接1本ずつアクセスするのではない。一旦、行あるいは列方向のビットライン（要素の語長が1ビットの1次元配列） W 本からなる横長あるいは縦長の短冊状配列データをアクセスし、これを介して、行、列両方のアクセスを実現する。この疑似的な2次元アクセス（以下、疑似2次元アクセスと呼ぶ。）における短冊状2次元配列データのアクセスは、ブロック単位のスキュード配列形式配列データ（図4. 1のメモリユニット配列に格納される $B_{1,1} \sim B_{N,N}$ のブロック配列データ参照）に対し、ブロックを単位として従来と同様な2次元アクセスを行うことによって実現される。ここで、ブロックとは各メモリユニットに格納される幅 W のワードデータの W 本の集りである。

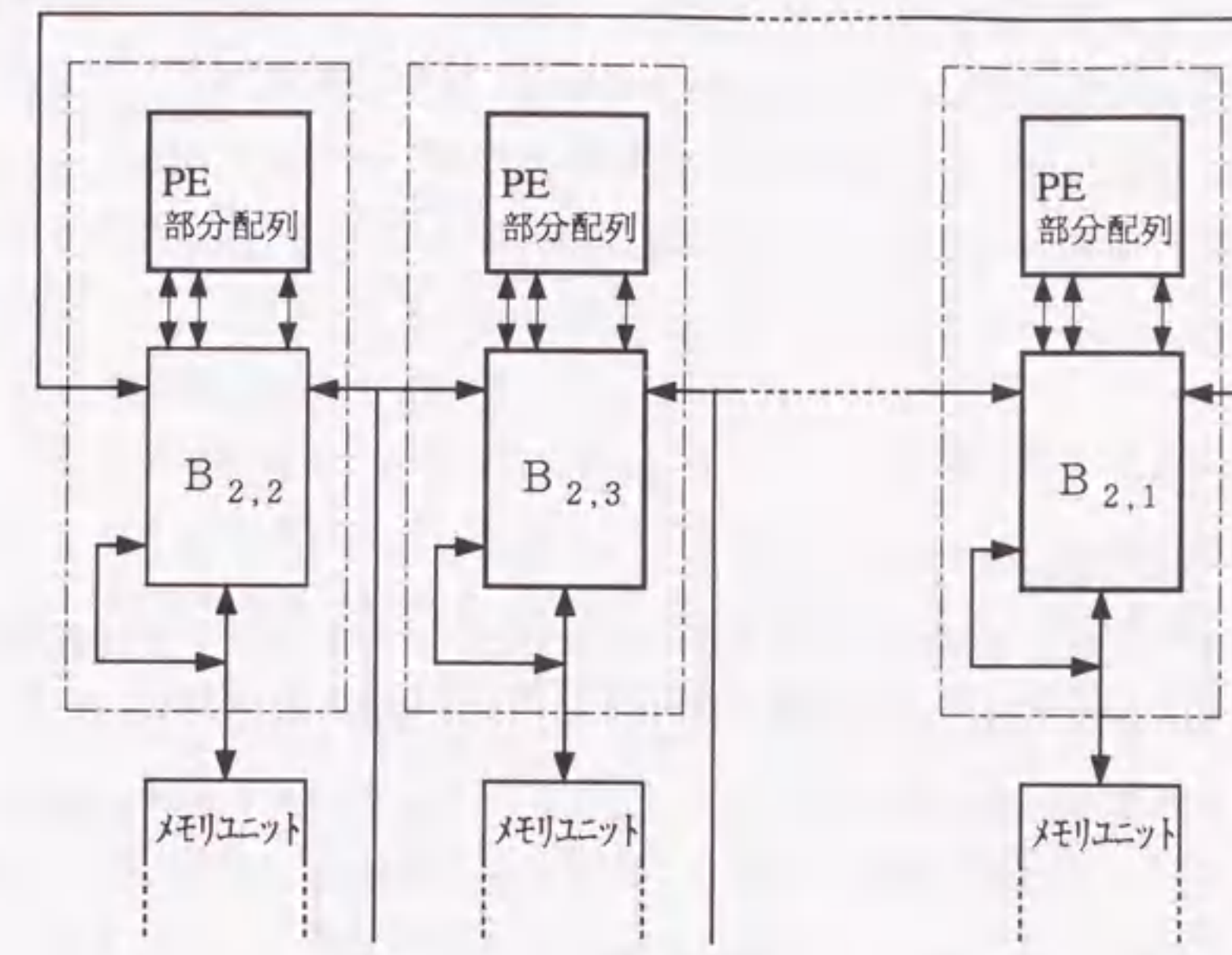
以下、図4. 1の配列データに対する短冊状2次元配列データのアクセスについて具体的に説明する。はじめに、この配列データに対し、各メモリユニット共通のアドレスで、例えば2段目のブロックを並列に読み出すと、RTRG部分配列の並びに図4. 4 (a)のブロック単位の配列データが得られる。この図から明らかなように、RTRG配列の転送機能により右方向に1ブロック分シフトし、ずれを補正すれば、目的の横長の短冊状配列データが得られる。また、次の2つの処理、

① 右上がりの対角方向の $B_{2,2}, B_{1,2}, B_{N,2}, B_{N-1,2}, \dots, B_{3,2}$ の各ブロック

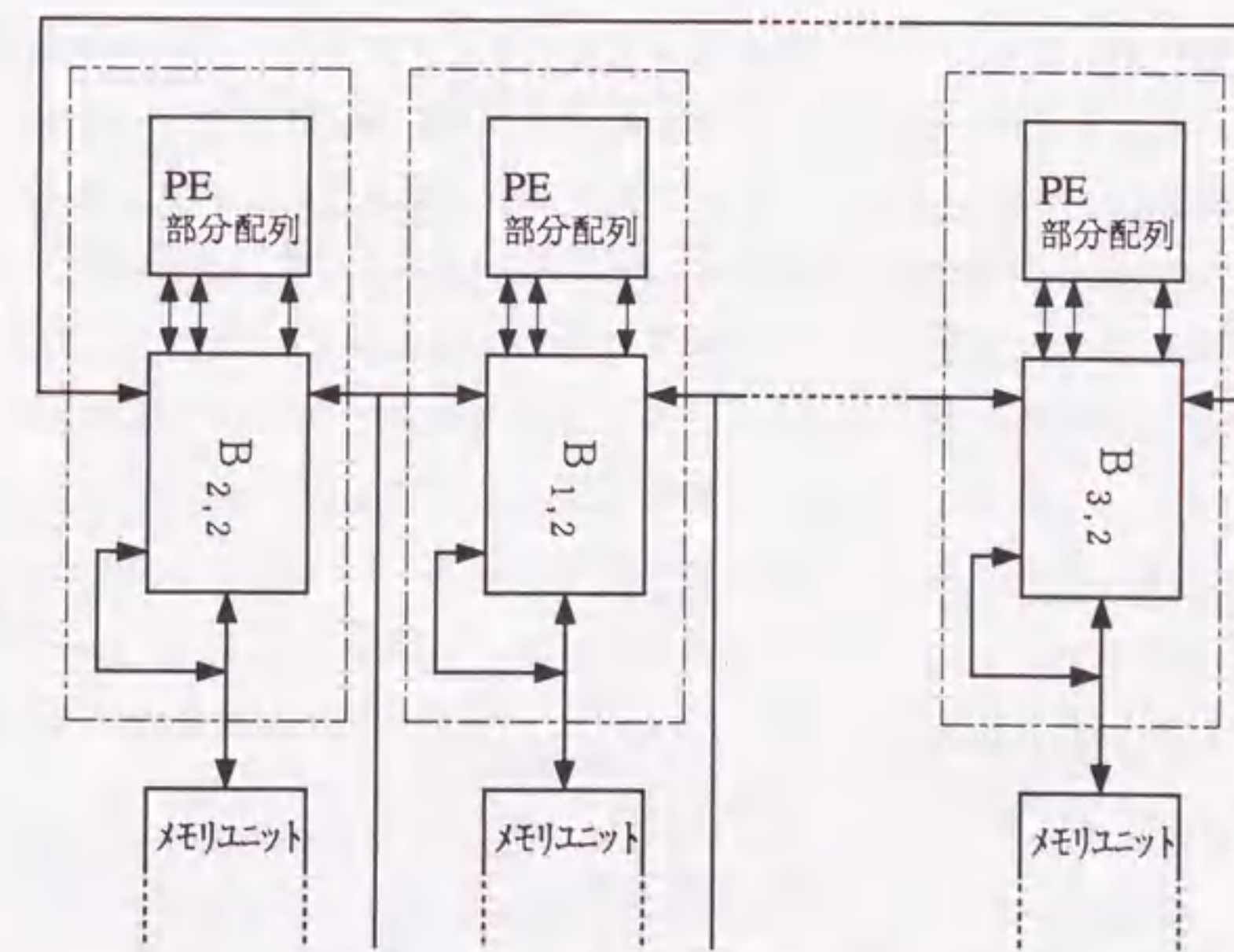
データを、先の間接アドレッシングによるデータアクセス機能を利用して（アドレス値は、各PEの演算機能を利用して算出する）必要なアドレスを分割単位ごとに生成することによって、並列に読出す、

② ①で読み出されたワード単位のデータを、左端のセレクタを介して、回転をかけながら列単位の配列データとして各RTRG部分配列に、順次、取込む

を同時に行うと図4. 4 (b)に示すブロック単位の配列データが得られる。これを先の場合と同様に、2ブロック分だけ左方向にシフトして、ずれを補正すると2列目の縦長の短冊状配列データが横長に変換された形で得られる。ここで、補正のためのシフト量は、容易に導かれるように、 N が偶数の場合、逆方向のシフトも利用すると短冊状配列データ当たりの平均で $(W \times N) / 4$ となる。従って、短冊状配列データのアクセスとそれから行、列を切出してアクセスすることを並行して実行できるようにハードウェアを構成すれば、短冊状配列データを構成する行、列を一通りアクセスする場合の1行



(a) 横長の短冊状配列データの読み出し



(b) 縦長の短冊状配列データの読み出し

図4. 4 転送回転型メモリアクセス機構による疑似2次元アクセス

および1列のビットライン当たりの所要サイクル数 C_{cp} , C_{cp} は、

$$\begin{aligned} C_{cp} &= \{W + (W \times N) / (4S)\} / W \\ &= 1 + N / (4S) \end{aligned} \quad (4.3)$$

$$C_{cp} = 1 + N / (4S) + A / W \quad (4.4)$$

となる。

4.4.4 補正無しの2次元アクセス

これまでの従来型および疑似2次元アクセスでは、アクセスデータのずれをRTRG配列を利用したシフト転送により一々補正することを前提としていた。しかし、行あるいは列を、その並びの順に順次アクセスすることによって行なう線順次処理⁴³⁾においては、アクセスデータのずれ補正の代りに、演算の途中結果の方をずらすことでも対応可能であり、これによってシフト転送量を低減できる場合がある。

実際、従来型の2次元アクセスでは、この方法でメモリユニットからのデータアクセスと、演算の途中結果の左あるいは右方向の1要素分のシフトを同時に行うことにより、ずれ補正のオーバーヘッドをまるまるなくすることができる。これは演算対象の2次元配列データのスキュード配列が、1行あるいは1列ごとに1要素ずつずれていく形式であり、1要素分ずつのシフトで、アクセスデータのずれに合せられるからである。従って、この場合の1行および1列のライン当たりのアクセスの所要サイクル数 C_{rc} , C_{cc} は、

$$C_{rc} = W \quad (4.5)$$

$$C_{cc} = W + A \quad (4.6)$$

となり、アクセス速度は従来型2次元アクセス機構に等しくなる。

しかし、疑似2次元アクセスについてはこの方法だけでは処理効率の改善はできない。これは、演算対象のスキュード配列形式の2次元配列データがW行あるいはW列ごとにW要素ずつ階段状にずれる形式であることから、演算の途中結果のシフトをW行あるいはW列分の処理ごとにW要素分まとめて行なう必要があり、ここまで前提としてきた1次元構成のPE配列では、結局、 W/S サイクルかかってしまうからである。

この疑似2次元アクセスを効率良く実行するには、本章で前提としている1次元の接続構成ではなく、2次元の接続構成が有効である。PEの2次元配列(例えば図1.1)の左右の端を一段ずつずらして接続し、各行を1つのPE部分配列として機能させれば、縦方向のPE間の隣接接続が、図4.1のW要素分のシフトをまとめて行うためのバイパスとして機能するからである。この縦方向の接続をバイパスとしてのみ用いる場合には、PE間の左右方向転送とは同時に動作することはないので、プロセッサ配列部の分割構成時にその接続に必要となる端子を、横方向の1次元配列を形成するために必要となる端子に多重化することができる。

従って、プロセッサ配列が2次元の接続構成を有し、かつ縦方向のPE間の接続のビット幅が、1である場合には、1行および1列のビットライン当たりの所要サイクル数 C_{rp} , C_{cp} は、横長の短冊状配列データの1回のアクセス(W本の行あるいは列のアクセス)に対し、1ブロック分の割合で演算の途中結果をシフトすることになるので、

$$\begin{aligned} C_{rp} &= (W + L / S) / W \\ &= 1 + L / (W \cdot S) \end{aligned} \quad (4.7)$$

$$C_{cp} = 1 + A / W + L / (W \cdot S) \quad (4.8)$$

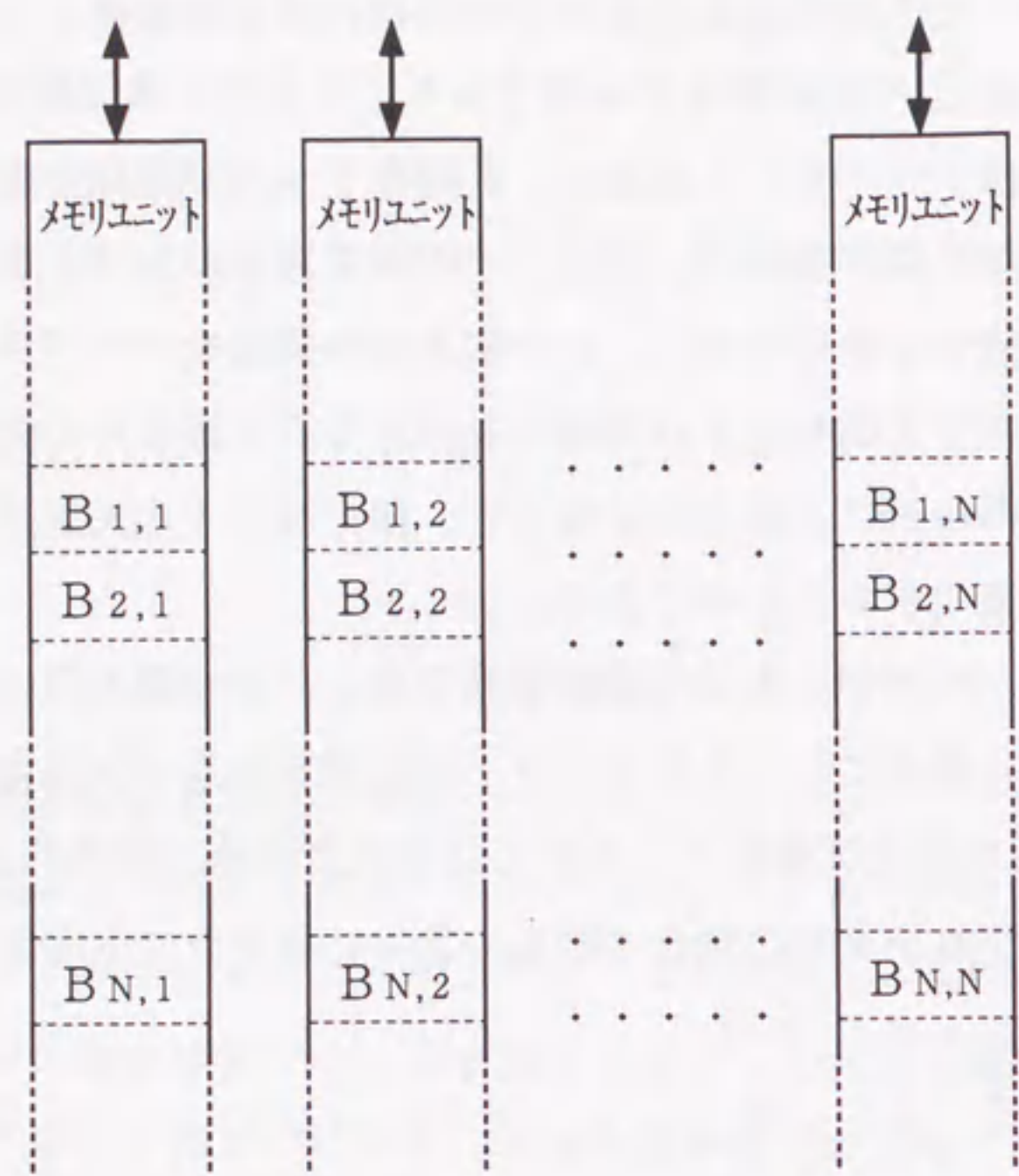
となる。ここで、Lは演算の途中結果の語長である。

4.4.5 間接的な2次元アクセスとそれに用いる90度回転

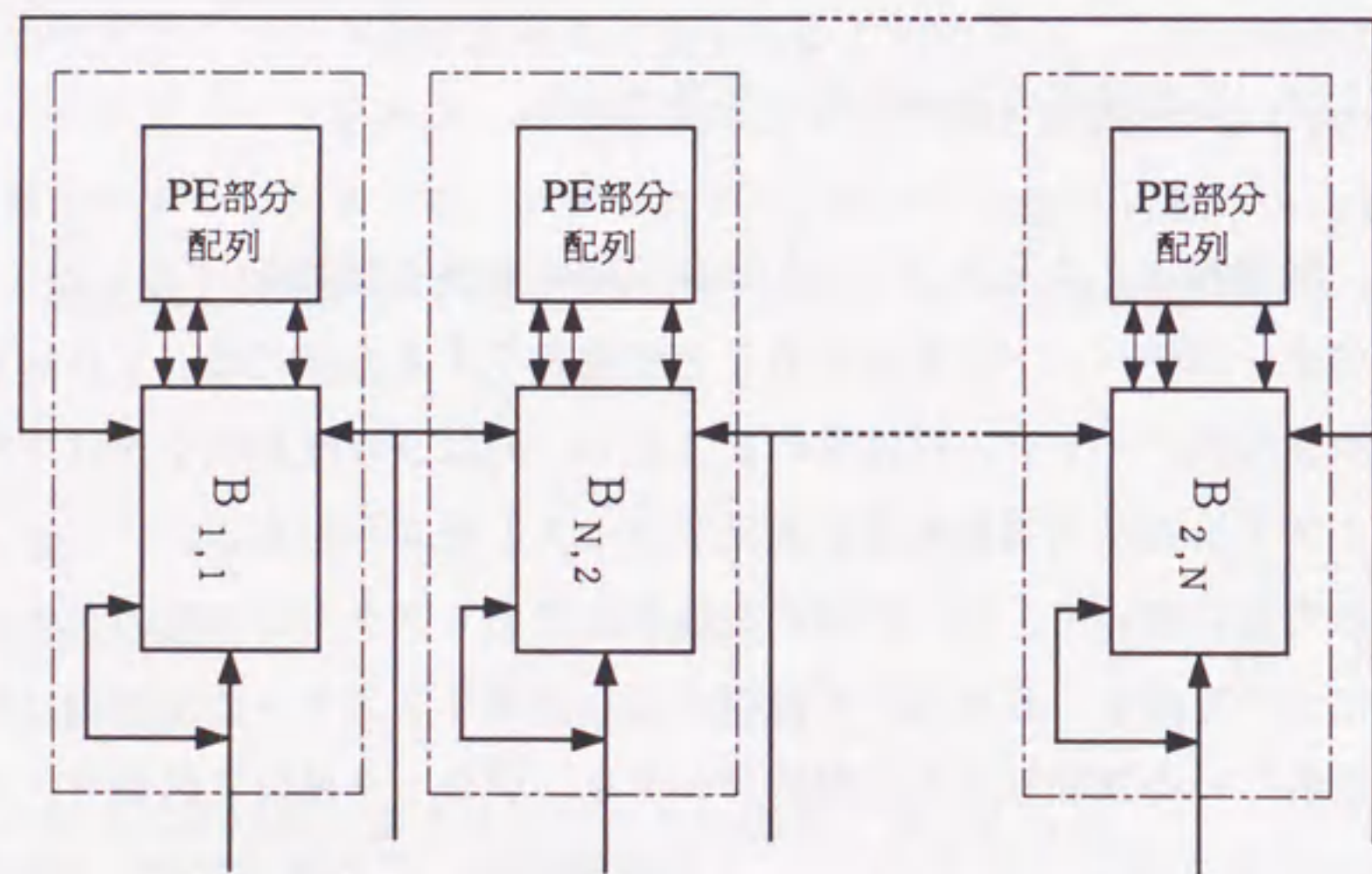
ずれ補正のオーバーヘッドを低減することで高速化する方法には、ずれ無しの形式で格納する2次元配列データを、行はそのままとし、列は90度回転をかけて行に変換してからアクセスする方法(間接的な2次元アクセスと呼ぶ)がある。

以下、この方法において、行、列間の変換手段でありアクセス速度の大半を決める90度回転について図4.5を用いて説明する。メモリユニットの配列内にずれ無しの形式で格納されている被回転2次元配列データを、ブロック単位に区切り、それぞれを $B_{1,1} \sim B_{N,N}$ とする(図4.5(a))。90度回転は、この被回転データに対し、

- ①左からj番目のメモリユニットから、被回転2次元配列データの $\{\text{mod}((i-j+N+1)/N), j\}$ の位置のブロックデータを、間接アドレッシングによるデータアクセス機能を利用して、RTRG部分配列に回転をかけながら読み出す(図4.5(b))。

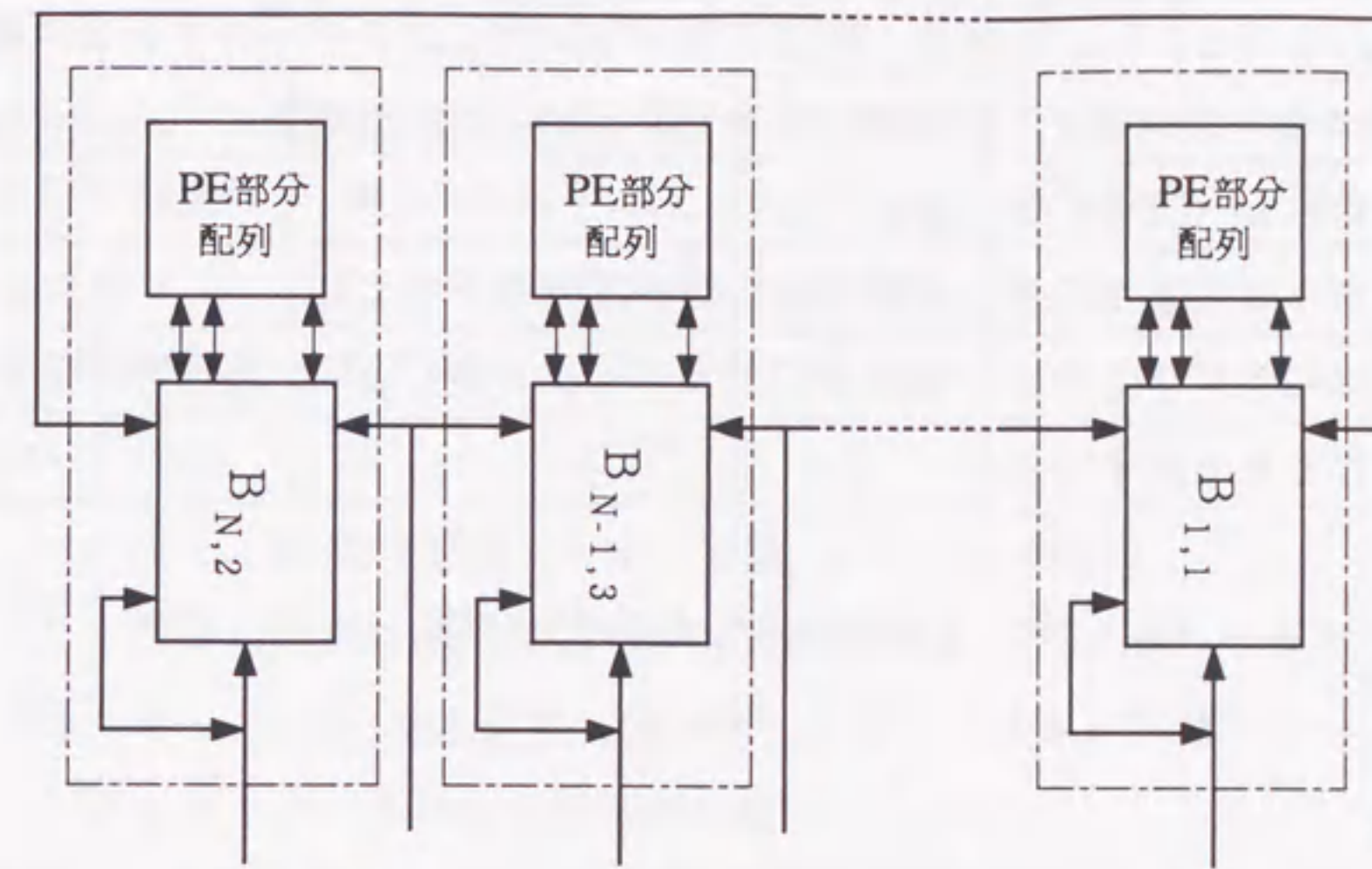


(a) メモリユニット配列内に格納される被90度回転データ

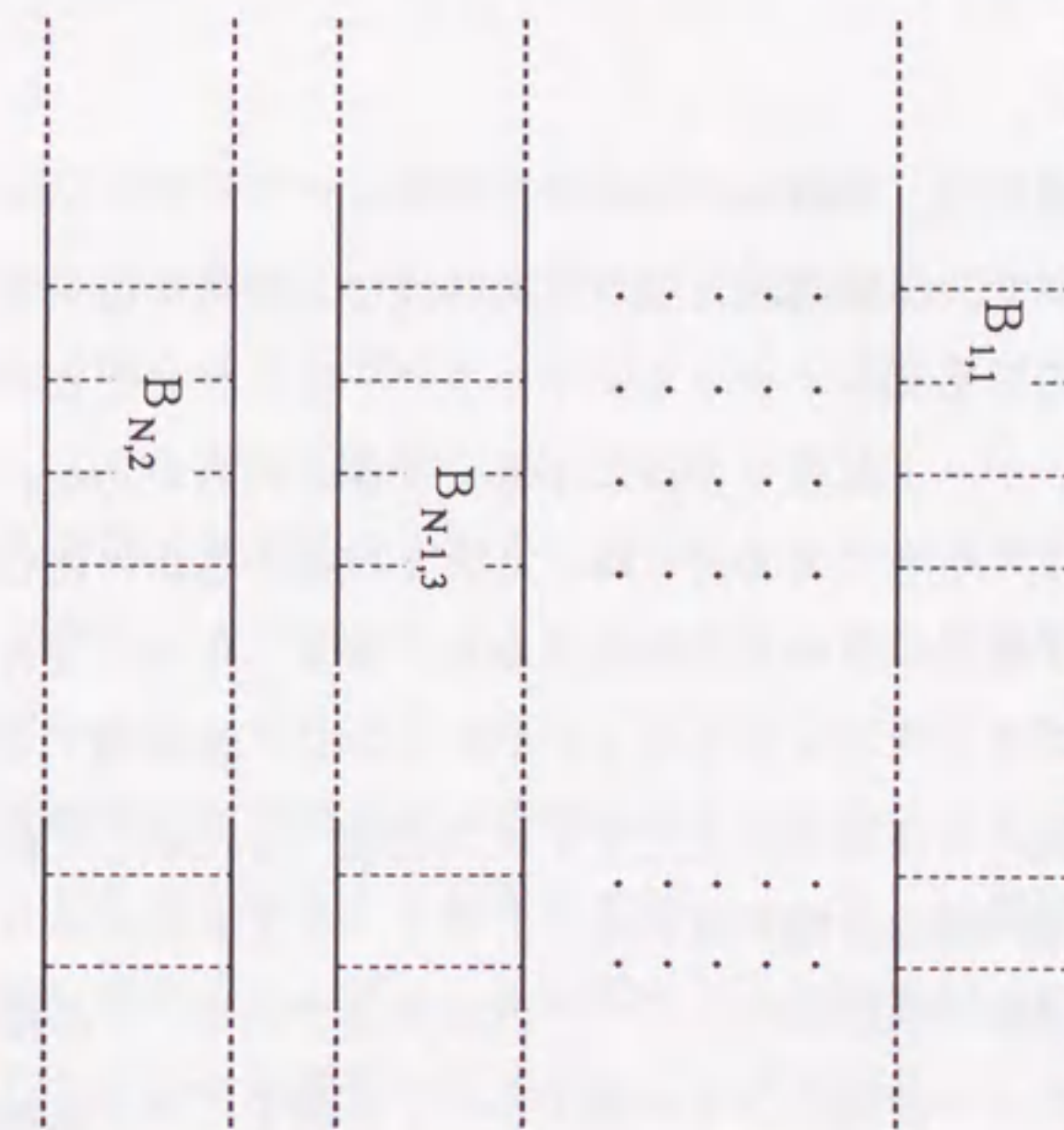


(b) ブロックデータ $B_{(\text{mod}((i-j+N+1)/N),j)}$ の R T R G 配列に対する回転をかけながらの読み出し。

図4.5 転送回転型メモリアクセス機構による90度回転の手順



(c) R T R G 部分配列に読み出したブロックデータの1ブロック ($i=1$) 分左方向にシフトした結果。



(d) (c) のシフト結果の $((\text{mod}((i+j)/N),j))$ の座標位置への書込み

図4.5 転送回転型メモリアクセス機構による90度回転の手順

②RTRG部分配列間で、①のアクセス結果を左方向に*i*ブロック分シフトし
(図4.5(c))、結果を、間接アドレッシングによるデータアクセス機能を利用
して、メモリユニットの回転後の2次元配列データの格納先の $\lfloor \text{mod}((i+j)/N), j \rfloor$ の
位置に書込む(図4.5(d))。

の処理を1から*N*の各*i*に対して繰り返し行って実現される。
行あるいは列のビットライン当たりの所要サイクル数*C₉₀*は、処理の単位がブロック
単位であることを考慮すると

$$C_{90} = 2 + (W \times N / (4S) + 2A + Q) / W$$

$$= 2 + N / (4S) + 2A / W + Q / W \quad (4.9)$$

となる。式(4.9)で、右辺第1項はメモリユニットとの間の読出し、書込みのサイ
クル数である。右辺第3項で*A*に2を乗じているのは読出し、書込みの両方で間接アド
レッシングによるデータアクセスが必要になるからである。右辺第4項の*Q*は、②の間接
アドレッシングによるデータアクセスでRTRGをアドレス生成に利用できるように、シ
フト結果を一旦ワークレジスタに転送するのに要するサイクル数である。

4.5 評価

4.5.1 構成の規則性、単純性

転送回転型メモリアクセス機構は、速度性能の点で理想的な従来型2次元アクセス機
構に比べると、高速のPE間ルーティングネットワーク、PEごとの個別のアドレス修
飾器等が不要であり、ゲート規模を格段に小さくできる利点がある。しかし、PE配列
部の分割単位のLSI化を考える場合には、LSI技術の進歩の著しい現在ではゲート
規模よりもむしろ端子数の多さの方が問題になる。表4.1は、この端子数について、
PE間のルーティングネットワークとして、パレルシフトを付加すると共に、アドレス
修飾器をPEごとに設ける従来型2次元アクセス機構³⁷⁾と比較した結果である。配列サ
イズが256のPE配列部を分割単位であるLSI(32PE/チップ)8個で構成す
る場合の端子数を、実際の数値例として[]内に示している。この構成条件では、端子
数が従来型2次元アクセス機構の304端子から96端子にまで低減されている。この
表の各式および実際の数値例から明らかなように、従来型データアクセス機構では、1
チップ化のための端子数制限を越えてしまう場合でも、転送回転型メモリアクセス機構
ならば容易に制限内に収められることがわかる。

表4.1 分割単位(LSI)の端子数の比較

		従来方式	本方式	本方式/従来方式
分割単位の端子数		$W(N+1)+a$ [304]	$3W$ [96]	$3/((N+1)+a/W)$ [0.32]
内 訳	転送用	$N \times W$ [256]	$2W$ [64]	$2/N$ [1/4]
	メモリアクセス用	W [32]	W [32]	1 [1]
	メモリアドレス用	a [16]* ²	0	-

注1) *W*:部分配列長, *N*:分割数, *a*:アドレスの幅

注2) 処理対象配列データの要素語長は1ビット。

*1: EOR Skewed Array方式³⁷⁾

*2: $a=10$ とした場合の値

表4.2 ビットライン当たりの所要マシンサイクル数*MS*の比較

	従来機構	本機構	従来機構/本機構
2次元アクセス	1	$1+N/(4S)+A/W$	$1/(1+N/(4S))$
90度回転	2	$2+N/(4S)+2A/W+B/W$	$1/(1+N/(8S))$

4.5.2 従来型2次元アクセス機構との高速性比較

4.4で示した各メモリアクセス法および90度回転の所要サイクル数を、従来型2
次元アクセス機構のそれと合せて表4.2に示す。このままでは比較が困難なので、実
際の数値を求めると、

PE内のアドレス計算、RTRGからPE内のワークレジスタへの転送等が比較的簡
単な演算であることから、 $A=2$, $Q=1$ のハードウェアを構成できるとする。これよ
り、 $A/W (=1/16)$, $2A/W (=1/8)$, $Q/W (=1/32)$ 等は小さい
として無視でき、従来型2次元アクセス機構に対する所要サイクル数の増分はほぼ
 $N/(4S)$ になることがわかる。これをもとに、4.5.1の構成条件で $S=1$ の場
合の所要サイクル数を求めると、表4.3に示すように従来型メモリアクセス機構に比
べ、 $1/2 \sim 1/3$ にアクセス速度は低下することがわかる。

しかし、増分の $N/4S$ が、すべて短冊状配列データのずれ補正に必要なRTRG配
列間のシフトに要する分であることから、サイクル当たりのシフト量*S*を大きくできれ
ば、アクセス速度は改善できる。このシフト量*S*を大きくすることは、RTRGの動作
速度がサイクルを決めるメモリアクセス速度、演算器の動作速度等より高速であるた

め、2~3程度までならば、比較的容易である。例えば、Sが1から2にできたとすると、表4.3に示すように、4.1の構成条件のアクセス速度の従来比は、2次元アクセスで1/3から1/2に、90度回転で1/2から2/3にそれぞれ改善される。

表4.3 本機構構成例と従来機構のビットライン当たりの所要サイクル数の比較

	従来機構	本機構		従来機構/本機構	
		S=1	S=2	S=1	S=2
2次元アクセス	1	3	2	1/3	1/2
90度回転	2	2	3	1/2	2/3

4.5.3 各アクセス法的高速性比較

前節では、従来型アクセス機構との比較を分かりやすくするために、転送回転型メモリアクセス機構については、最も効率の良い場合の所要サイクル数を用いた。しかし、各アクセス法には、それぞれ制約があり、これを満たすか否かにより、アクセス速度が左右されるため、従来型2次元アクセス機構との速度性能の差はもう少し大きい。その制約を以下にまとめて示す。

①従来型2次元アクセスの制約

列あるいは行の要素の語長がメモリユニットのアクセス幅Wに一致していること。この制約は、メモリユニットの1ワードに1要素を格納することを前提としていることによる。

②疑似2次元アクセスの制約

列あるいは行の個別のアクセスが、一旦アクセスした短冊状配列データの範囲の中で繰り返し行われること。メモリユニットからの最小のアクセスの単位が横長あるいは縦長の短冊状配列データであることによる。

③ずれ補正なしの2次元アクセス

行あるいは列を順次アクセスして処理を進めること（線順次処理）を前提としているからである。

④間接的な2次元アクセスの制約

行あるいは列を、ある程度以上、一方的にアクセスすること。これによって、必要な90度回転の頻度が小さくなるからである。

これらの制約から、本転送回転型メモリアクセス機構において、単純に各アクセス法的高速性を比較することはできない。いずれのアクセス法が最適であるかは処理内容に

よって変わるからである。しかし、実際の応用では、間接的な2次元アクセスが有利な場合が多いと推定される。これは、一般的な処理では行あるいは列の一方を続けてアクセスする傾向が強く90度回転の頻度を小さくできるからである。

いま、90度回転のオーバーヘッド分を含めた間接的な2次元アクセスによる列あるいは行の平均のアクセスサイクルを C_I とすると、式(4.9)より、

$$C_I = 1 + r \times (2 + N/4S) \quad (4.10)$$

のように表すことができる。ここで、右辺第1項は通常のメモリアクセスの所要サイクル数、右辺第2項は90度回転のオーバーヘッド分のサイクル数である。90度回転の付加サイクルである $2A/W$ および Q/W はやはり小さいとして無視している。また、 r は回転対象の2次元配列データの構成本数分($N \times W$ 本)の行あるいは列をアクセスするごとに必要となる90度回転の回数であり、列あるいは行の1本分のアクセスに加わる平均的な90度回転のオーバーヘッドの割合を表している。図4.6は、この式

(4.10)で与えられる間接的な2次元アクセスの所要サイクル数と従来型2次元アクセスあるいは疑似2次元アクセスの所要サイクル数が等しいことを条件として r と N の関係性を求めたグラフである。この図で、グラフの曲線の上側が、疑似2次元アクセスが速い領域、下側が間接的な2次元アクセスが速い領域をそれぞれ示している。この結果より、転送回転型メモリアクセス機構が、端子数低減、ハードウェア規模低減等で意味を持つ4以上の分割数 N に対して、間接的な2次元アクセスが高速になるのは、条件の悪い $S=2$ で、 r が0.2以下の場合であることがわかる。従って、この値を目安に従来型2次元アクセスあるいは疑似2次元アクセスと間接的な2次元アクセスを使い分けることによって処理の高速化をはかることができる。

ところで、これまでのセルラ配列型1次元SIMDプロセッサの文字認識への応用の経験(8章, 9章参照)からすると、一般的な処理では行あるいは列のいずれかを一方的に処理することが多く、 r の値を0.2以下にすることは難しくない。従って、従来型2次元アクセス、疑似2次元アクセスとの整合性の高い場合、線順次処理が適用可能な場合等を除くと、一般には間接的な2次元アクセスの利用を前提に処理途中の90度回転の回数低減をはかるのが効率的な処理を実現する近道といえる。

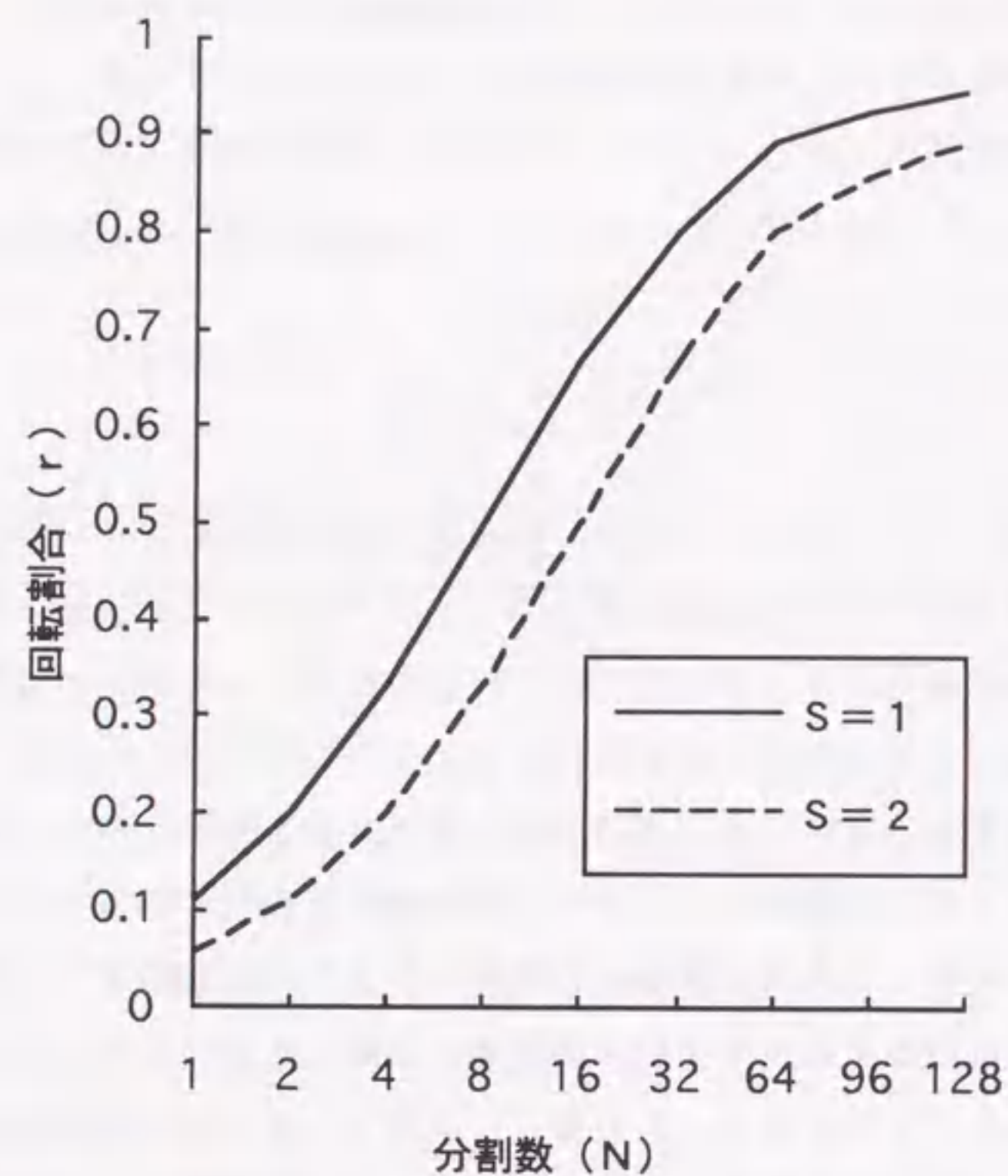


図4.6 疑似2次元モードと間接2次元モードのアクセス速度が等しくなる回転割合 (r) と分割数 (N) との関係

4.6 むすび

基本のSIMDプロセッサにわずかな改造を加えるだけで実現される転送回転型メモリアクセス機構を提案し、その各種のアクセス法を明らかにすると共に、従来型の2次元アクセス機構と比較して、端子数、アクセス速度等を評価した。

はじめに、従来型のメモリアクセス機構が多くの端子数を必要とし、チップ当たりの搭載PE数拡大の妨げになることを示した。次いで、提案する転送回転型メモリアクセス機構として、PE配列を等分割したW個のPEからなるPE部分配列の内部に形成されるWビット幅のルーティングレジスタW個の並び(W x Wの2次元のレジスタ要素配列とみなせる)に対して、

- ①列、行両方向の入出力パスを設け、90度回転器として利用できるようにすること、

- ②ルーティングレジスタの縦続接続からの出力を、PE部分配列に付加されるメモリユニットに対するアドレスとして供給できるようにすること、

の2点の改造を加えたSIMDプロセッサ向きメモリアクセス機構の構成を示した。

次いで、このメモリアクセス機構を用いたPEごとの間接アドレッシングによるデータアクセス、従来型の2次元アクセス、横長あるいは縦長の短冊状の2次元配列データを単位とする疑似的な2次元アクセス、線順次処理を前提とする補正無しの2次元アクセス、2次元配列データの90度回転とこれを用いる間接的な2次元アクセス等のメモリアクセス法を示すとともにその所要サイクル数算出式を求めた。また、これらの2次元アクセス法のいずれが高速であるかについても、所要サイクル数算出式をもとに簡単な解析を行い、行あるいは列のそれぞれを、一定回数(回転対象の2次元配列の構成行[列]数の5倍)以上続けてアクセスする形で行える一般的な処理では、90度回転を利用する間接的な2次元アクセスが高速となることを導いた。

さらに、このメモリアクセス機構を、構成PE数256、分割数8の条件で、バレルシフタとPEごとの個別のアドレス生成器を付加する速度性能の点で理想的な従来型2次元アクセス機構と比較した。その結果、2次元アクセス速度は、2/3~1/2程度に低下するものの、端子数が1/3程度まで低減されることが明らかとなった。

このように、転送回転型メモリアクセス機構は、基本のSIMDプロセッサにわずかなハードウェアを付加するだけで構成される上に端子増も小さく押さえられるので、汎用性の高いセルラ配列型SIMDプロセッサを小形、経済的に構成する上で極めて有用なPE配列構成技術といえる。

第5章 1ビットセルラ配列型SIMDプロセッサ システムの設計と試作

5.1 はじめに

本章では、前章までの検討結果をもとにセルラ配列構成の利点をすべて引出すことを目指して試作した小型高並列プロセッサLISCAR (Line Scannable Cellular ARray processor) と大規模セルラ配列型SIMDプロセッサAAP (Adaptive Array Processor) の設計、構成について述べる。

設計に当たっては、LISCARでは、小型経済的な高並列処理システムを目指し、PE数が 10^2 程度の規模のPE配列を含む全体を1ボードに実装することを、AAPでは、超並列の利点を比較的小型のシステムで享受できるようにPE数が $10^4 \sim 10^5$ 程度の規模のPE配列を1台のラックに実装することを、それぞれ目標とした。これらの目標を達成するために、1.3節(3)の具体的な取り組みの方針に従い、アーキテクチャとしては面単位とライン単位の並列処理の両立が可能な2次元のPE配列を、可変構造とバイパスで補強する最もシンプルなセルラ配列構成を採用することとした。

本章では、はじめに、このLISCARとAAPで規模以外は共通であるPE配列について、その構成要素であるセルラ配列型LSIの構成、設計を中心に述べる。次いで、LISCAR、AAPのそれぞれについて、具体的なハードウェア構成と性能評価、プログラム開発環境等について述べる。なお、セルラ配列型LSIとしては、1981年にはじめて64個の1ビットPEが搭載できることを示したAAP1 (nMOS版)⁴⁴⁾に続き、AAP1 (CMOS版)⁴⁵⁾、AAP2と試作を重ねてきたが、ここではそれらの集大成であるAAP2-LSIを中心に述べる。

5.2 PE配列とその構成要素であるセルラ配列型LSIの設計

5.2.1 PE配列とLSIの構成

PE配列はセルラ配列構成の前提に従い、1チップに数十PEを搭載するLSI (セルラ配列型LSI: AAP2-LSI) とLSI毎に付加するメモリICの対を、正方形格子に配置する単純で規則的な構成とする。図5.1は、そのAAP2-LSIとメモリICの対とPE配列構成との関係を示している。この図からも明らかのように、PE配列の構成・設計とは、セルラ配列型LSIの構成・設計にはかならない。また、LSIの構成自体も、必然的に 8×8 のPE部分配列とその周辺回路とからなる極めて単純なものになる。

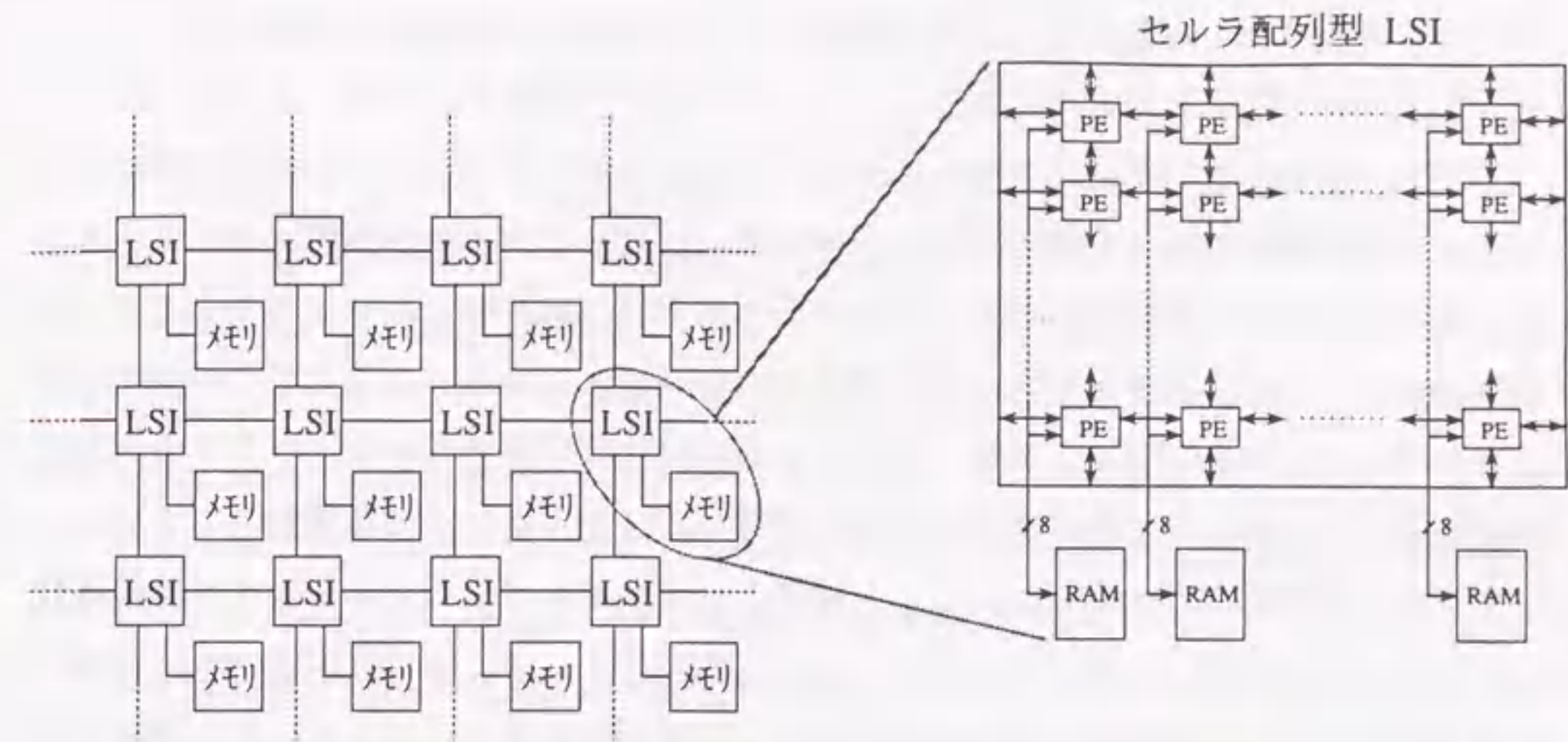


図5.1 セルラ配列型LSIとメモリICの規則的配列で構成されるPE配列

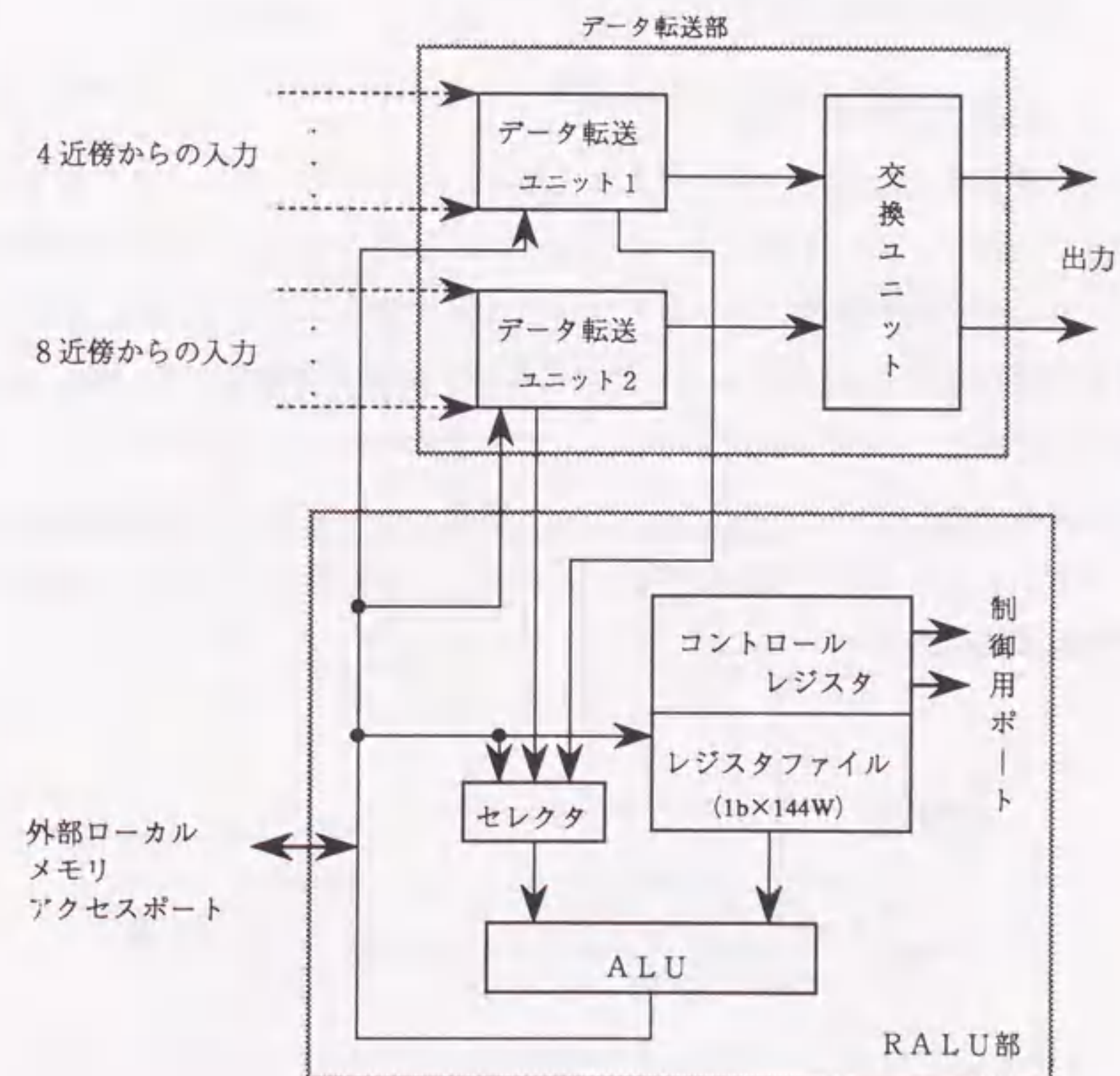


図5.2 PEの概略構成

5. 2. 2 PE構成

図5. 2のブロック構成図に示すようにPEは、レジスタ付きALU部とデータ転送部とからなる。

レジスタ付きALU部は1ビットのALU、入力セレクタ、8ビット構成の制御レジスタ、一部制御レジスタを兼ねるマルチポートレジスタファイル、外部入出力ポート等からなる。8ビット構成の制御レジスタへのロードは、あらかじめレジスタファイルに格納されているデータを8ビット幅の読み出し専用ポートから一括して移す構造により1クロックサイクルで行なえるようにしている。外部入出力ポートは、LSIの入出力端子に直結されており、各PEから直接外部の局所メモリをアクセス可能としている。

データ転送部は二つの転送ユニットを内蔵し、それぞれPE間で4隣接、8隣接接続のネットワークを構成している。パスの転送方向は、3. 2. 4で述べたように制御レジスタによりデータ転送部の入力マルチプレクサを切り換えることにより、各PEで個別にプログラム可能としている。2系統のパス間のデータ交換についても、3. 2. 3で述べたように制御レジスタによって個別に切り換え可能な交換ユニットによってPEごとに選択的に指定できるようにしている。

5. 2. 3 伝搬演算機構の構成

伝搬演算の最も単純な高速化機構である1段のバイパスを、図5. 3に示すようにLSI単位で、各行、各列に設けている。バイパスは端のPEから反対側の端のPEの出力セレクタに途中の7個のPEをスキップする構造として、図2. 6に示されるバイパス系専用の演算器P₁₂を単なるセレクタに置換する構成を採用している。この構成では、図2. 6のP₁₁だけでなく、P₁₂、P₂₁についてもPEでエミュレーションする必要があるため、単なる伝搬転送ではなく演算を伴う場合には多少処理が煩雑になる欠点がある。しかし、PE配列にバイパスとセレクタを付加するだけで構成される点で、セルラ配列構成向きといえる。

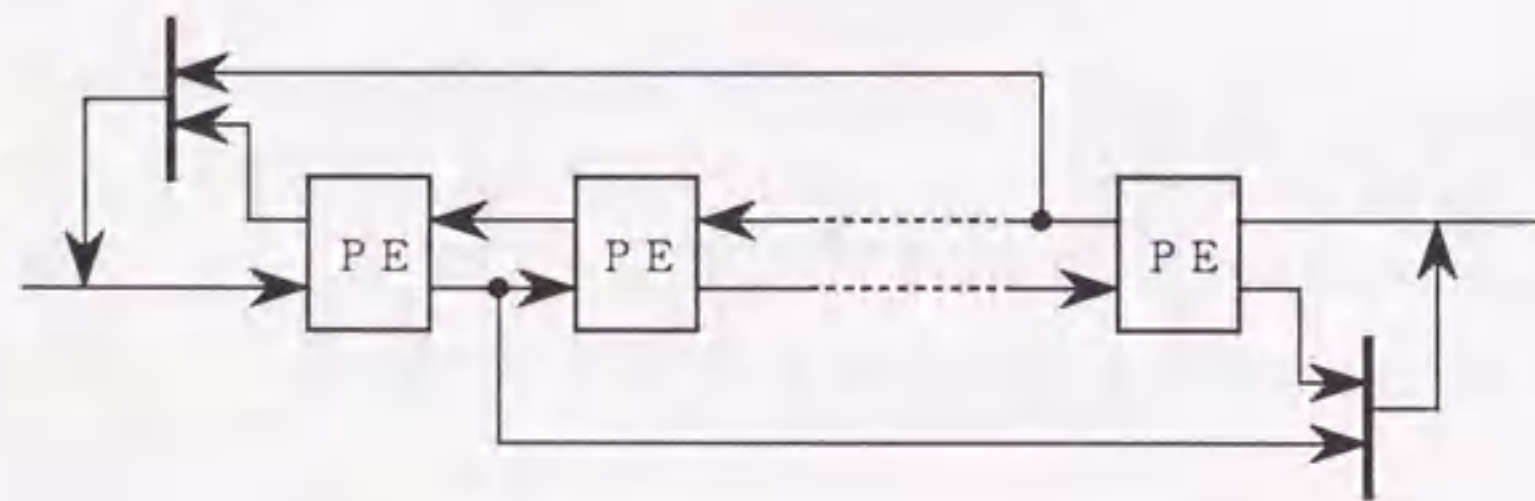


図5. 3 行方向LSIレベルバイパス

5. 2. 4 回路技術

(1) 出力レベル補償型nMOS構成パストランジスタ論理

人手によるレイアウトを前提に、片(n)チャネル型MOSトランジスタ単独構成のパストランジスタによるセレクタ形論理をインバータで受けるタイプの回路(図5. 4)をPE論理部で多用した。例えば、ALUもパストランジスタ論理を用いた4-1セレクタで構成している⁴⁶⁾。従来からのフルCMOS型のNAND、NOR等のみで構成する場合に比べると、ゲート換算の集積度は数倍以上高められるものの、CMOSとPMOSを対抗させる規則的な配置がとれないため、レイアウトは難しくなる⁴⁷⁾。しかし、PEの回路規模は小さくかつ繰り返し使われるので、全体から見た設計の負担増はわずかであり、その割に大きなチップ面積低減効果が得られる。

nMOS構成のパストランジスタ論理では出力のハイレベルが十分上がりきらないため、次段のCMOS回路に貫通電流が流れる問題があった。この問題に対して、入力論理しきい値を超えた時点で自分自身で入力レベルを引き上げる能力を持つ回路(単安定回路)を、次段に配置することによって解決している⁴⁸⁾。図5. 4に示すように、単安定化は、インバータの場合、pMOSトランジスタを一個追加することによって簡単に実現されるので、それによる面積増加、速度劣化はわずかである。

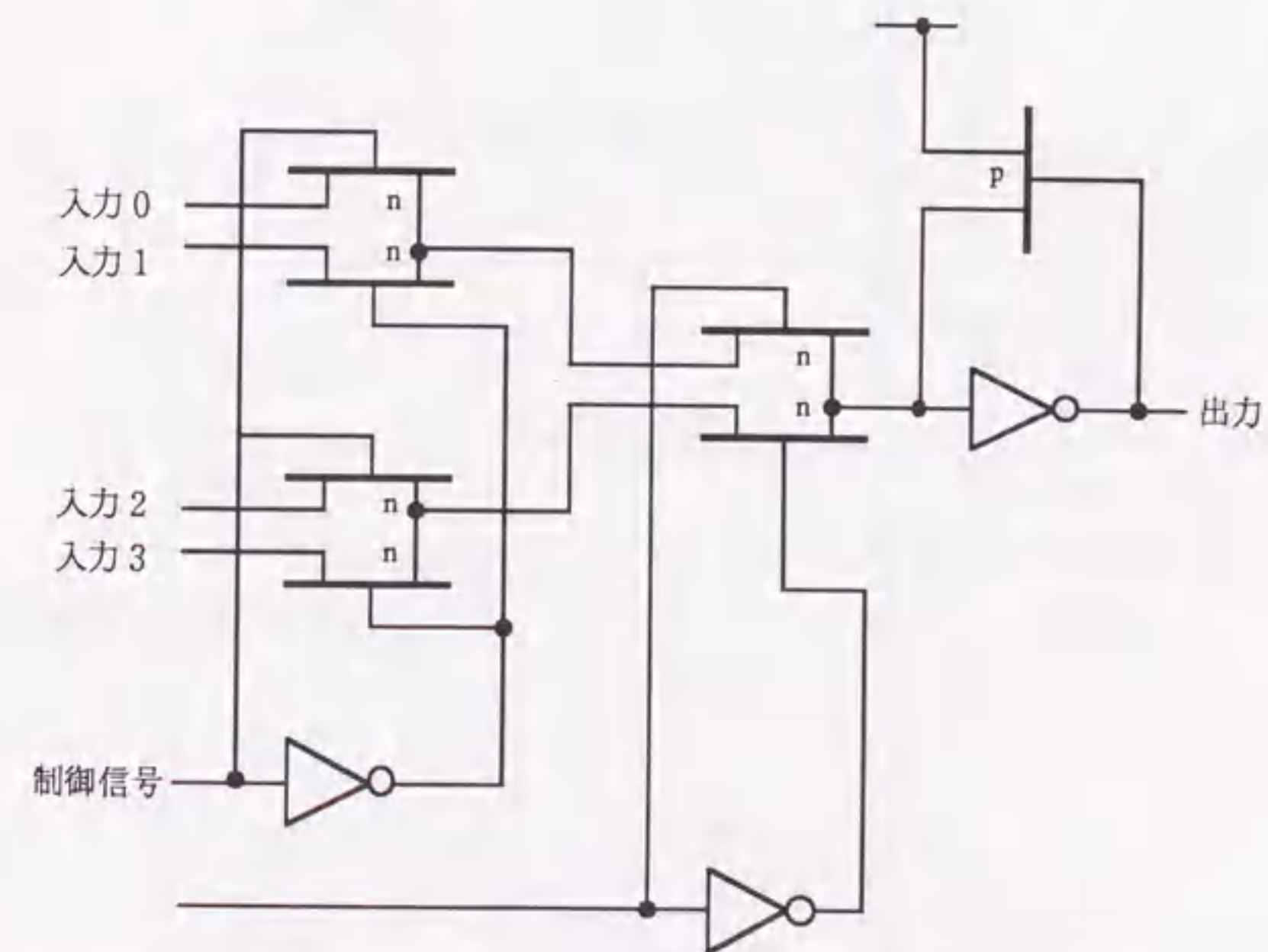


図5. 4 片(n)チャネルMOSトランジスタ単独で構成するパストランジスタを用いた4-1セレクタ

(2) マルチポートレジスタファイル

メモリ配列部を9列x16行の標準的なSRAM(6トランジスタメモリセルの完全CMOS型)構成とするのに加え、アドレスデコーダ部を複数PE(16PE)で共通化することにより、占有面積の低減をはかっている。デジタル線のチャージアップはプリチャージによるダイナミック型とし、アドレスの多重選択の防止と読み出しの高速化をはかっている。

読み出し書き込みポートは、通常の1ビットアクセスRAMの読み出し書き込み用の第1ポートに加え、図5.5に示すようにメモリ配列の各列につながる8個のセンスアンプSAからなり行単位の並列読み出しが可能な第2ポート、16行目のメモリセルから直接制御信号として引き出される第3ポートの3つから構成している。この第2、第3のポートは、読み出し専用で、レジスタファイルの一部を制御レジスタとして利用するために設けてある。特に第2のポートは、制御レジスタの内容をバイト単位で書き換えることを可能にしており、これによって配列の構造を瞬時に切り換える高速の可変構造制御が実現される。

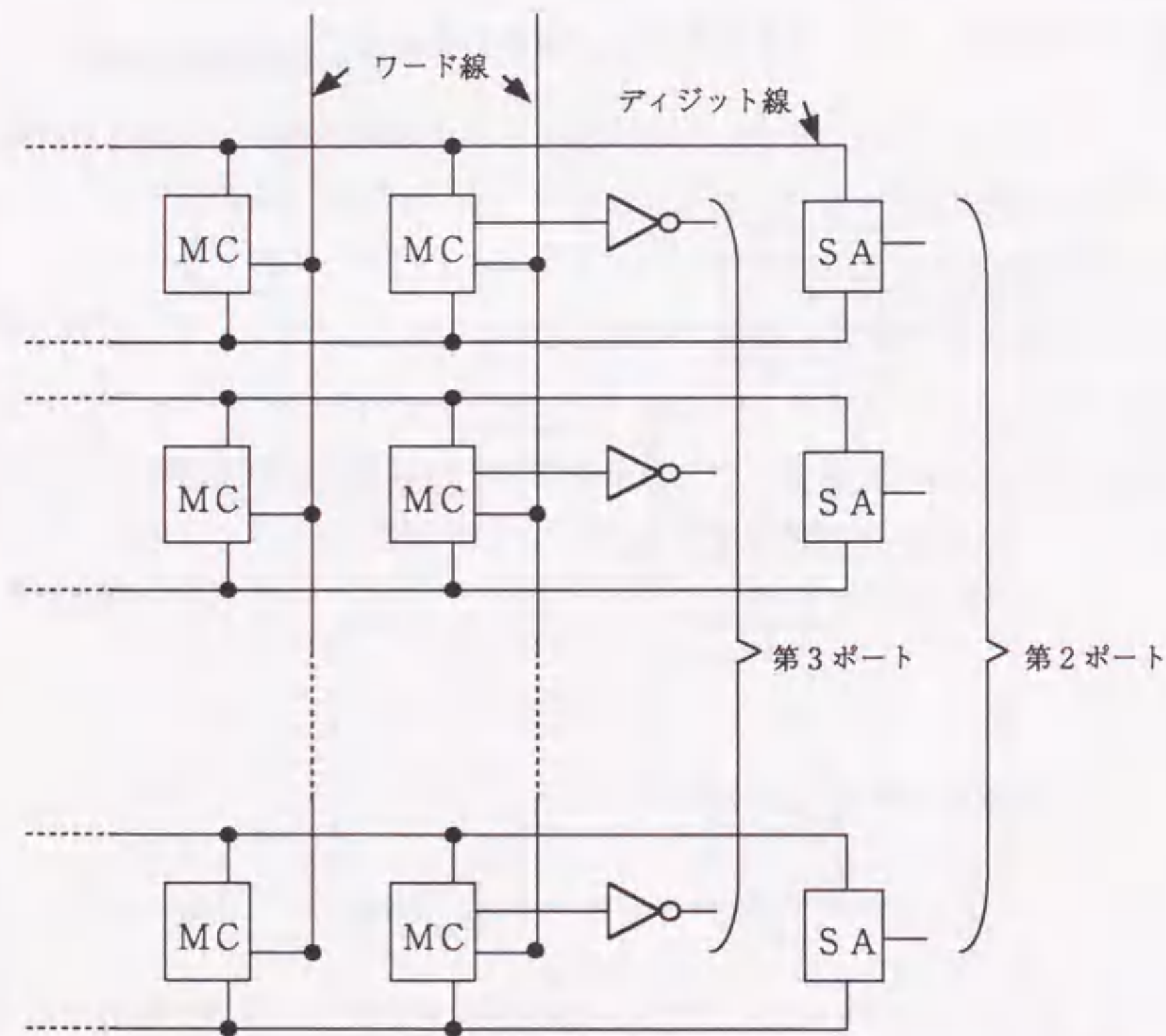


図5.5 レジスタファイルの第2、第3ポートの構成

5.2.5 チップレイアウト

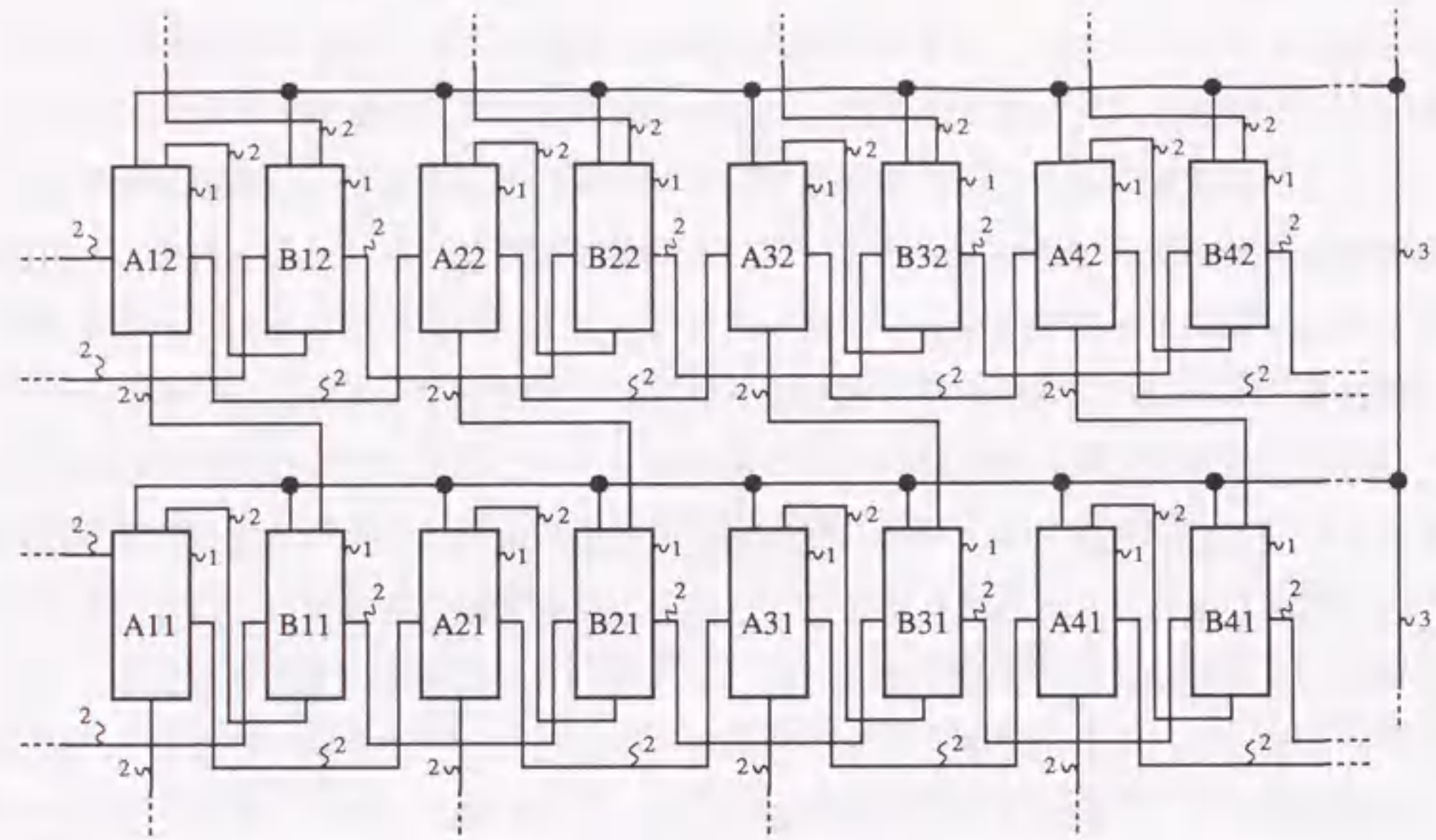
8x8の論理的配列を4x16の物理的配列で実現することにより配線のみ占有される領域を大きく低減している。図5.6にその概略レイアウト構成を示す。

4x16のPE配列と、その左右両側のドライバー配列部およびA12層配線からなる共通制御信号の束とで構成している。ドライバー配列部の各ドライバーからの出力は、各PEの論理部の上を通過させ、16個のPEに対して制御信号を供給している。また、共通制御信号の束の下には、PE配列全体に関わる制御論理を埋め込んでいる。

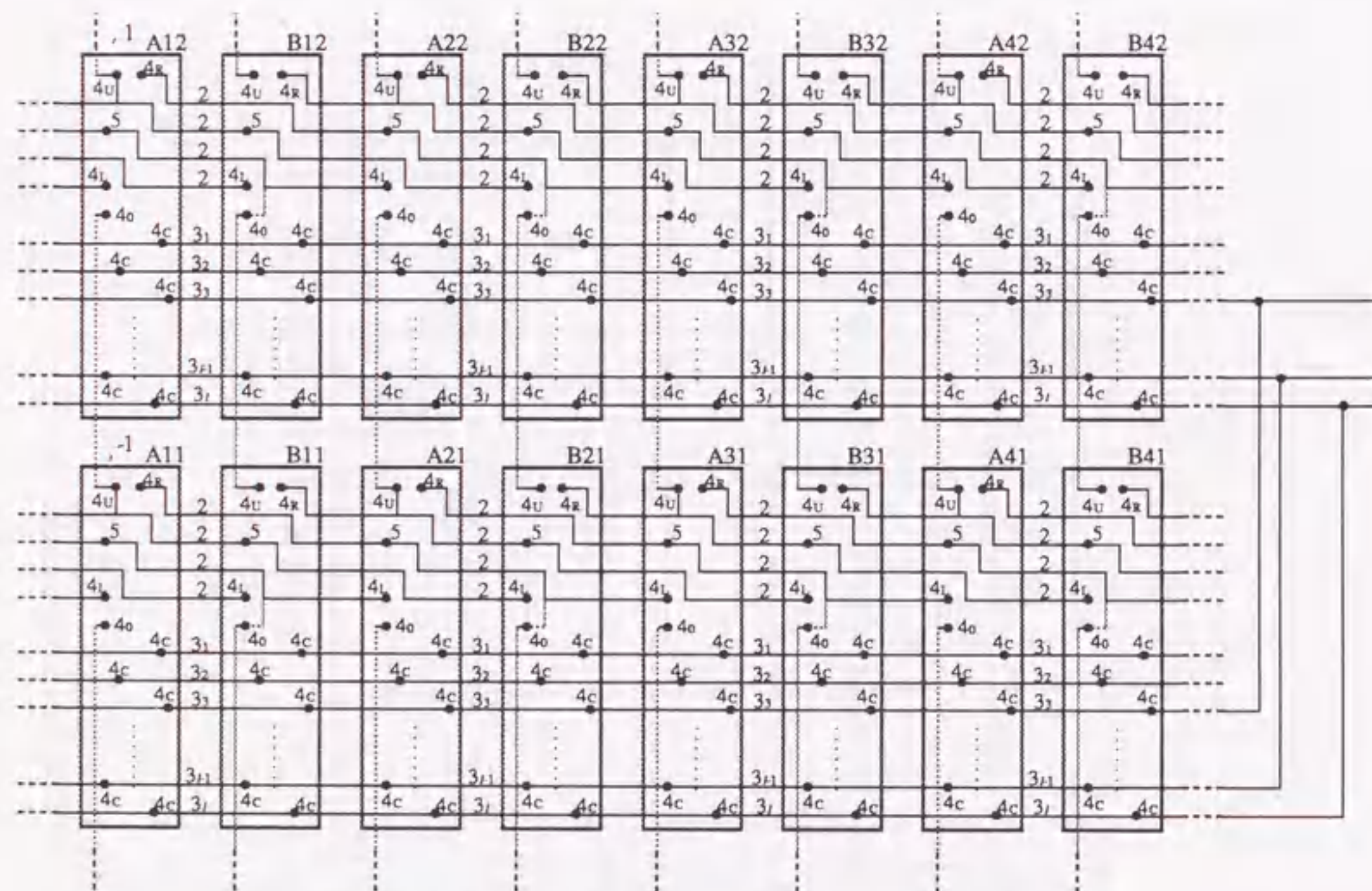
この図5.6から明らかなように、このレイアウトの利点はPEの形状が細長くできるためにPE上の長手方向とは直角の方向にPE上を通過する配線トラックが多数確保できることにある⁴⁹⁾。この構成により、共通の制御信号線、PE間接続線等のみによって専有される配線領域は大きく低減され、PE配列部の集積度が大幅に高まる。一方、このレイアウトのかなめである4x16のPE配列は、PE間配線を複雑とし、かえって面積の増大につながるようにも見える。しかし、この推測は正しくない。図5.7は4x16の配列をとった場合のパス1に関するPE間接続の概略を示している。上PEと下PEを規則的に並べるだけで構成されており、配線の複雑化は生じていない。



図5.6 PE配列の概略レイアウト

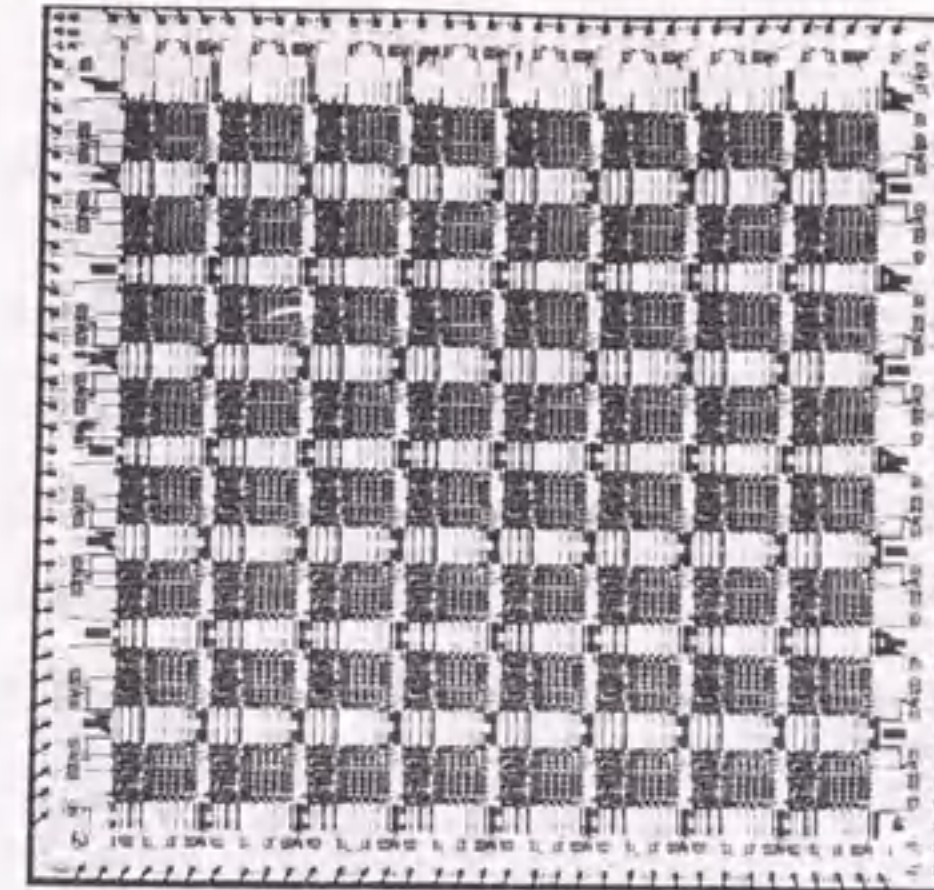


(a) ブロック構成図

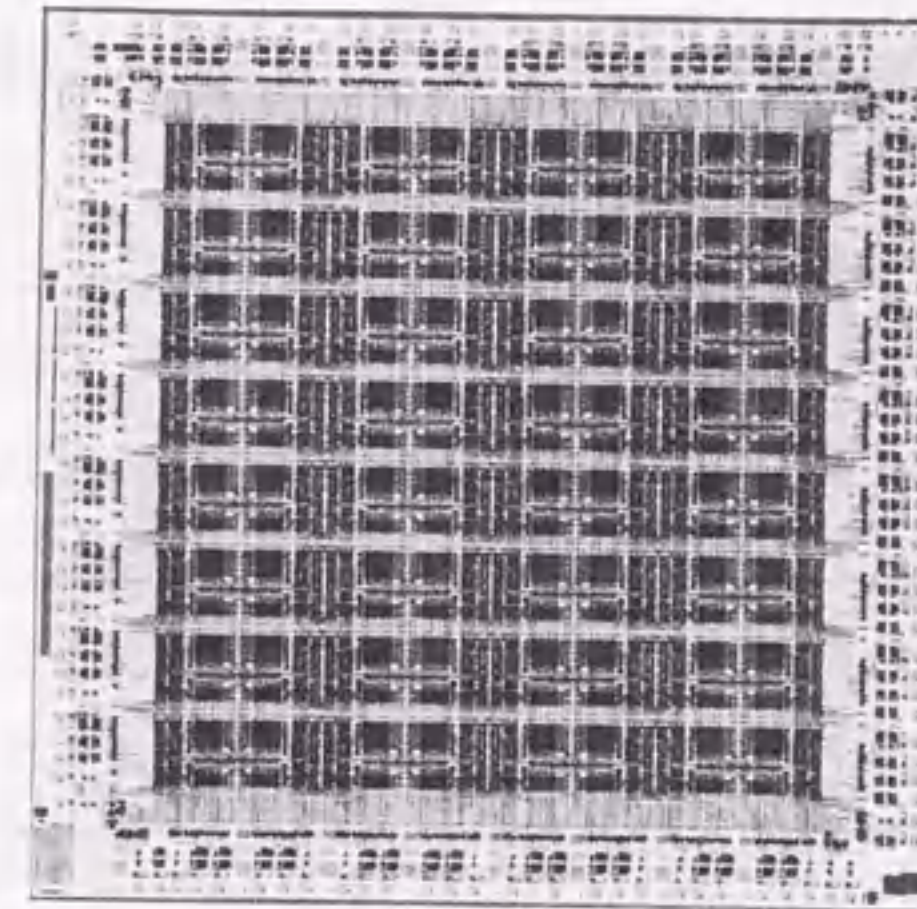


(b) マスクパターンの概略レイアウト図

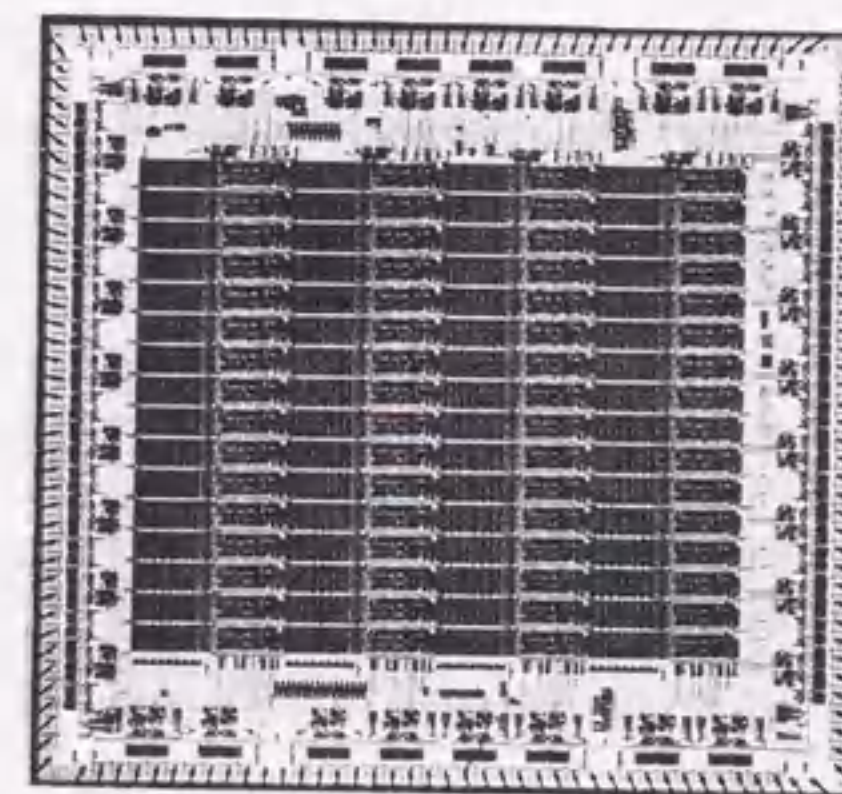
図5.7 パス1に関するPE間接続のレイアウト構成



(a) AAP1



(b) AAP1改



(c) AAP2

図5.8 各AAP-LSIのチップ写真

5. 2. 6 設計・試作結果の評価

2 μ m pウェル2層AlのCMOS技術を用いて実現した本LSIの写真と諸元を図5. 8と表5. 1に、AAP1-LSI (3 μ m Al1層nE/DMOS版)とそれに転送機能の一部を強化したAAP1-LSI (2 μ m Al2層pウェルCMOS版)と並べて示す。AAP2-LSIは、AAP1に比べ、データ転送機能の強化、ALUのファンクションの個別設定機能の付加、レジスタファイル容量の拡大(1.5倍)等の改良を加えている。それにも関わらず、チップサイズ自体は小さくなっている(表5. 1参照)。これは、前節で述べたレイアウトの工夫が効いている。実際、チップ写真から明らかのように、PE配列の中でPE間に走る信号配線の束のみで占められる面積が、AAP1 (3 μ m Al1層nE/DMOS版)では50%程度、AAP1 (2 μ m Al2層pウェルCMOS版)で30%程度をそれぞれ占めるのに対し、AAP2ではほぼ10%程度にまで低減されている。この信号配線専用領域の低減により、AAP2の集積密度は大きく向上している(表5. 1参照)。また、このチップサイズ縮小には

表5. 1 各AAP-LSIの諸元比較

比較項目	AAP1(nMOS版)	AAP1(CMOS版)	AAP2
搭載PE数	8×8	8×8	8×8
プロセス技術	3 μ m Al1層 nE/DMOS	2 μ m Al2層 pウェルCMOS	2 μ m Al2層 pウェルCMOS
素子数	81,000	111,900	104,900
サイクルタイム (typ.)	200 ns	55 ns	55 ns
電源電圧	5V単一	5V単一	5V単一
入出力レベル	TTL	TTL	TTL
特徴的な回路技術	・nMOSパストラ ンジスタ論理多用	・標準フルCMOS回路	・出力レベル補償型 nMOSパストラ ンジスタ論理多用 ・ダイナミックプリ チャージ型SRAM
チップサイズ	10.0×9.4 mm ²	9.85×9.75 mm ²	8.8×8.16 mm ²
集積密度	862素子/mm ²	1,170素子/mm ²	1,401素子/mm ²
端子数	141	144	160
PE当たりの 伝搬演算時間 (typ.)	55 ns	17 ns	18 ns

表5. 2 ゲート当たりの素子数

PE論理部	1.86
周辺回路	2.48
LSI全体	1.97

表5. 3 集積度一覧

チップ全体	1,401 素子/mm ²	
RAM部 を除外 した 場合	PE論理部	1,639 素子/mm ² (1,062 G/mm ²)
	配列部全体	1,401 素子/mm ² (879 G/mm ²)
	チップ全体	809 素子/mm ² (410 G/mm ²)

出力レベル補償型のnMOS構成パストランジスタ論理を多用したPE部の人手設計も効いている。表5. 2はゲート当たりのトランジスタ数を、nMOS構成パストランジスタ論理を多用したPE論理部と、通常のフルCMOS型の論理で構成した周辺部とで、比較したものである。周辺部に比べ、PE論理部が30%近くもゲート当たりのトランジスタ数が少ないことがわかる。パストランジスタ論理はゲート当たりの実現機能が高いので、実際の素子数低減効果は、これ以上と考えられる。また、表5. 3には、得られた集積密度を示している。人手設計と相まって、RAM部を除外した配列部全体では、同程度のデザインルールのスタンダードセルベースのLSIに比べ4~5倍高い集積密度が得られている。

一方、動作速度については、リードモディファイライトサイクルで55 nsと、既存のコントローラLSIと組み合わせるのに十分な速度性能が得られている。AAP-LSIのように単純な構成の伝搬演算機構を採用する場合には、PE当たりの伝搬演算時間も性能上重要となるが、これについても表5. 1に示すように18 nsとサイクルタイムの1/3程度の演算時間を確保している。

5. 3 小型システムLISCAR1の構成

パソコンレベルのホストコンピュータと組み合わせる経済的なシステムを指向して、AAP2-LSIを用いた小型の高並列プロセッサLISCAR1を試作した。本節では、そのハードウェア構成を中心に述べる。

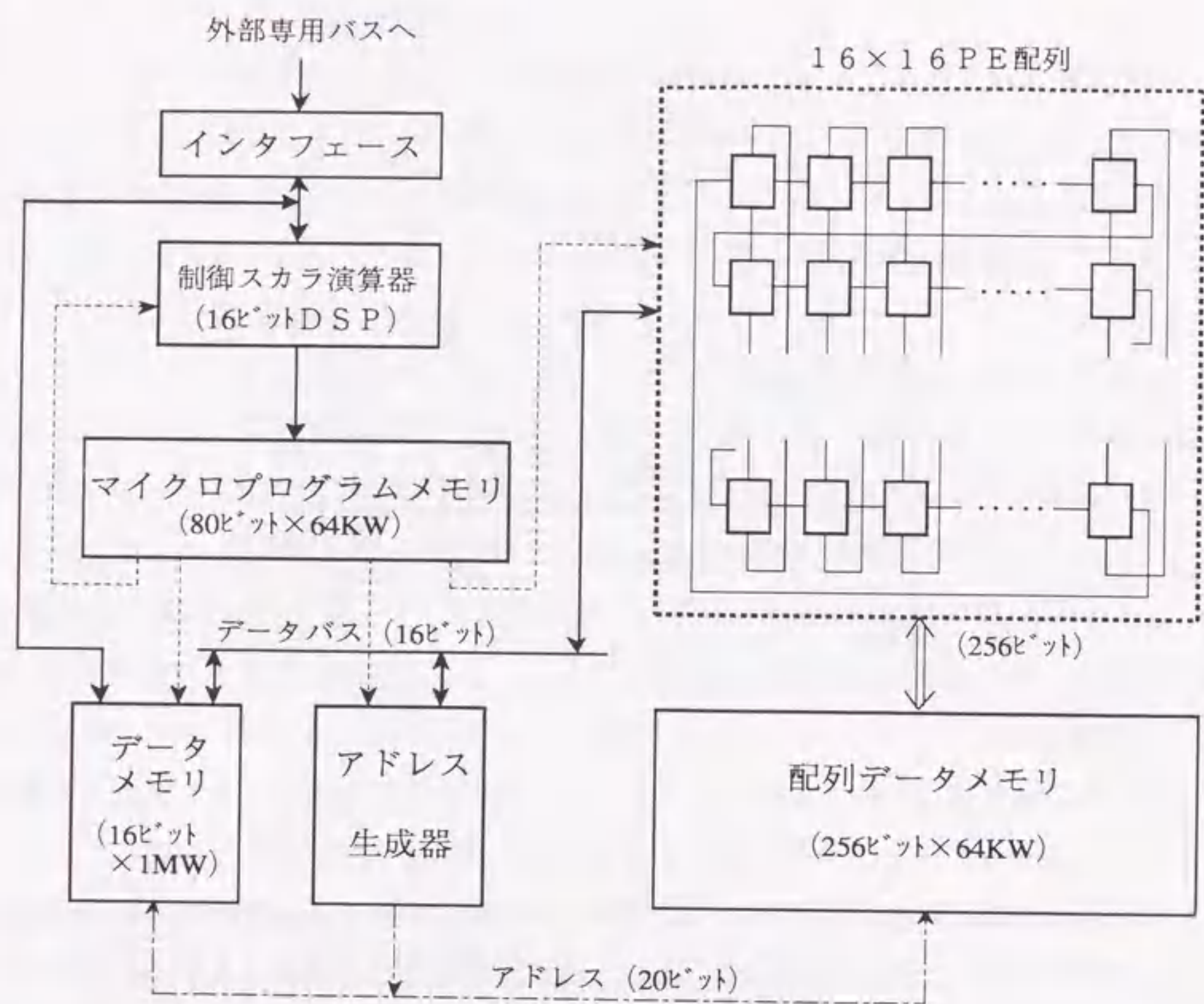


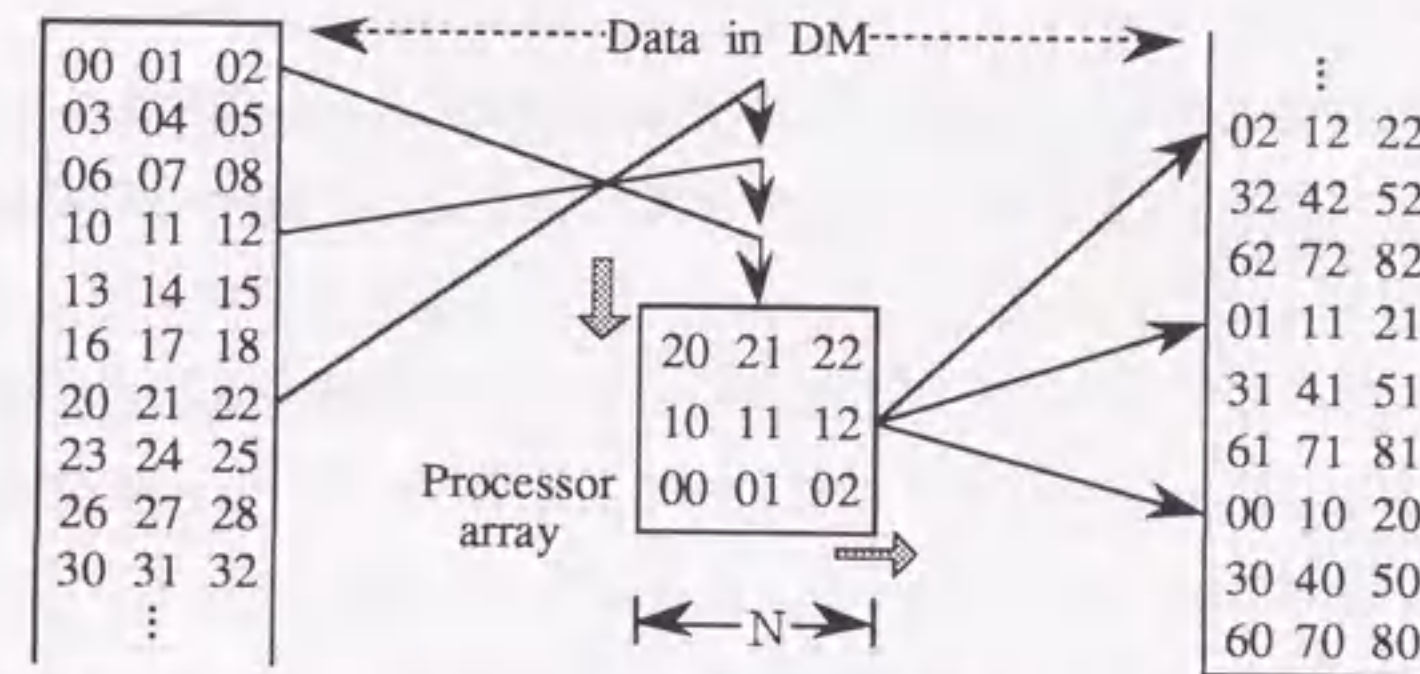
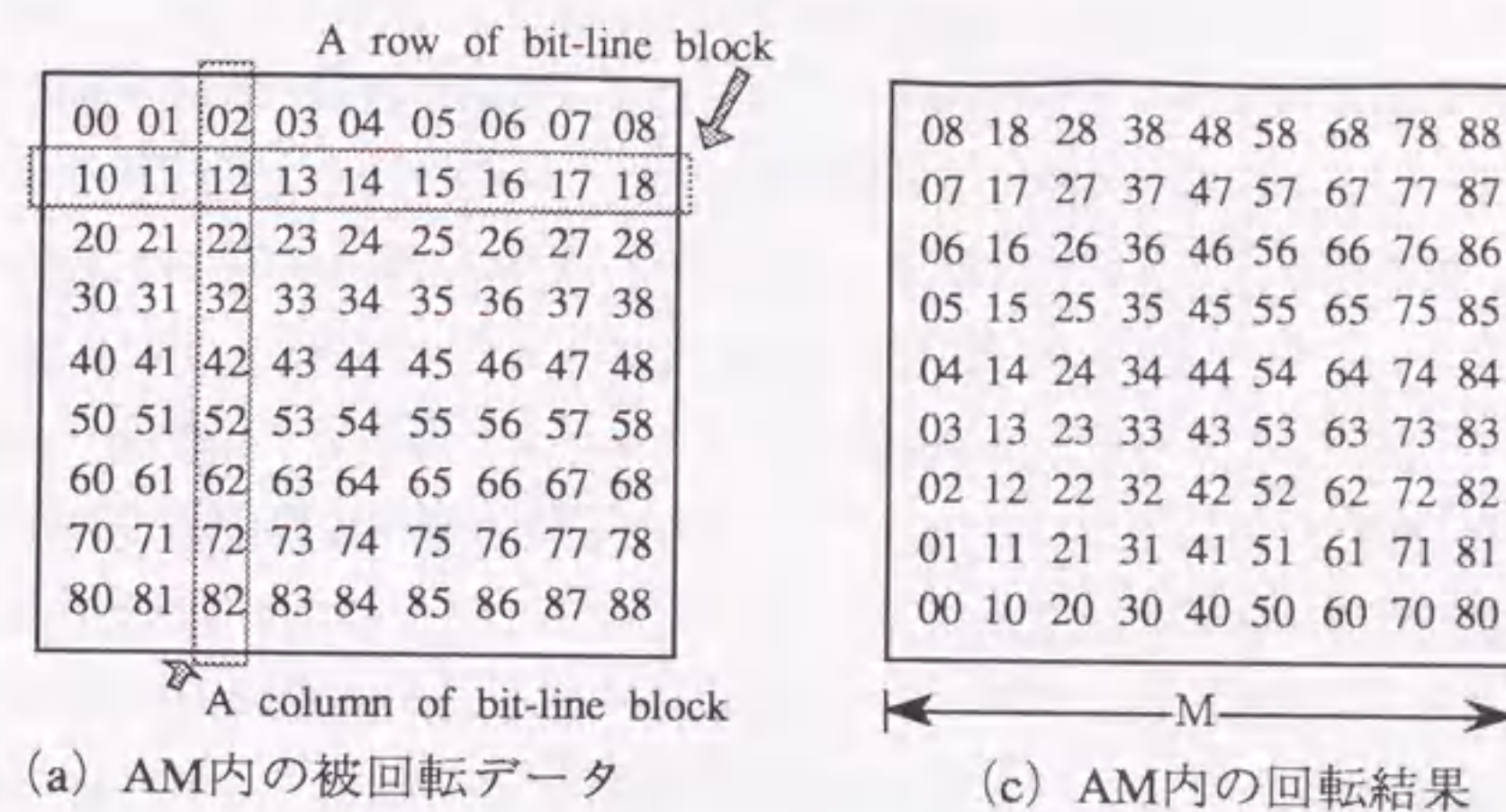
図5.9 LISCAR1の構成

5.3.1 全体構成

LISCAR1は、図5.9に示すように、16x16 PEからなるPE配列部と、制御スカラ演算器、アドレス生成器、プログラムメモリ、データメモリ、配列データメモリ、インタフェース等で構成している。ここで、制御スカラ演算器は、プログラムメモリと対で、全体を一括制御するためのマイクロ命令を生成するとともに、スカラ演算を分担する。アドレス生成器は、データメモリ、配列データメモリをアクセスするためのアドレスを生成する。データメモリはスカラ演算用のメモリであり、配列データメモリはPE配列部の各PEの局所メモリの配列である。

5.3.2 1次元/2次元両用のPE配列

図5.10に図示しているように、PE配列は、各行、各列のそれぞれで、一方の端ともう一方の端を結ぶループ接続により全体としてトーラス型の2次元隣接結合ネット



(b) PE配列とDM間のデータ転送を利用した90°回転

図5.10 LISCAR1における90度回転

ワークを形成している。ただし、2系統の転送系の一方については、各行の端と端の接続を1段ずらす構成(ネステッドトラス)をとっている。これは、16x16のサイズの2次元配列を単位とする動作モードと長さ256の1次元配列を単位とする動作モードの両方を可能とするためである。

5.3.3 90度回転機構

4.1で述べたように1次元のSIMDプロセッサにおいて、メモリユニット配列内の2次元配列データに対し、行単位にでも列単位にでもアクセスできる2次元アクセス機能は、汎用化に必須ともいえる機能の一つである。しかし、LISCAR1は1次元の動作モードをサポートしていながら、行/列の両方向のアクセスを可能とする2次元アクセス機構も、5章で明らかにした転送回転型メモリアクセス機構も搭載していない。このため、配列データメモリ(AM)から一旦データメモリ(DM)上に移した2

次元配列データを、2次元のPE配列との間で入出力する際の16x16のブロック単位の回転機能により90度回転することによって、等価的に、行、列両方向のアクセスを実現している。その回転手順は、以下の通りである(図5.10参照)。

- ①縦、横の長さがPE配列の長さ(256)に等しいAM内のビットプレーンデータ(図5.10(a))を行単位(行方向のビットライン単位)でPE配列上に読みだしてから、PE間のシフト転送により16ビット幅のワード単位に分割してDMに順次転送し、DM内に図5.10(b)の左側に示すように各行が折り畳まれた形式のデータ配列を得る。
- ②図5.10(b)に示すように、左側のDMから、PE配列のサイズに一致する16x16のブロックの2次元配列を順次読みだし、図5.10(b)の中央のPE配列内にPE間のシフト転送を利用して入力する。
- ③PE配列上のブロックの配列データを、入力方向とは直角の方向にシフト転送により出力し、図5.10(b)の右側に示すように、DM内に各列が折り畳まれた形式で格納する。
- ④DM内に得られている列形式の配列データを、列単位でシフト転送によりPEアレイに転送し、行単位の配列データとして、図5.10(c)のようにAMにもどし、90度回転後のビットプレーンを形成する。

以上の90度回転の1行当たりの所要ステップ数Nsは、

$$\begin{aligned} N_s &= (1+W) + W + W + (W+1) \\ &= 4W + 2 \end{aligned}$$

となる。ここで、右辺の第1項~第4項は①~④に対応する1行当たりの所要ステップ数である。また、WはPE配列を2次元配列として見た時の1辺の長さで、LISCAR1ではデータメモリの語長でもある。ただし、PE間のシフト転送が1PE分のシフト当たり1ステップで実行可能で、かつDMのアクセスと並列に実行可能であるとしている。この所要ステップ数は、5章で述べた転送回転型メモリアクセス機構に比べると1桁近く大きく、回転の頻度の高い応用では、性能低下の原因になる。

5.3.4 制御構成

ハードウェア上での命令に階層を設けず、マイクロ命令で全ての構成要素を制御するマイクロコードマシンの制御構成を採用した。この理由は、1命令で複数の構成要素を並列に制御できる高速性の利点だけでなく、高速かつ大容量のメモリが比較的容易に入

手可能な現在では、開発コスト低減の点でも有利と判断したからである。

マイクロ命令の語長は80ビットで、内16ビットをコントローラ用の汎用DSPの制御フィールドに、残りをPE配列、アドレスジェネレータ、各メモリ等、その他の構成要素の制御フィールドに割り当てている。

5.3.5 装置化

(1) 1ボードプロセッサ構成

有力な応用先である文字認識プロセッサとして、小形経済性で十分な競争力の確保をめざしプロセッサ全体をB4版サイズのボード1枚に実現した。そのLISCAR1ボードの写真を図5.11に、諸元を表5.4に示す。このコンパクトな実装は、

- ①PE配列を1チップに64PEを搭載したAAP2 LSIを4個で構成する、
- ②SRAMに比べ1個当たりの容量が4倍以上と大きい標準DRAMを高速アクセスモードで動作させる構成によって、メモリICの使用個数を低減する、
- ③コントローラとスカラー演算器を兼ねる1チップDSPを採用することで、制御スカラー演算部の部品点数を低減する

等により達成している。なお、採用したDSPではアドレス空間が64Kと狭く、データメモリの全領域(1M)を直接アクセスできない問題が生じた。そこで、64Kを超える領域に対しては、1Mまでのアドレッシング可能なレジスタ付きALU LSIからなるアドレス生成器の側からアクセスできるようにして解決している。

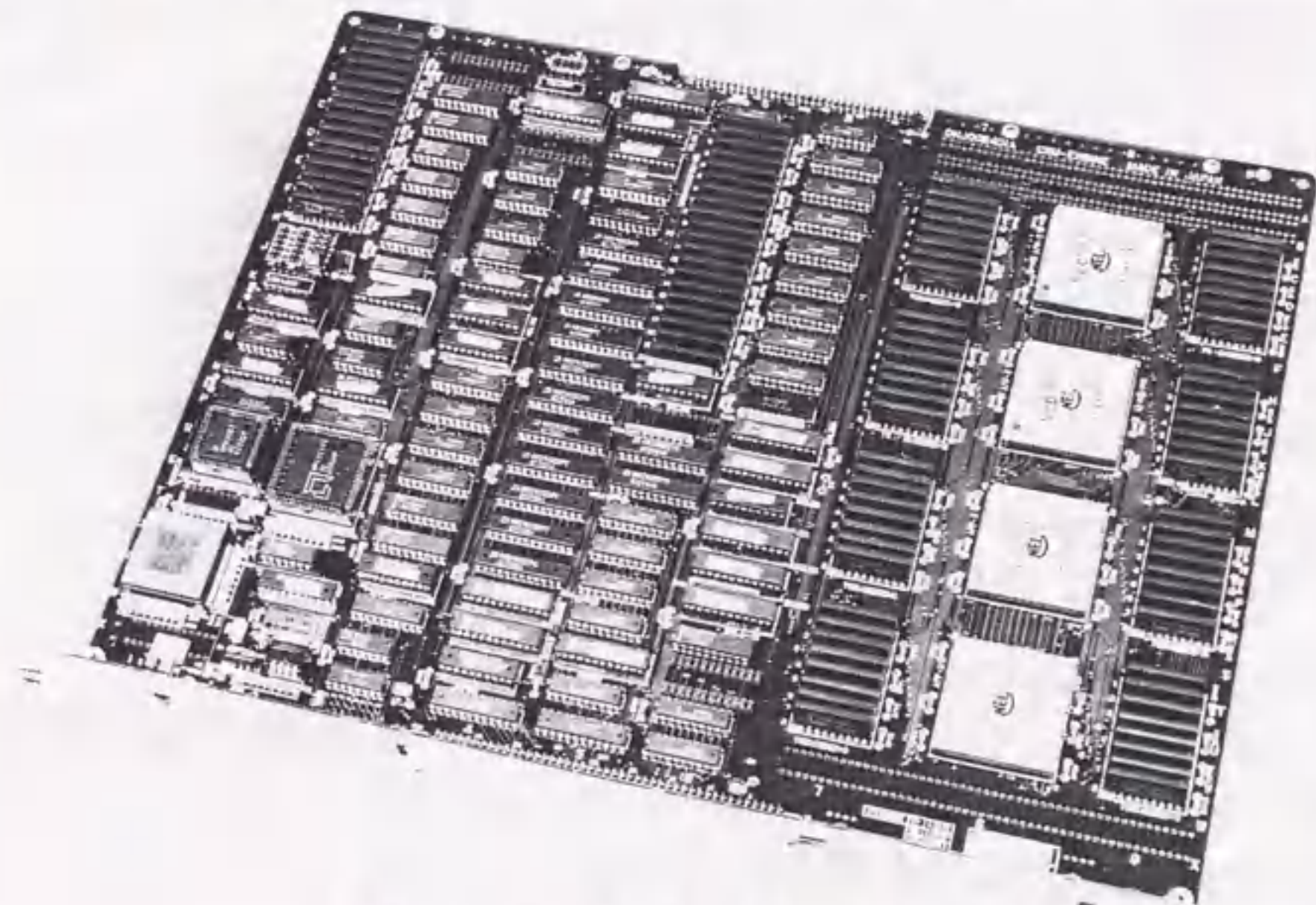


図5.11 LISCAR1ボード

表5.4 LISCARボードの諸元

項目	内容
PEアレイサイズ	16×16 (256×1)
メモリ容量	
プログラムメモリ	640 Kb
データメモリ	2 MB
アレイデータメモリ	2 MB (8 MB)
最小命令実行時間	80 ns
電源電圧	5V ± 0.25 V
消費電力	27.5 W
ボードサイズ	366.7 × 280 mm ²

配列データメモリは、各PEがその1ビット分を占有できるように、アクセス幅をPE数に等しい256ビットに選んでいる。この構成は、高速化には有効であるが、メモリICの個数やコストがネックとなるために、メモリ容量の変更の自由度は低い。そこで、16ビット構成の1Mと4MのDRAMを使い分けて、同一構成のボードで2Mバイトと8Mバイトの配列データメモリの容量を実現している。

速度性能については、最小命令実行時間が最初の試作機では200 nsであったのに対し、回路構成の見直し、使用部品的高速化等により、143 ns版を経て、最終的には表5.4に示す80 ns版までの高速化を達成している。

(2) 装置構成

基本的な装置構成を、図5.12に示す。SCSIインタフェースとこれに専用バスを介してつながるLISCAR1ボードでLISCARユニットを構成し、これをSCSI経由でホストコンピュータに接続する形である。LISCAR1ボードは、専用バスに4枚まで接続可能で、この範囲でLISCARユニットの性能を要求に合わせることができる。また、複数のLISCARユニットを、SCSI経由でホストコンピュータに接続することによっても、性能の向上が可能である。

ここで、各LISCARボードは、ホストコンピュータのバックエンドプロセッサとして機能し、ホストコンピュータよりプログラム、データを受け取り、必要な処理を実行した後、結果をホストコンピュータに転送する。

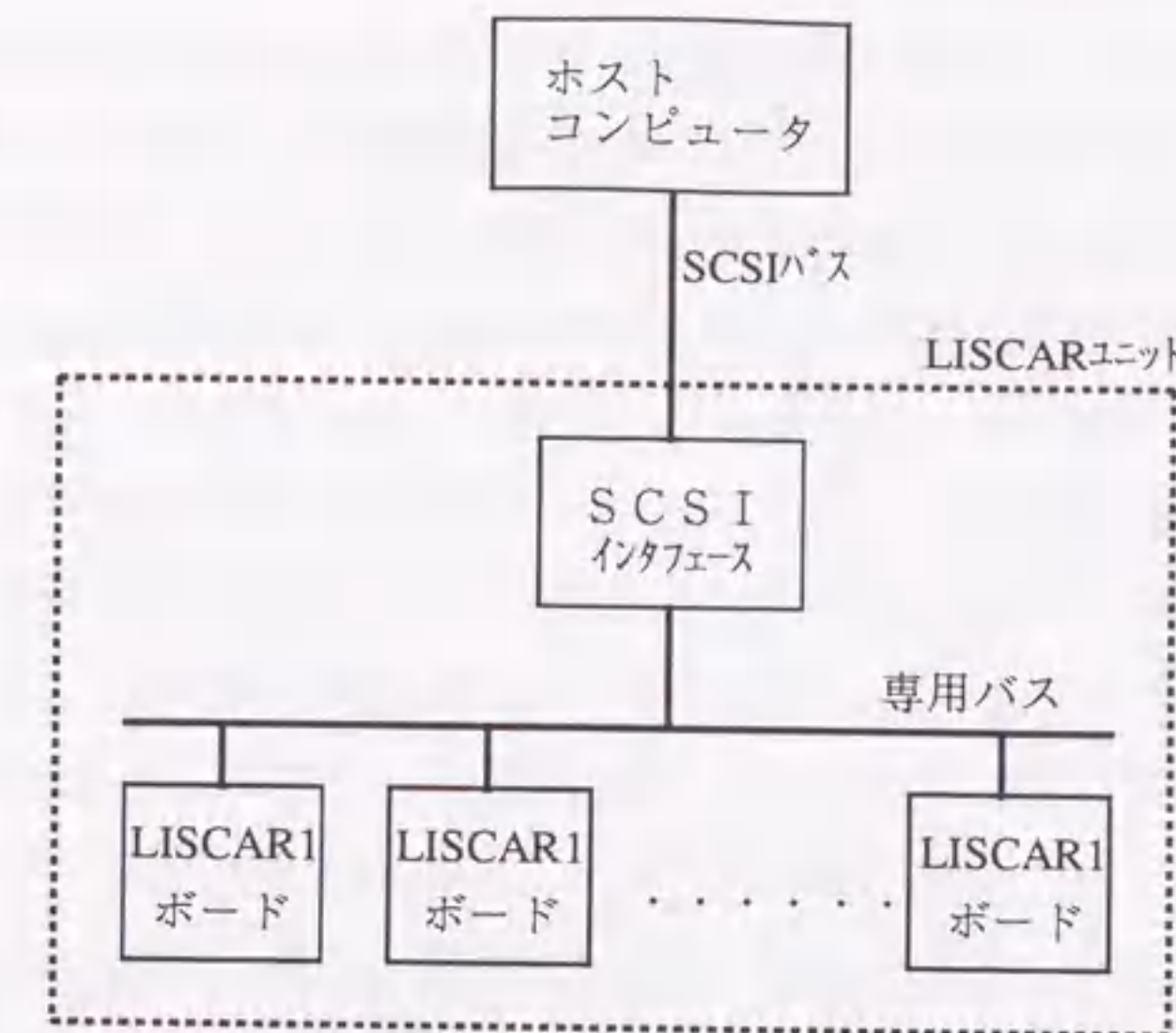


図5.12 装置構成

5.3.6 ソフトウェア開発環境

プログラムのクロスコンパイル、LISCARに対するコンパイル結果、処理結果等のロード・ストア、LISCARの起動、対話型のデバッグ等を行なうことのできるソフトウェア開発環境を、VAX上に構築した。

(1) プログラミング言語

プログラミング言語は、FORTRANの部分ルーチンとしてコールするマクロ命令の集合で構成している。このため、プログラムは形式的にFORTRANプログラムとなり、FORTRAN習熟者ならば容易に修得できる。マクロ命令は、やはりFORTRANサブルーチンとしてコールするマイクロ命令で記述されたマイクロプログラムであり、必要に応じてユーザが作成できる。

変数形式としては、16ビット長のワード変数に加えて、PE配列の1次元の演算モードのデータ形式に対応するライン変数(長さは256の倍数)をサポートしている。PE配列による並列処理は、このライン変数に対する演算として記述する。

(2) デバッグ

マクロ、マイクロ命令の両階層のデバッグに対応するシンボリックデバッグであり、マクロ、マイクロ命令の混在したプログラムのデバッグが可能である。走行中のLISCARに直接割り込んで、状態観測、状態設定を行なう方式を採用している。このため、十分な応答速度が得られ、スムーズな対話型デバッグが可能となった。

5. 3. 7 基本性能評価

PE配列が1ビットプロセッサの配列であることから、特にビットラインに対する性能が高い。長さが256のビットライン間の基本演算では、例えば、算術論理演算および左右両方向のシフトで、実行時間が、それぞれ、90 ns (2,860 MOPSに相当)、320 ns と非常に短い。グレイスケールライン (要素の語長が2ビット以上のライン) に対する演算でも、乗算が入らない場合、例えば8ビットグレイスケールライン間の差の絶対値の累積では、4.1 μs (180 MOPSに相当) と、かなり高い性能が得られる。これに対し、乗算の入る8ビットグレイスケールライン間の差の2乗の累積では、36.4 μs (21.1 MOPS) とかなり低くなる。ここで、グレイスケールライン間の差の絶対値の累積と差の2乗の累積は、それぞれ文字認識処理の識別処理に有用なシティブロック距離計算とユークリッド距離計算である。これらの性能値はいずれもピーク値であり、実際の応用で得られる性能はこれより低くなるが、それでも数十MOPSのレベルにある汎用のDSP等と比べると、はるかに高い性能と言える。表5.5に基本的な2値画像処理の性能を示す。文書画像処理、文字認識処理等で多用する周辺分布、黒点面積等を求める処理では画素当たり数nsと極めて高速に処理できることがわかる。ただし、すべての演算を論理算術演算、データ転送等の基本演算の繰り返しによって実現しているために、平滑化、90度回転などでは処理時間が若干大きくなっている。

この低いほうの性能に注目すると、機能の高い演算でも画素当たり十数ns~数十nsで実行する最近の画像処理プロセッサ⁵⁰⁾の方が優れているように思える。しかし、これらのプロセッサも、すべての演算を数十nsで実行できるわけではない。専用の演算機構を組み込むことによって、高速化を達成しているので、高速化可能な演算は特定

表5.5 画素当たりの演算時間

演算内容	演算時間 (ns)
論理演算	0.95
90度回転	39.2
輪郭抽出	13.4
平滑化	62.7
周辺分布	3.3
黒点面積	3.4

注) 対象画像: 256 x 256の2値画像

のものに限られている。あらかじめ対応していない演算に対しては、結局、数十nsかかる基本演算の組み合わせで実行せねばならず、基本演算を数ns以下で実行するLISCAR1に比べると、はるかに演算時間は長くなってしまふ。従って、文書画像処理、文字認識処理等のように処理内容が多様で、サポートする演算の内容を絞り切れない場合にはLISCARのような汎用的な構成が適しているといえる。

5. 4 大規模システムAAPの構成

LSI-CADの自動配線における迷路法のように大規模な2次元配列データを直接処理する必要があり、逐次処理プロセッサでは十分な速度性能の得られない処理の超高速化を指向して、AAP1-LSIおよびAAP2-LSIを用いた大規模セルラ配列型SIMDプロセッサAAP1⁵¹⁾およびAAP2を試作した。AAP2は、AAP1に対し、高速化と高機能化をはかった改良版であり、構成自体はほぼ同一である。そこで、本節では、AAP2の構成および設計について述べる。

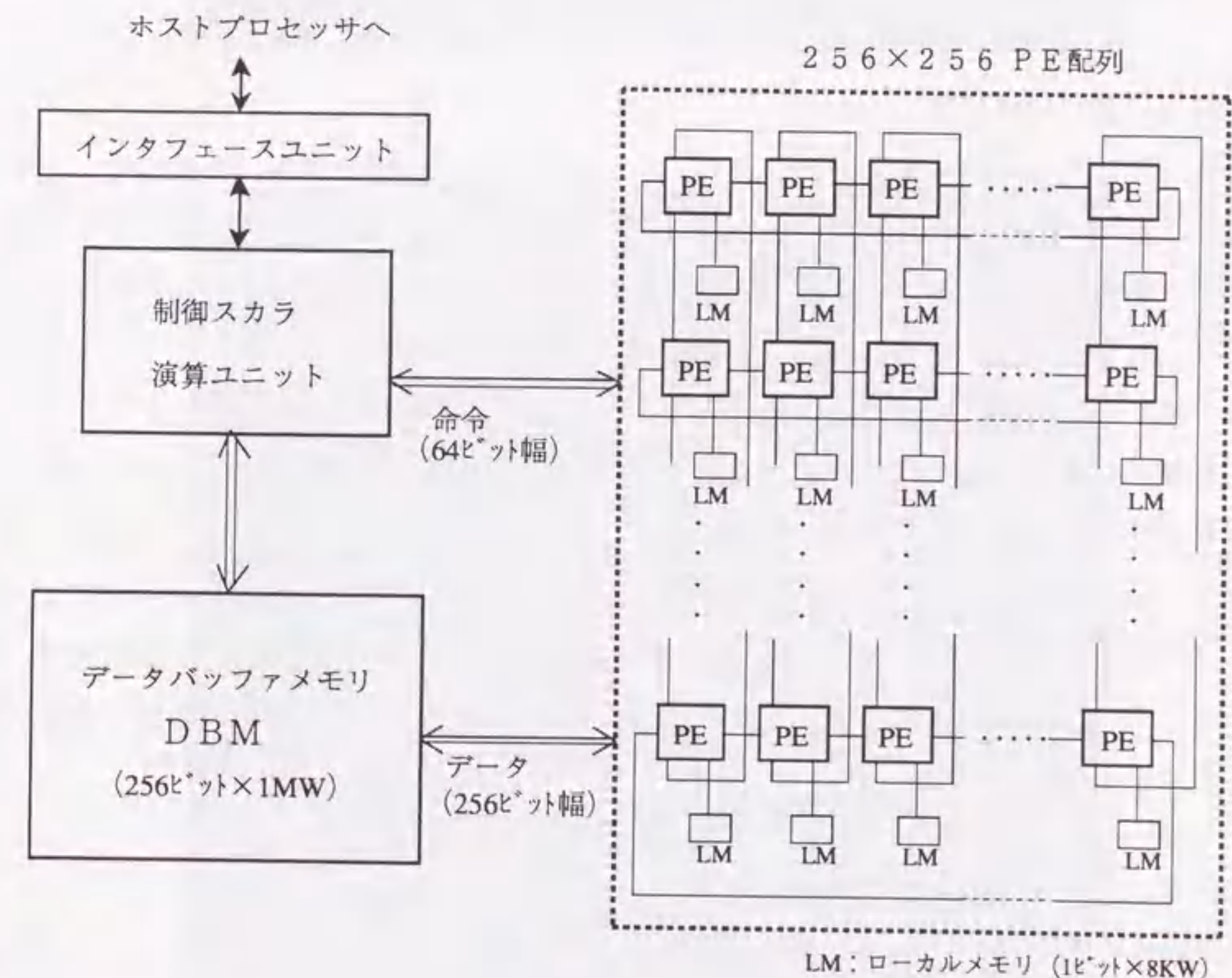


図5.13 AAP2の全体構成

表5.6 AAP2のハードウェアの諸元

項目	内容
配列サイズ	256×256 (AAP2-LSIを64個搭載したボード16枚で構成)
メモリ容量	
データバッファメモリ (DBM)	32Mバイト (256ビット×1Mワード)
インストラクションメモリ (IM)	1Mバイト (64ビット×128Kワード)
配列データメモリ (LM配列)	64Mバイト (65536ビット×8Kワード)
マシンサイクル (最小命令実行時間)	330 ns
基本演算性能	
8ビット固定小数点加算	7,355 MOPS
32ビット固定小数点加算	2,006 MOPS
ラックサイズ	1.3×1.55×0.95 m ³

5.4.1 全体構成

AAP2の全体構成を図5.13に、ハードウェアの諸元を表5.6に示す。256×256のPEの2次元配列からなるPE配列、制御スカラ演算ユニット、デュアルポート構成のデータバッファメモリ、インタフェースユニット等で構成している。

5.4.2 制御スカラ演算ユニット

各種の搭載メモリに対するアドレス・PE配列用の制御信号等の発生、およびスカラ演算等を行うユニットであり、シーケンサ、インストラクションメモリ、スカラプロセッサ、2ポートレジスタファイル、カウンタ、シフタ等で構成している。本ユニットの各部を効率良く制御するため、LISCAR1同様、階層的な制御構造は採用せず、水平型のマイクロ命令に近い形式の64ビット長の命令でハードウェアの各部を直接制御する方式を採用している。

5.4.3 PE配列とその実装技術

大規模なPE配列をコンパクトに実装するために、AAP2-LSIチップ1個と局所メモリ(LM)用にSOP(Small Outline Package)タイプの64KビットSRAM

8個を表と裏に搭載したQFP(Quad Flat Package)タイプのセラミックのモジュール(図5.14参照)を開発した。このモジュールを64個載せた図5.15に示すプリント基板16枚で、256×256のサイズのPE配列を、わずか0.2m³のスペースに実現している。このモジュールは、高密度実装が可能だけでなく、不良時の交換も比較的容易であるという利点ももっている。

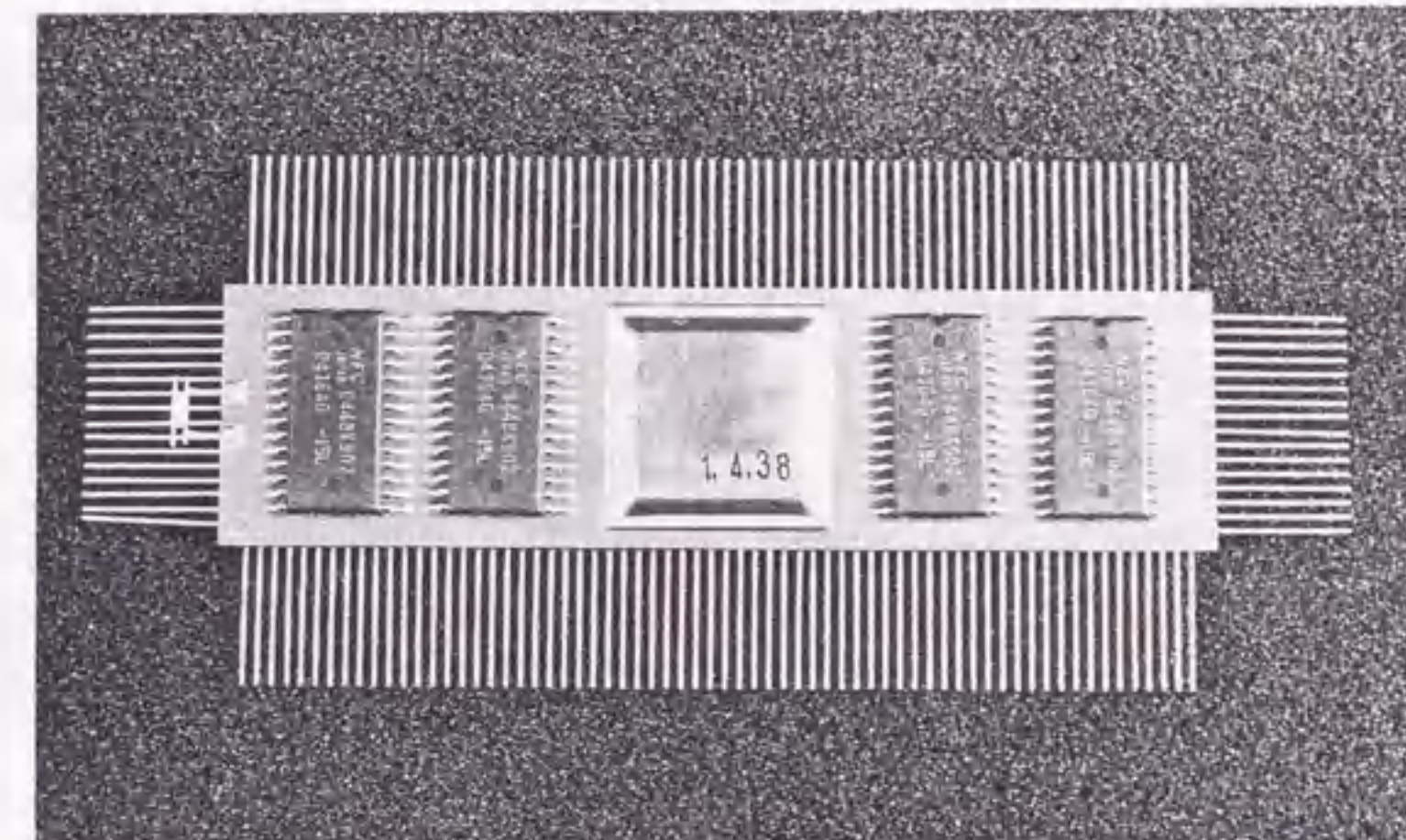


図5.14 AAP2モジュール

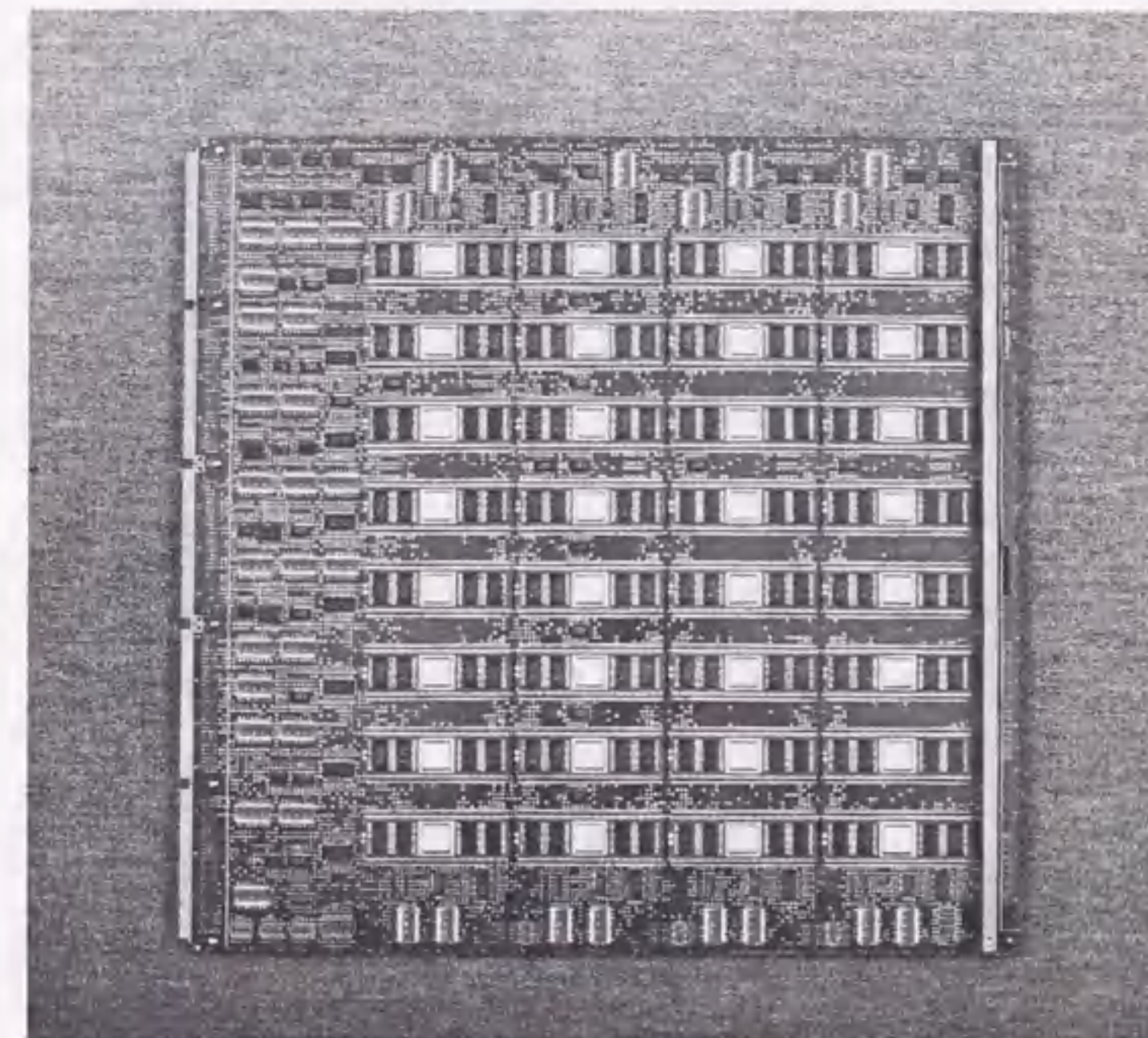


図5.15 AAP2 PE配列ボード

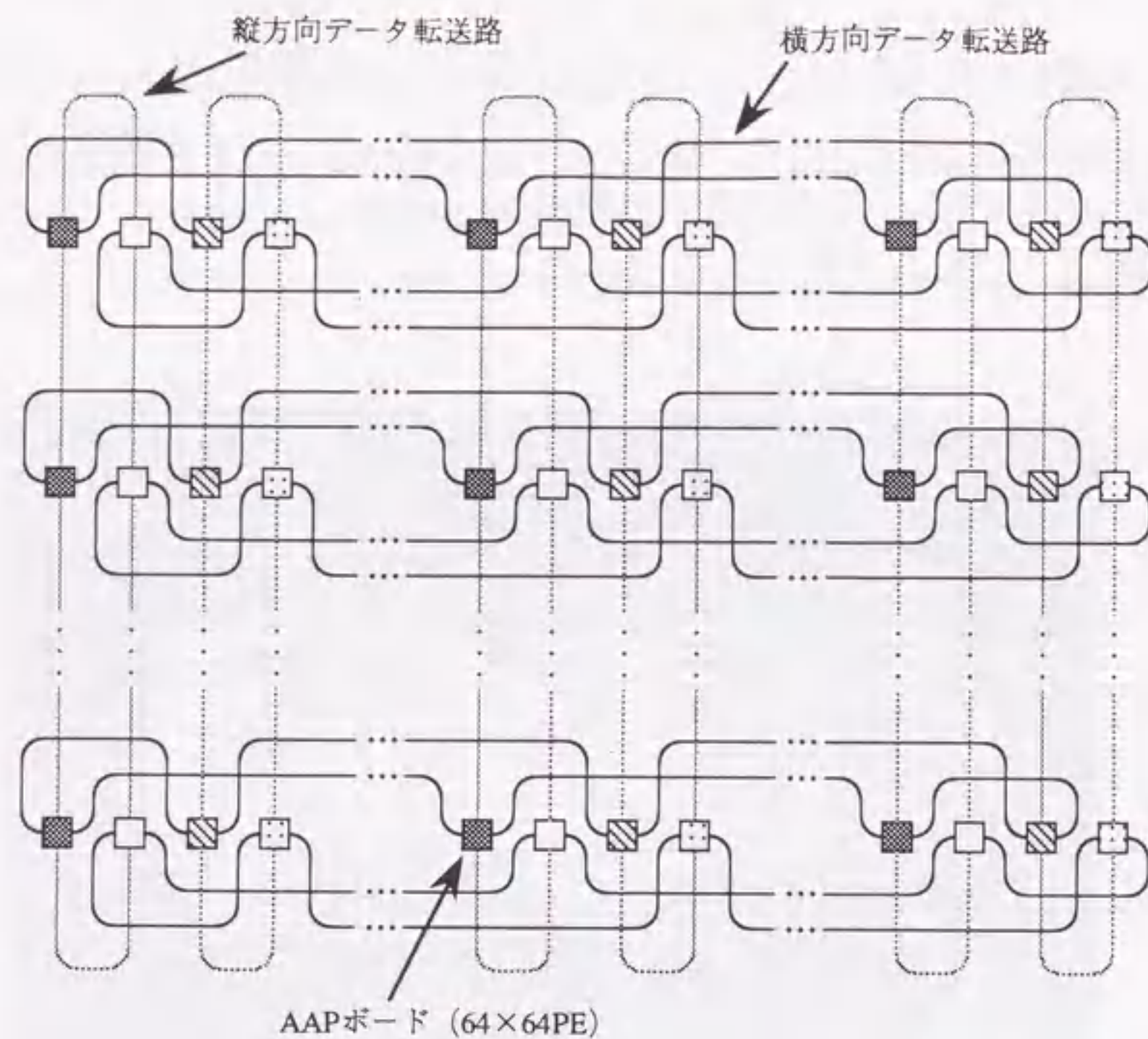


図5. 16 ボード間接続図

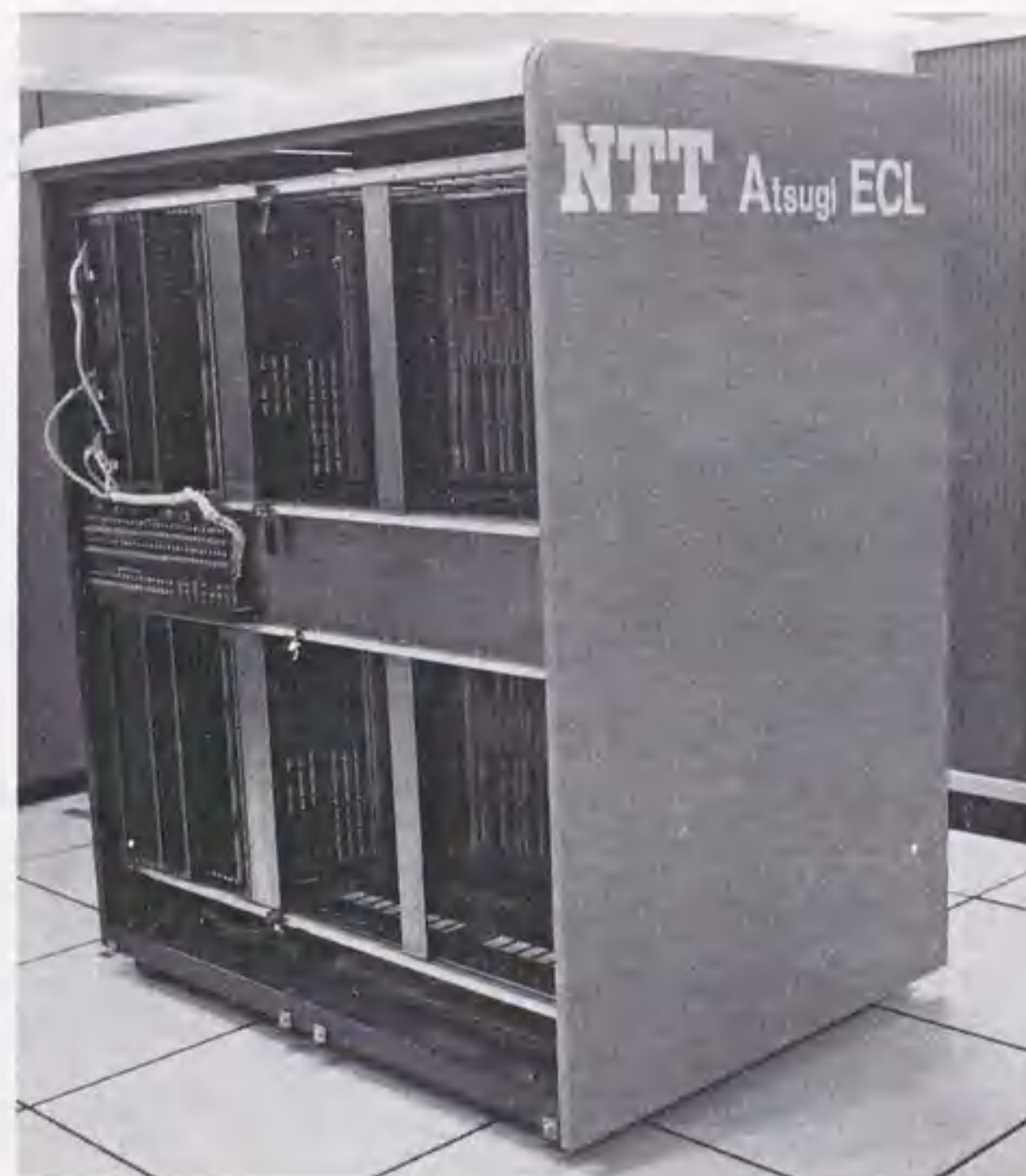


図5. 17 AAP2システム

基板間の接続は、モジュラリティ確保のために、図5. 16のように接続している。この接続構成により、実装時に特定部分の接続距離増大のないボード間接続距離均一の条件を満たすトラス結合を実現している。また、接続の規則性が高く、ボードの増設により、PE配列のサイズを容易に拡大することができる。

試作した装置のきょう体(図5. 17参照)は、制御ユニット、データバッファメモリに加え、256x512までのPE配列が実装できるようにしている。

5. 4. 4 ソフトウェア開発環境

AAP2は専用指向が強く、高速化可能な処理は限定される。しかし、その配列データに対する超高速性をうまく引出すことができれば、効果は非常に大きい。このための試行錯誤が効率良く行えることが必要である。このような観点から、AAP2のプログラム開発環境として、APL風の専用言語AAPLとIBMの汎用コンピュータ上で動作するコンパイラを開発した⁵²⁾。AAPLは、AAPのビットプレーン単位の並列演算と伝搬演算機能を、APLにならい、それぞれ2次元配列間の演算として簡潔に表現可能とするとともに、伝搬演算機能を『走査』の一部として表現できるようにしている。また、AAP2の役割の異なる4つのメモリ領域に対応する変数形式を設け、効率のよいプログラムが記述できるようにしている。

5. 4. 5 基本性能

各PEは局所メモリ内にある2つのオペランドの間でビット直列型の論理あるいは算術演算を行い結果を局所メモリ内の別の領域に得る基本的な演算をビット当たり3クロックサイクルで実行できる。これをもとに例えば8ビットの固定小数点データに対して、PEの稼働率が100%の場合の加算速度を求めると表5. 6の試作したハードウェアの諸元を示すように7, 355MOPS(Mega Operation Per Second)となる。

5. 5 むすび

1ビットプロセッサの配列からなるセルラ配列型LSI(AAP2-LSI)、さらに、このAAP2-LSIを用いた2種類のセルラ配列型SIMDプロセッサLISCAR1(Line Scanable Cellular ARray processor 1)およびAAP2(Adaptive Array Processor 2)を、設計、試作した。

AAP2-LSIは、セルラ配列構成の設計上の利点をすべて引出すことを主眼に以下の3つの設計手法

- ①1ビット構成で、かつ算術論理演算機能、PE間のシフト転送機能、最も単純なバ

イパス型伝搬演算機構等の必要最小限の機能、機構を組込むにとどめ、PEの単純性およびPE配列の規則性を確保する方法、

②トランジスタ当たりの実現機能の高いnチャネルのMOSトランジスタ単独で構成するバストラジスタ論理を多用する方法、

③ $N \times N$ のPE配列を $(N/2) \times (2N)$ 配列としてPEのレイアウト形状を縦長とすることによって共通の制御信号をPE上にはしらせ、配線専用領域を低減する高密度のレイアウト方法、

を採用して、スタンダードセルベースのフルカスタムLSIの4~5倍の高集積度を達成し、 $2\mu\text{m}$ のCMOS技術で $8.8 \times 8.16\text{mm}^2$ のチップに 8×8 PEの搭載を実現した。

LISCAR1は、小型のセルラ配列型SIMDプロセッサをめざし、4個のAAP2-LSIに16ビットの1チップDSP、ビットスライスALU、標準DRAMを組み合わせることによって、B4版サイズのボード1枚に実現した。その性能は、長さが256のビットライン単位の算術論理演算で2,860MOPSの性能を、8ビットのグレイスケールライン間の差分絶対値累積処理(シティブロック距離計算)では180MOPSの性能をそれぞれ得ている。2値画像に対する性能も非常に高く、基本的な演算では数ns/画素以下で、平滑化や90度回転でも数十ns/画素で実行できる。

AAP2は、超高速のセルラ配列型SIMDプロセッサをめざしたもので、1,024個のAAP2-LSIと8,192個のSRAMからなるPE配列部に、主にディスクリートのSSI、MSIで組み上げた制御部、SRAMで構成したインストラクションメモリ、データバッファメモリ等を組み合わせることで構成している。PE配列部は、AAP2-LSIチップとSRAMを直付けするQFPタイプのセラミックモジュールを両面実装した多層ボードと、そのボード間を均等距離でつなぐ接続構成とで、わずか 0.2m^3 のスペースに実現した。性能は8ビットの固定小数点データに対する加算で最大7,355MOPSの性能を得ている。

第6章 8ビットセルラ配列型SIMDプロセッサ システムの設計と試作

6.1 はじめに

LISCAR1、AAP2で採用した単純で規則性の高い1ビット構成のPE配列は、大規模かつ高集積のLSIが容易に設計できる点で優れている。しかし、処理データの語長が大きくなると演算の効率が低下する、回転や間接アドレッシングなどの汎用化に必須とも言える機能が効率良く実行されない等の欠点があり、応用対象が限られたり、性能/コストの面で並列処理の利点を活かさない場合がある。

これに対し、最近ではLSIの設計技術が着実に進歩し、そのおかげで、多ビット構成のPE配列も比較的容易に設計できるようになっており、AAP2-LSI程には単純な構成にこだわる必要がなくなってきた。そこで、AAP2-LSIの欠点を克服すべく、従来からのAAPの可変構造に加え、第2章で導いたセレクトティブツリー構成伝搬演算機構、第4章で示した転送回転型メモリアクセス機構を組んだ8ビット構成のPE配列を搭載したAAP3-LSIを開発した。また、種々の機器へ手軽に組み込み可能な程度までの小型化、経済化を指向したセルラ配列型SIMDプロセッサLISCAR2を試作した。

本章では、このLISCAR2の実現上の鍵となったPE配列部用LSIを中心に、ハードウェア構成、基本性能評価等について論じる。

6.2 小型経済化のためのアプローチ

小型経済性を確保するには、ハードウェア規模の低減に加え、部品コストの低減も必要である。そこで、アーキテクチャを設計する上で、次の3つの構成条件を設定した。

①PE配列部は、一種類の専用LSIと局所メモリ用のDRAMを規則的に並べるだけで構成する。

②各メモリは原則としてDRAMで構成する。

③PE配列部以外の構成には極力汎用のLSIを活用する。

ここで、①は、全構成ゲート数の大半を占めるPE配列部のハードウェア規模低減のための条件である。この条件を満たすには、プロセッサ間転送用ネットワークや2次元アクセス機構等もPE配列とともに1チップに搭載しなければならない。このために、端子数やハードウェア規模が制限されるが、許される範囲内で、極力、性能や汎用性の向上をめざす。

②は、大容量のメモリを小型経済的に構成するための条件である。DRAMは、

SRAMに比べ、速度性能が劣るので、性能向上の制約となる。この制約に対しては、アクセス幅の拡大、メモリアクセスと演算の並列実行等により対応する。それでも、プロセッサの性能に追いつかない分については、演算性能がアクセス速度に見合うまで、プロセッサ側の構成を単純化する。高性能化よりも小型経済化を優先するからである。

③は、ハードウェア規模低減の効果が小さい部分について、開発コストを低減することで、間接的に経済化をはかるためである。

6.3 SIMDプロセッサ構成

6.2節の方針に基づいて設計した本SIMDプロセッサの全体構成を図6.1に示す。本プロセッサは、並列演算部であるPE配列部と、制御部、データメモリ、アドレス生成器等からなる標準的な構成のSIMDプロセッサである。制御部は、コントローラに16ビット固定小数点DSPを用い高速の制御機能とスカラ演算器機能を小型経済的に実現している。マイクロプログラムメモリ(80ビットx64KW)とデータメモ

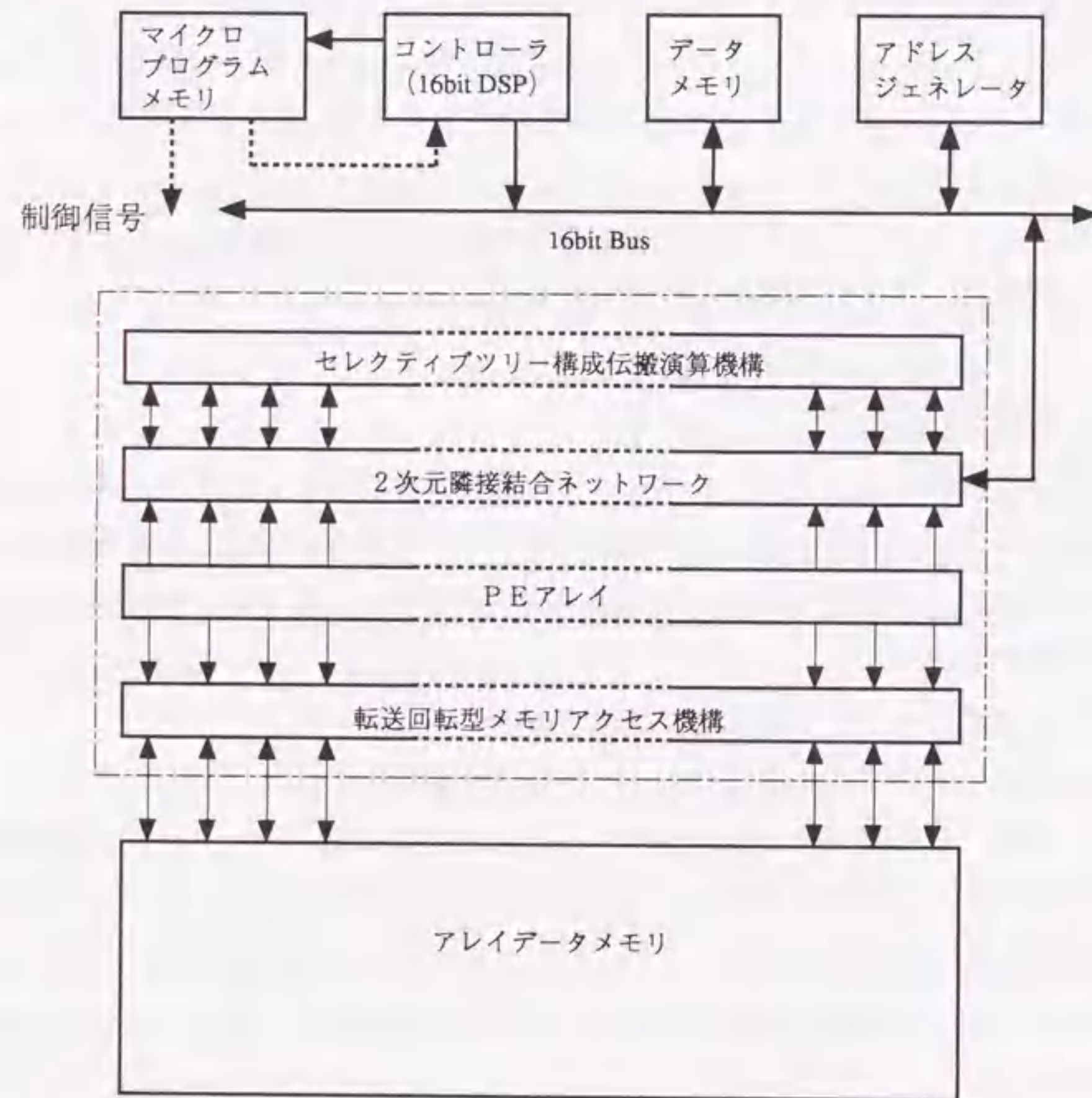


図6.1 LISCAR 2のPE配列のブロック構成

リ(16ビットx1MW)は1Mないし4MビットのDRAMを用いることで小型化を図っている。以下、本プロセッサの性能上のかなめであるPE配列部の構成とその構成要素であるプロセッシング要素(PE)の構成について説明する。

6.3.1 PE配列部の構成

このPE配列部は、図6.1に示されるように8x8の2次元接続のPE配列にセレクトティブツリー構成の伝搬演算機構と転送回転型のメモリアクセス機構を付加した構成を採っている。セレクトティブツリー構成の伝搬演算機構は、図2.12に示す3個の2入力1出力のセクタからなる伝搬演算エレメントの縦続接続である伝搬演算ユニット(SPU)をAAP3-LSI内でツリー状に2段接続して構成している(図6.3のSTN部参照)。LSI間については、外部の部品点数を減らすために、ツリー型ではなく、単純な縦続接続としている。また、転送回転型メモリアクセス機構(TRMAP)については、図4.3に示す転送回転型メモリアクセス機構をそのままAAP3-LSIごとに組込んでいる。

6.3.2 プロセッシングエレメント(PE)の構成

図6.2にPEのブロック構成を示す。2個の転送演算部が互いに共有するレジスタファイルRF、IOBで結合される複合的な構成を採っている。ここで、RFは主に演算に用いるのに対し、IOBはLSI外部に設ける配列データメモリを効率良くアクセスするのに用い、配列データメモリからのロードストアを他の演算と並行して実行することができる。

一方の転送演算部は、画像処理や文字認識処理において、処理の大半を占める演算精度に合せ、演算幅を8ビットに設定している。もう一方は、PE内のビット処理と、PE間のビット転送をスムーズに行うために設けたもので、演算幅を1ビットに設定している。1ビットの転送演算部は、もちろん8ビットの転送演算部に組込むことが可能であるが、このように独立した構成を採ることにより、元になったLISCAR1との互換性を考慮した高度のビット処理機能、ビット転送機能を比較的簡単に設計することができる。

8ビットの転送演算部は、ALU、シフタ、ルーティングレジスタ(RTRG)、乗算器等で構成している。乗算器は、ALU、RF等による累積処理と並行に動作できるようにしており、識別処理やニューラルネットワークのシミュレーションに必要な積和演算を効率良く実行できる。

1ビットの転送演算部の特徴は、2個の1ビットALUを持つことである。これは、

セレクトティブツリー構成伝搬演算機構に対し、先行的な2組の演算結果を供給するために設けたものである。

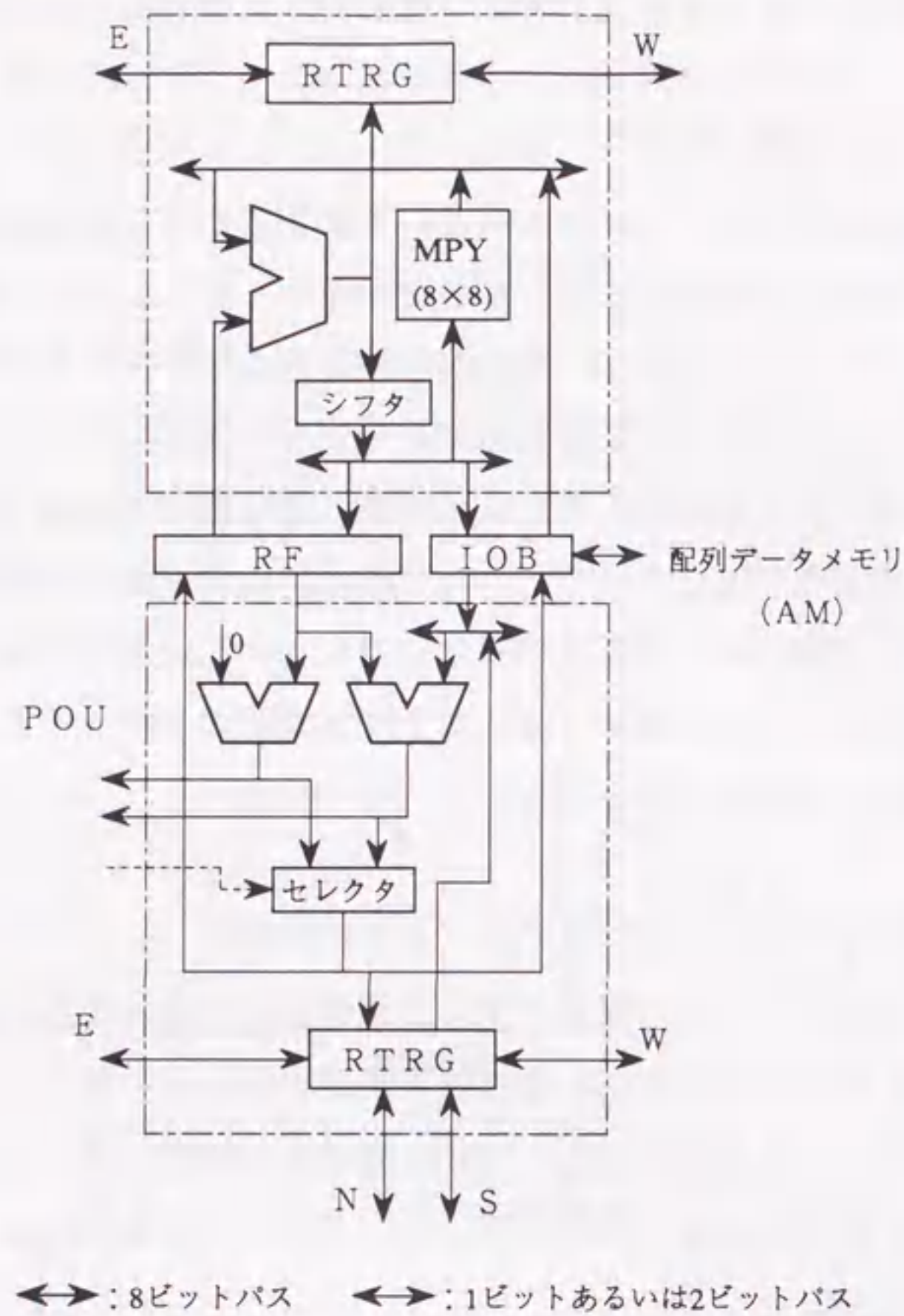


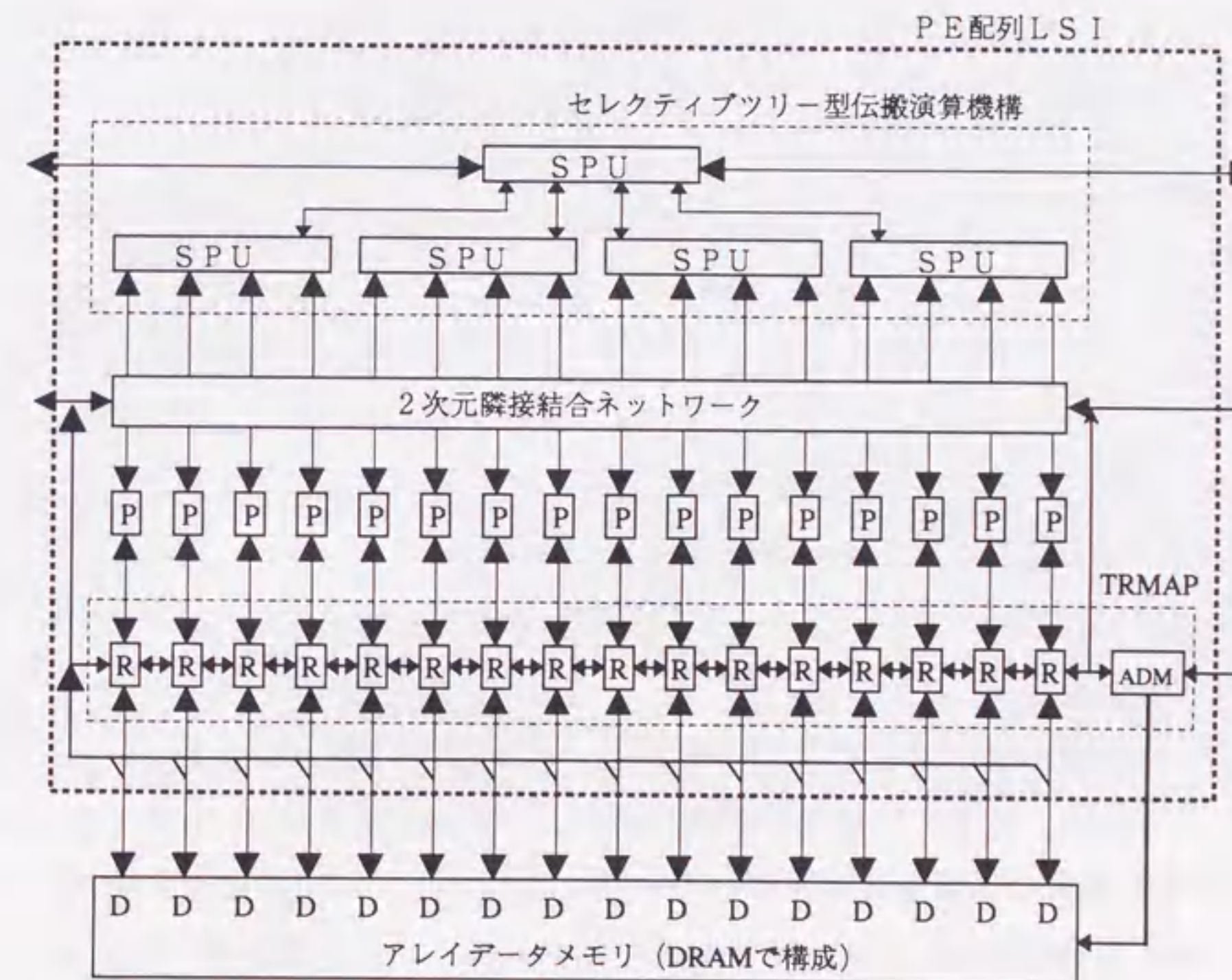
図6.2 PE構成:8ビットバス :1ビットあるいは2ビットバス

6.4 PE配列LSIの設計

この種のSIMDプロセッサのPE配列部は、通常、フルカスタムLSIで構成される。これは、PE配列部が、性能、小型経済性の決め手となるからである。しかし、標準DRAMの使用を前提にすると、データのアクセスネックからフルカスタム構成により得られるLSI内部の高い演算性能を活かし切ることは容易でない。そこで、開発コスト、開発期間の面で有利なゲートアレイを用い、標準DRAMのアクセス速度にバランスする性能の達成をめざした。ゲートアレイは、ローゲート数128Kの0.8 μ m Al2層のCMOS SOG (Sea Of Gates) を用いた。LSIの最小サイクルタイムは、標準DRAMを高速ページモードで動作させる場合のサイクルタイムに適合する55ns (保証値) が達成された (表6.1参照)。

表6.1 LSIの諸元

PE数	16
サイクルタイム	55 ns (保証値)
電源電圧	5 \pm 0.5 V
パッケージ	160 pin QFP
プロセス技術	0.8 μ m Al 2層 CMOS
基本セル構成	SOG
セル数	128 kセル
チップサイズ	11.76 \times 11.82 mm ²



- SPU : セレクトティブ伝搬ユニット (Selective Propagation Unit)
- P : 8ビットプロセッサ
- R : ルーティングレジスタ (Routing Register)
- ADM : アドレス修飾器 (Address Modifier)
- TRMAP : 転送回転型メモリアクセス機構
(Path for Transmission, Rotation, and Memory Access)

図6.3 AAP3-LSIの構成

6. 4. 1 全体設計

PE配列LSIの構成を図6. 3に示す。このLSIを規則的に並べるだけで、PE配列部全体を構成できるように、プロセッサ部分配列と共に、セレクトティブツリー構成伝搬演算機構、2次元の隣接接続ネットワーク、転送回転型メモリアクセス機構等をすべて搭載している。ただし、この図では理解し易くするため、図6. 1と同様に、プロセッサ部分配列と、2次元隣接接続ネットワーク、転送回転型メモリアクセス機構を別々に示しているが、実際にはハードウェア規模低減のために、PE部分配列に一体化する形で構成している。

6. 4. 2 PE設計

ゲートアレイでは、ハードマクロやPE配列の規則性を利用したレイアウトの高密度化は困難であり、同一チップサイズのフルカスタムLSIに比べると、搭載可能なゲート規模は数分の1以下と小さい。従って、全体のゲート数の大半を占めることになるPE配列部の構成単位であるPEについては、ハードウェア規模低減が必須といえる。ここでは、高速化との両立の観点から特徴的なレジスタファイルと乗算器の構成について説明する。

レジスタファイルは、8ビットx32ワード構成と通常のレジスタを組み合わせる構成法ではLSI全体で見るとかなりのゲート数になる。そこで、次の2つの方策、

- ①アクセス機能を制限して、同一サイクルに1オペランドのみ読みだす構成とし、レジスタファイルの単純化をはかる、
- ②64ビット幅のハードウェアマクロのRAM2個を、各PEが8ビット分づつ分割して用いる

によりゲート規模の低減をはかっている。

また、乗算器は、大半の処理で被乗数を配列データメモリからロードするために、そのロード速度以上の演算速度は不要との判断に基づいて、ハードウェア規模低減をはかっている。具体的には、8ビットの加算器とブースのリコーダを用いたシリアル型の構成とし、8x8の乗算を、被乗数のロード時間に等しい4クロックサイクルで実行できるようにしている。

6. 4. 3 セレクトティブツリー構成伝搬演算機構の設計

図6. 3に示されるように、伝搬演算をビット単位で行うビット直列型のセレクトティブツリー構成の伝搬演算機構を搭載している。LSI内で閉じたツリーとすることによって、LSI間を単純に隣接結合するだけで、伝搬演算系を構成できるようにしてい

る。なお、このビット直列型のセレクトティブツリー構成伝搬演算機構を各ビットに設ける並列構成により、伝搬論理演算の高速化が可能であるが、ここではゲートアレイのハード規模制約から採用していない。

6. 5 並列プロセッサシステムの実現

外部バスインタフェース部を含む並列プロセッサ全体を、VME規格のボード1枚に実現した。さらに、この1ボードプロセッサ複数枚を、ワークステーションのバックエンドプロセッサとして動作させる実験システムを開発した。

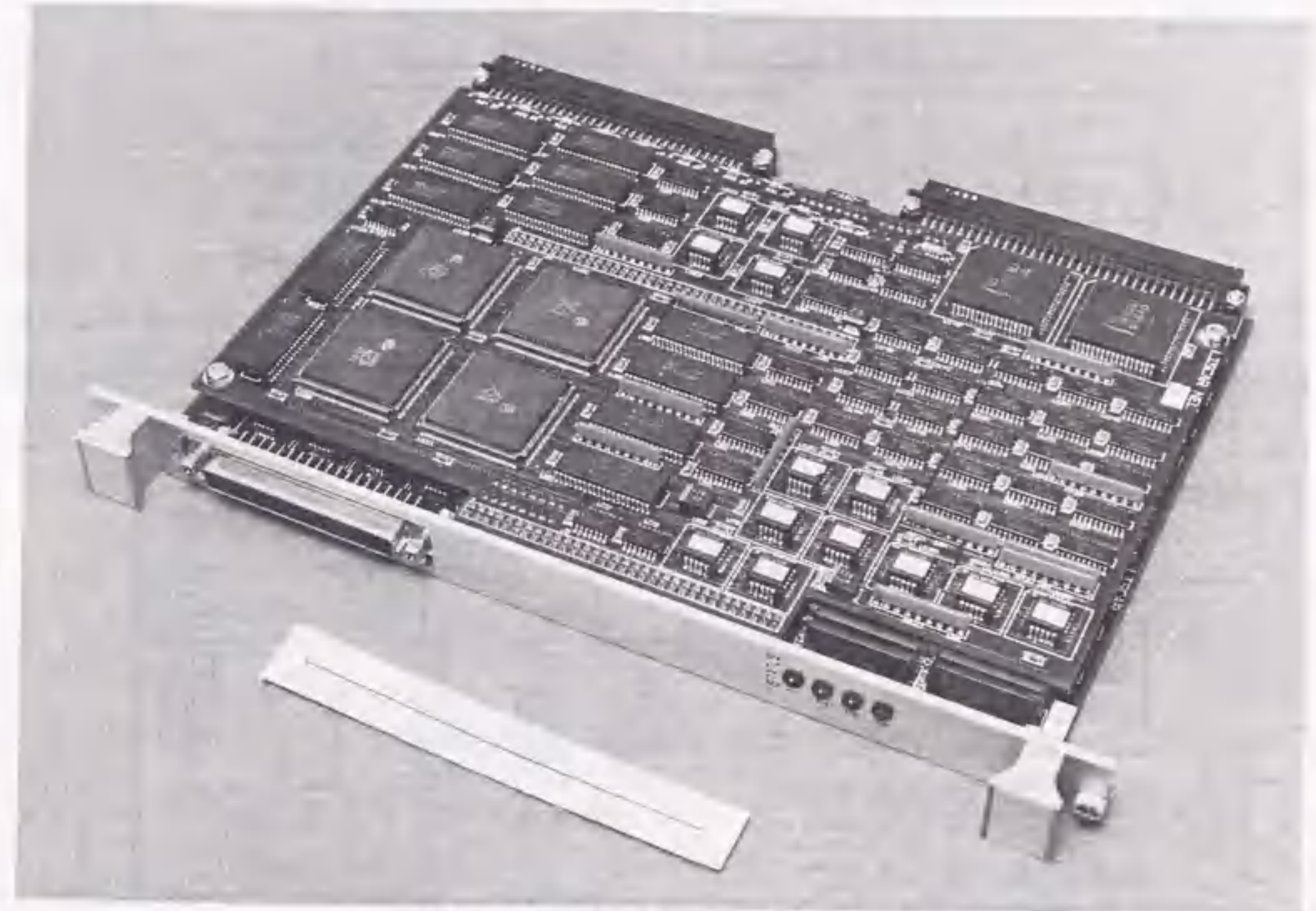


図6. 4 LISCAR 2 ボード

6. 5. 1 ボード

図6. 4に試作したボードの写真を示す。主に、外部バスインタフェース部、制御部を搭載したマザーボードと、PE配列部、アドレス生成部、マイクロプログラムメモリ等を搭載したドーターボードとを、積み重ねて構成している。配列データメモリと80ビット幅のマイクロプログラムメモリには、高速ページモードで動作させる16ビット幅の4Mおよび1MビットDRAMを用いて、経済化をはかっている。また、アドレス生成部は16ビットスライスのALUを用いて構成し、アドレス生成以外の演算にも利用できるようにしている。ボードの諸元を表6. 2に示す。

表6.2 ボードの諸元

サイクルタイム	62.5 ns
搭載LSI数	4
メモリ容量	
配列データ	4 MB
データ	2 MB
プログラム	640 KB
ボードサイズ	233.3 × 16 mm ²
ボード規格	VME

6.5.2 システム

図6.5に、試作したシステムのブロック構成を示す。ワークステーションからVMEバスを介して、複数のプロセッサボードを動かす構成を採っている。適当な周辺ボードと組み合わせることによって種々の用途に適用可能である。ワークステーションには、LISCAR用から記述性の向上をはかったマクロアセンブリ言語に対応するマクロアセンブラとシンボリックデバッガを搭載している。

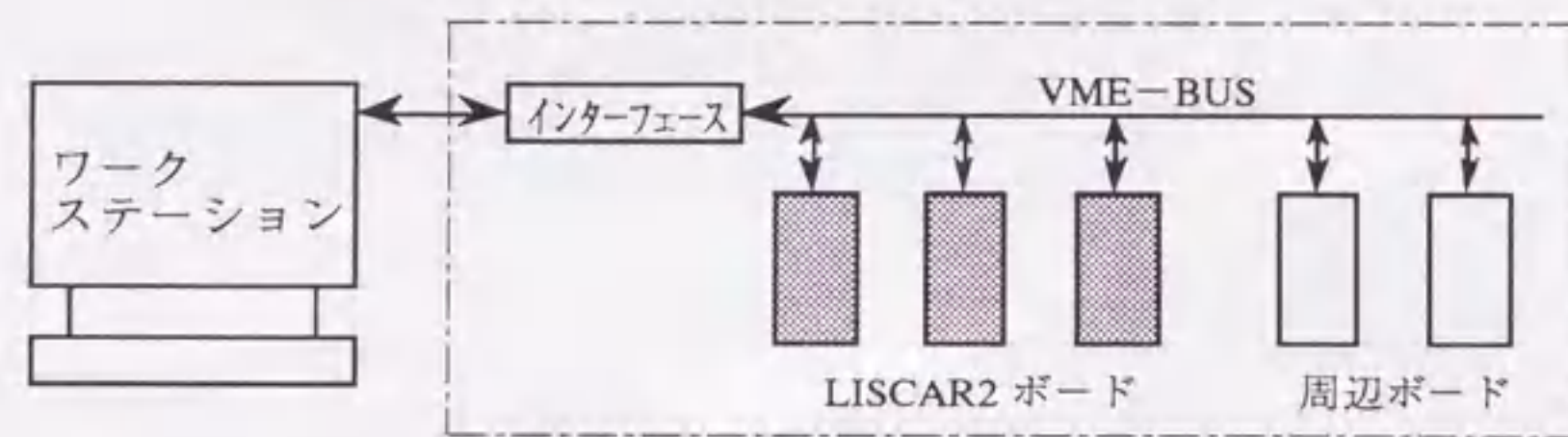


図6.5 LISCAR2システムの構成

6.6 基本性能評価

基本演算性能を示すとともに、パターンデータ処理の重要な演算の1つである90度回転性能を明らかにした。

6.6.1 基本演算

表6.3に基本性能をまとめた結果を示す。配列に対する算術演算、論理演算の性能は、並列処理のおかげで非常に高い。ALUと並列に設けた乗算器により、高い積和演算性能を確保している。表6.4に示す例からも明らかのように、これらを組合わせて

行う配列演算の処理時間もかなり短い。乗算器の効果の出るユークリッド距離計算では、533MOPSとLISCAR1の25倍以上の性能が、乗算器の効果の出ないシティブロック距離計算でも379MOPSとLISCAR1の2倍以上の性能が得られている。この高い基本演算性能を活かすことによって、ニューラルネットワーク処理では20MCUPSの性能が得られる⁵³⁾。

また、伝搬演算についても、セレクトティブツリー構成伝搬演算機構のおかげで、ゲートアレイを用いながら、2ns/PEと、PE間を直接伝搬させるLISCAR1に比べ1桁高速化されている。

表6.3 配列データに対する基本性能

演算内容	性能	精度
算術論理演算	1024 MOPS	8ビット固定小数点
乗累算	512 MOPS	乗算：8ビット固定小数点 累算：32ビット固定小数点
伝搬演算	2 ns/PE	ビット単位

表6.4 基本的な配列演算に対する処理時間

配列演算の内容	演算性能 (MOPS)	演算時間 (ms)	演算の条件	
距離計算	シティブロック距離	379	8.3	配列要素の語長：8bit ベクタの次元数：256 参照ベクタの数：4,096
	ユークリッド距離	533		
行列・ベクタ間の乗算	357	2.2	配列要素の語長：8bit 行列のサイズ：1,536×256 ベクタの次元数：4,096	

6.6.2 90度回転処理

90度回転は、4.4.5で述べたように転送回転型メモリアクセス機構を用いて、効率良く行える。図6.6は、その処理アルゴリズムの概要を示したものであり、

- ①図の上部に示す64×64の2値データ配列（ビットプレーン）から、対角方向に並ぶ16×16の部分ビットプレーン4組（図ではA, N, K, H）を、90度回転をかけながら、TRMAPのルーティングレジスタ（RTRG）配列に読みだす、
- ②読み出した部分配列4個を、左方向に1個分シフトする、

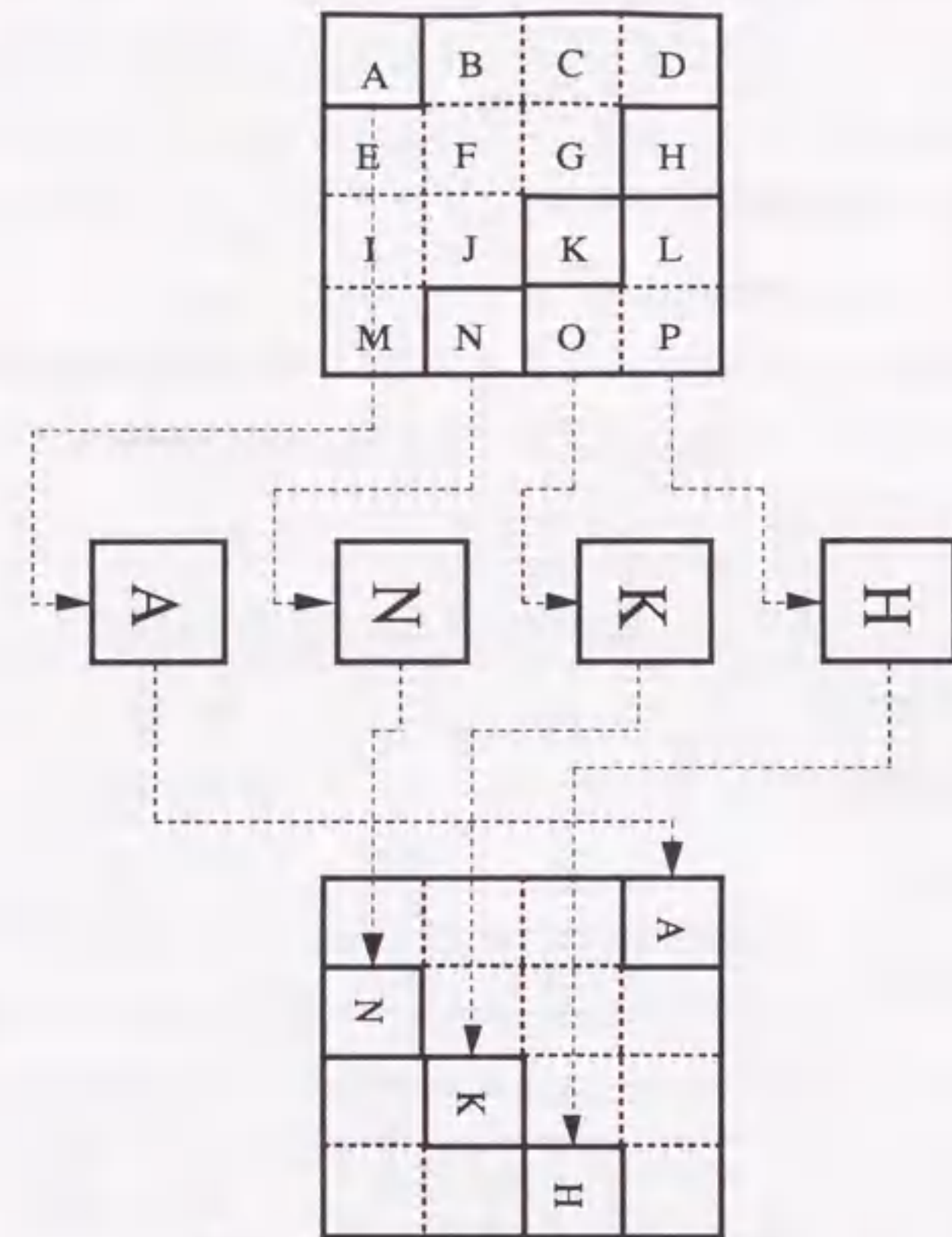


図6.6 90度回転アルゴリズム

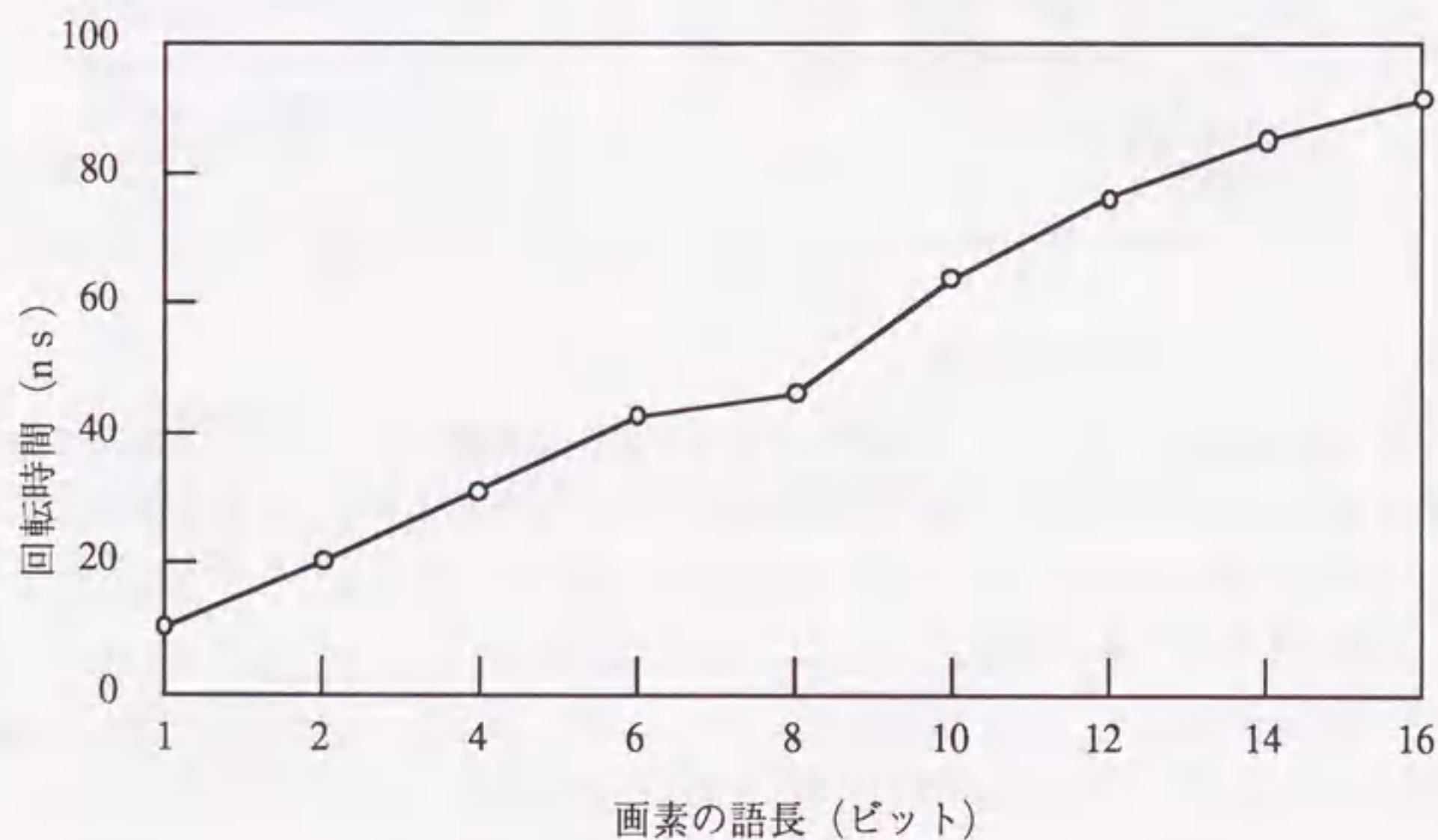


図6.7 画素当たりの90度回転時間

③シフト結果 (図ではN, K, H, A) を読みだし時の対角方向と直交する方向に書き戻し、図下部の回転後の64x64のビットプレーンを形成する

の3ステップを、読みだし位置と書き込み位置を更新しながら、4回繰り返すことを行う。図6.7は、このアルゴリズムをもとに作成した回転マクロについて、64x64のプレーンに対する90度回転の処理時間を測定した結果である。この図から明らかなように、転送回転型メモリアクセス機構は、従来の転送機構にわずかなハードウェアの追加で実現されているにもかかわらず、64x64のビットプレーンが回転器への読みだしの64クロックサイクル、回転器からの書き込みの64クロックサイクルのみで実行される理想的な構成の90度回転器の1/3~1/4の速度で、対象画像のグレイトーンによらず回転されている。この結果、LISCAR1の4倍以上の回転速度が達成されている。

しかし、図6.7から明らかなように、要素の語長が大きくなるにつれて速度がかなり低下する。この原因は、転送回転型メモリアクセス機構の部分配列単位の回転が16x16の部分ビットプレーンであるためである。従って、必要に応じて、各PEでRTRGを複数持つ構成とし、部分配列単位で複数の部分ビットプレーンを同時に回転することによって高速化可能である。

6.7 むすび

PE配列部をゲートアレイLSI (SOG) 4個と標準DRAMのみで構成することによって、小型化、経済化をはかった1ボードSIMDプロセッサLISCAR2を実現した。

はじめに、各メモリは原則DRAMで構成する、PE配列部以外は汎用のLSIで構成する等の小型経済化、開発コスト低減のための方針を明らかにした。次いで、これらの方針に基づいて設計したプロセッサ構成、PE配列部用LSI (AAP3-LSI) の構成、ボード構成、システム構成を示した。プロセッサ構成は基本的にLISCAR1のそれを踏襲しながら、性能向上の鍵となるPE配列部について、

- ①PEを1ビット転送演算部と8ビットの乗算器内蔵の転送演算部の複合構成とする、
- ②PE間の2次元の隣接結合ネットワークにセレクトティブツリー構成の伝搬演算機構と転送回転型メモリアクセス機構を組み合わせる

の抜本的な改良を加えるとともに、プロセッサアレイのサイズを256から64に縮小している。AAP3-LSIは、伝搬演算機構、転送回転型メモリアクセス機構を一体化した16個のPEの部分配列からなる構成とし、0.8μmのSOG型ゲートアレイ

の技術により実現している。ボードは、4個のAAP3-LSIと8個の16ビット幅の4MDRAMでPE配列を構成して、プロセッサ全体をB5版サイズのボードに実現している。

さらに、このLISCAR2の性能を評価し、基本的な演算に対しては、算術論理演算性能が1024MOPS、伝搬演算速度が2ns/PEと非常に高い性能の得られること、複合的な演算に対しても、乗算器が有効に働くユークリッド距離計算で、533MOPSとLISCAR1の25倍以上の性能が、乗算器の効果の出ないシティブロック距離計算でも379MOPSとLISCAR1の2倍以上の性能がそれぞれ得られることを示した。また、転送回転型メモリアクセス機構による90度回転の性能を評価し、LISCAR1の4倍以上、理想的な構成の従来型90度回転器と比べても、1/3~1/4の回転速度の得られることを示した。

このようにLISCAR2は、単に算術演算、積和演算の性能が高いだけでなく、セレクトティブツリー構成伝搬演算機構や転送回転型メモリアクセス機構を利用することによって、PE配列の局所ごとの総和演算、PE配列全体に対する放送、90度回転等を効率良く実行できるので、配列演算を主体とする応用では、専用構成のハードウェアに匹敵する小型化、経済化が達成されよう。

第7章 大規模セルラ配列型SIMDプロセッサAAPのLSI-CADへの応用

7.1 はじめに

5.3節で述べたAAPは、そのセルラ配列構成から明らかなように、大規模な2次元配列データの処理に適している。この意味で、規模の大きな配列データを扱うLSI-CADは、AAPに適した応用分野の一候補といえる。しかし、配列データを扱うといっても、AAPの得意とする短語長の2次元配列データに対する演算のみで処理できるわけではない。むしろ、従来はたとえ処理対象データがもともと短語長の2次元配列データであっても、処理中は、ベクトルデータに変換して処理することが多かった。これは、逐次型の汎用プロセッサでは、データ構造の規則性よりも、データ量の少なさ、すなわち演算量が少ないほど、高速に処理できるからである。このような状況は、2次元配列データのまま直接処理を進めるアルゴリズムの発展を疎外してきた。従って、AAPの性能を十分に引出すには、処理アルゴリズムの見直しから行わなければならない。

本章では、LSI-CADの中でも完全性を追究しようとする汎用プロセッサでは処理時間が大きくなりすぎる論理シミュレーションと、セルラ配列型SIMDプロセッサに特に適すると考えられるマスクパターンレイアウト設計における自動配線処理について、1ビット構成のPEからなるAAPに適した処理アルゴリズムを提案する。さらにこれらのアルゴリズムに基づき、AAP1を用いて行った処理実験の結果について述べる。

7.2 論理シミュレーション

論理ゲートをPEで、論理ゲート間の結線をPE間の伝搬転送でそれぞれエミュレートすることにより、AAP上で論理シミュレーションを行うことができる。

図7.1は、8ゲートからなる論理回路を4x4のPE配列に割付ける例を示している。ここで、それぞれの升目はPEを表している。また、未割付のPEは、データ転送専用のPEである。各PEは、次式で示す論理演算を実行することによって、3入力までの任意のゲートをエミュレートすることができる。

$$G_{out}(T+1) = IV_{out} \bigoplus \bigwedge_{i=1}^3 \{ (ID_i(T) \oplus IV_i) \vee MA_i \} \quad (7.1)$$

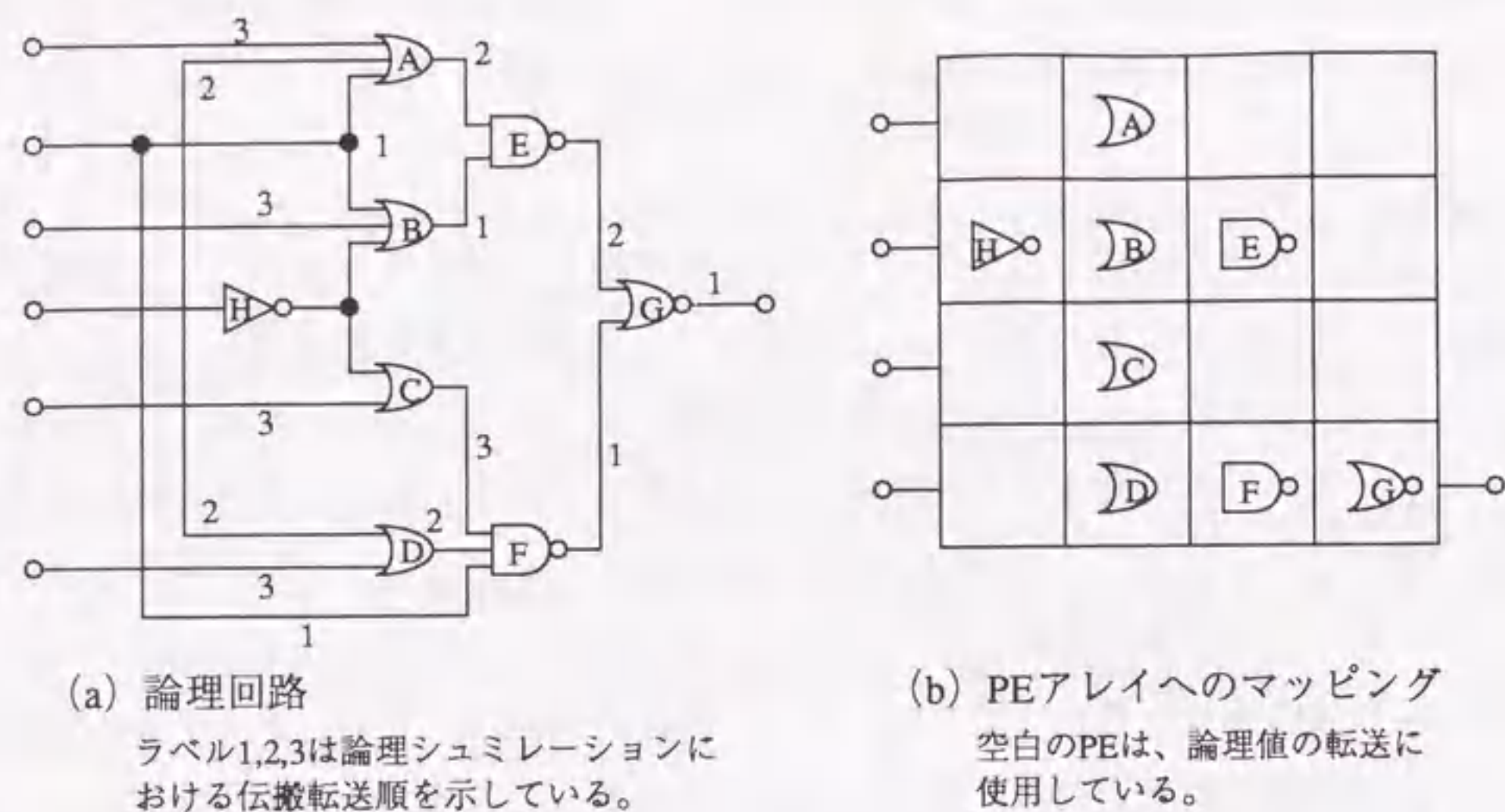


図7.1 論理回路のPE配列へのマッピング

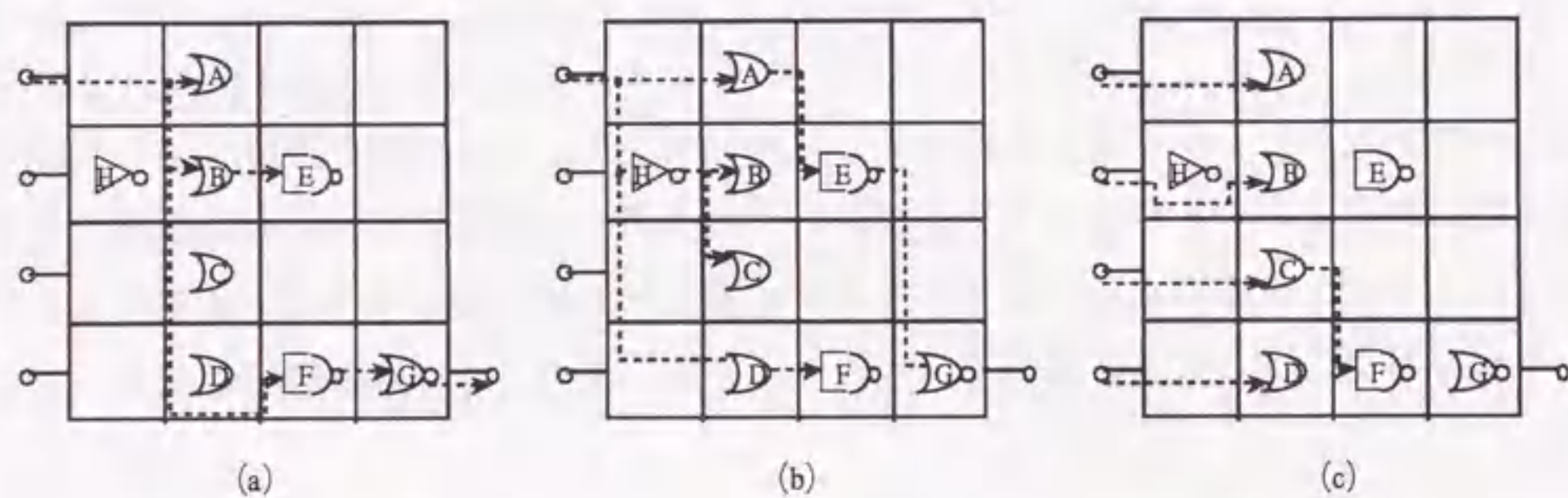


図7.2 伝搬転送を利用しゲート間の結線のエミュレーション
(a), (b), (c)は図7.1(a)のラベル付けされた1, 2, 3の結線に対応

ここで、 IV_i と IV_{out} は、入力信号と出力信号に対する反転制御用の係数データである。 $ID_i(T)$ は、時刻Tにおけるi番目の入力端子に対する入力データ、 MA_i はi番目の入力端子を活性化するか否かを示すマスクデータ、 \oplus は排他的論理和である。また、 $G_{out}(T+1)$ は、時刻T+1における論理ゲートの出力である。論理ゲート間の結線は、図7.1(a)で1, 2, 3のラベルで指定される結線に対応し、右方向と下方向の2回の伝搬演算を組として、図7.2に(a), (b), (c)の順で繰り返すことで、

表7.1 シミュレーション時間の比較

回路の名称		加算器	浮動小数点データ 正規化回路
ゲート数		2,105	337
シミュレーション時間 (ユニット遅延)	AAP1	4.0 ms	0.89 ms
	7MIPS汎用 コンピュータ	MIN 32.0 ms MAX 260.0 ms	MIN 4.2 ms MAX 49.0 ms

エミュレートする。

以上の論理ゲートのエミュレーションと論理ゲート間の結線のエミュレーションを繰り返すことによって単位遅延の論理シミュレーションが実現される。

このアルゴリズムの原理確認のプログラムを、クロック周波数0.67MHzのAAP1に組み込み、シミュレーション実験を行った。その結果得られたシミュレーション性能と7MIPSの汎用コンピュータでのシミュレーション性能との比較を表7.1に示す。汎用コンピュータではイベントの生起に関わるゲートのみ論理値を評価するイベントドリブン型のシミュレーションアルゴリズムを採用しているため、シミュレーション時間のイベントの発生頻度が高い場合の値(上側)と少ない場合の値(下側)は、10~100倍も変化している。これに対し、AAP1は無条件に全ゲートを評価するアルゴリズムなので、イベントの発生頻度に関わらずシミュレーション時間は一定で、汎用コンピュータの5から60倍の性能が得られている。AAP1のクロック周波数が低い高速化の程度は、それほどでもないが、汎用コンピュータ並みの10M~20MHzのクロック周波数が実現されれば、さらに10倍以上の高速化が可能である。また、AAP2では転送方向とALUファンクションのPE毎の個別設定機能を利用する4.3.2の方法で、もう一段の高速化が可能になる³²⁾。

7.3 自動配線

マスクパターンレイアウト設計における自動配線は、パターンデータを扱う点で自動配置⁵⁴⁾とともにセルラ配列型SIMDプロセッサに適した処理といえる。しかし、従来は、汎用コンピュータの使用を前提に、基本的にベクトルデータの世界で逐次処理のアルゴリズムが発展してきた。従って、これらの自動配線処理をそのままAAPに持ってきても高い性能を得ることは難しい。それでも、中には本来並列処理向きで、有効性は明白であっても演算量の多さから局部的にしか使用できなかったり、あるいは簡略化せ

ざるを得ないアルゴリズムも存在した。その代表的なアルゴリズムが最短経路を見つけることのできる迷路法であり、古くから専用ハードウェアあるいは並列プロセッサによる高速化の検討がなされてきた^{55,56)}。

本節ではこの迷路法と処理量の関係から従来より広く用いられてきたライン探索アルゴリズム⁵⁷⁾を組合わせた新しい配線アルゴリズム⁵⁸⁾と、それによる自動配線のAAP1での実験結果について述べる。

7. 3. 1 新配線アルゴリズムの概要

はじめに新アルゴリズムについて3つの端点間の配線を例に説明する。この例の配線処理は次の7ステップを順に実行することで実現される。

ステップ1：端点T1から、昇順のラベルを振りながら波面を最も近い端点T2に到達するまで広げる [図7. 3 (a)]。

ステップ2：ステップ1とは逆に端点T2から、降順のラベルを振りながら波面を端点T1に到達するまで広げる [図7. 3 (b)]。

ステップ3：ステップ1, 2で得られた昇順と降順のラベルの論理積をとり、互いのラベルが同一の領域(網点部)を抽出する [図7. 3 (c)]。この領域は、波面の伝搬の性質から明らかなように、T1, T2間の最短経路の集合である。

ステップ4：ステップ3で得た網点領域の周辺から、昇順ラベルの波面を端点T3に到達するまで広げる [図7. 3 (d)]。

ステップ5：ステップ4とは逆に端点T3から降順ラベルの波面を網点領域の周辺に到達するまで広げる。

ステップ6：ステップ3同様、ステップ4, 5で得た昇順と降順のラベルの論理積をとり、互いのラベルが同一のハッチ領域(ライン)を抽出する。このハッチのラインは、網点領域と端点T3の間の最短経路である。このラインに接する網点領域の点をPとする [図7. 3 (e)]。

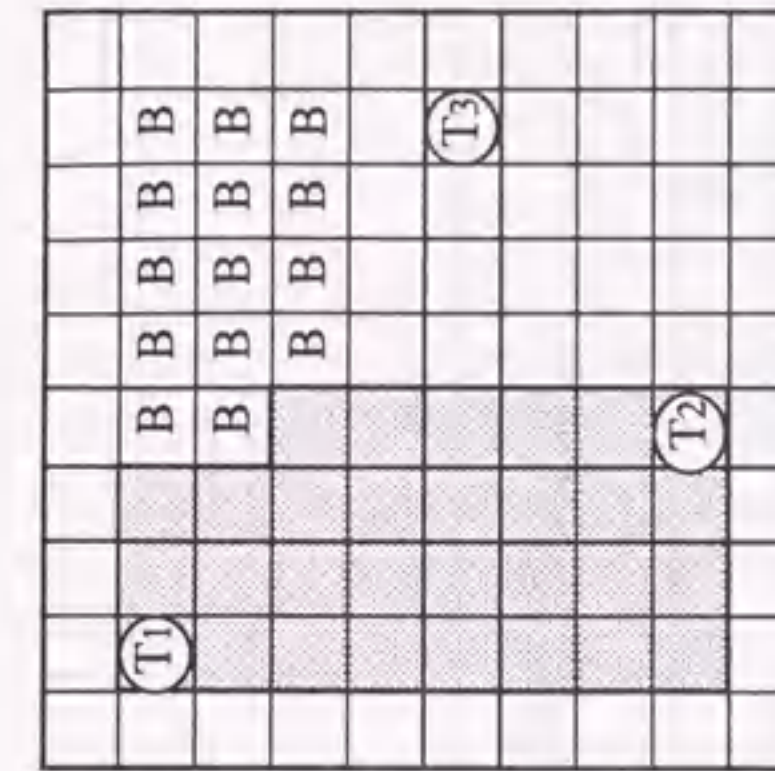
ステップ7：網点領域、ハッチ領域に限定し並列のライン探索アルゴリズムによって、点Pから、端点T1, T2, T3に対する経路を見つける [図7. 3 (f)]。



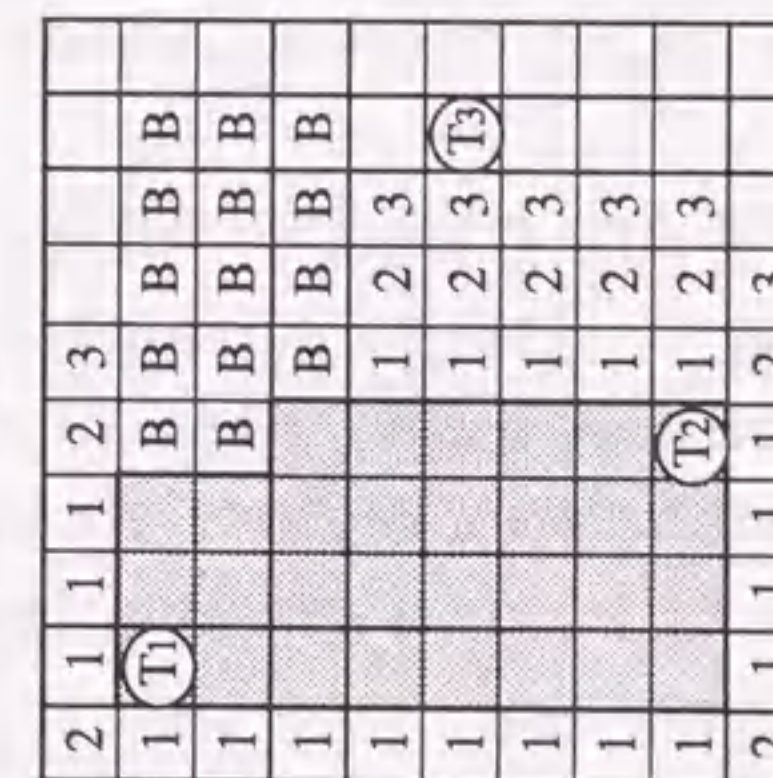
(a) T1から最近傍のT2に向かう波面の拡張



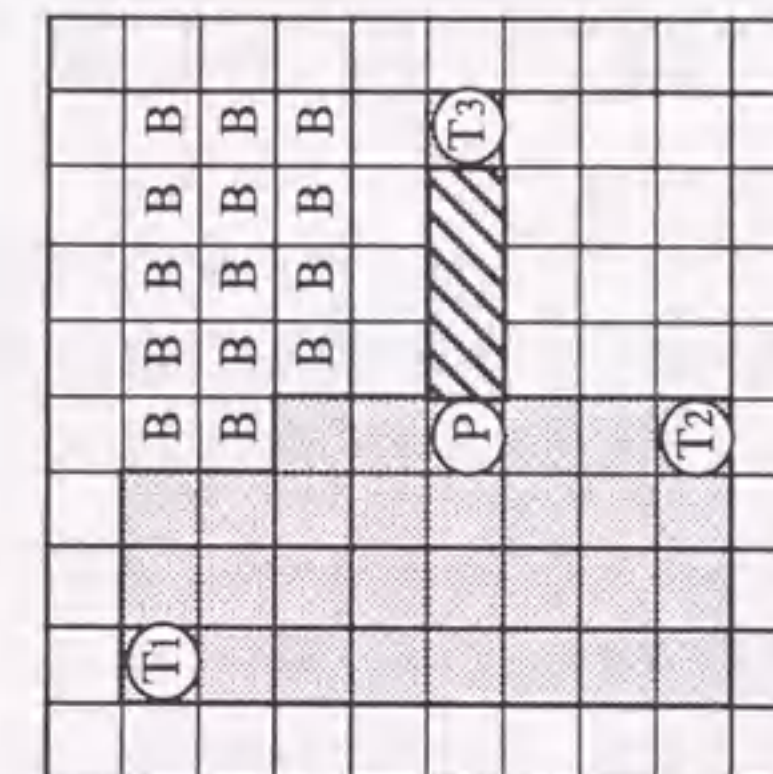
(b) T2からT1に向かう波面の拡張



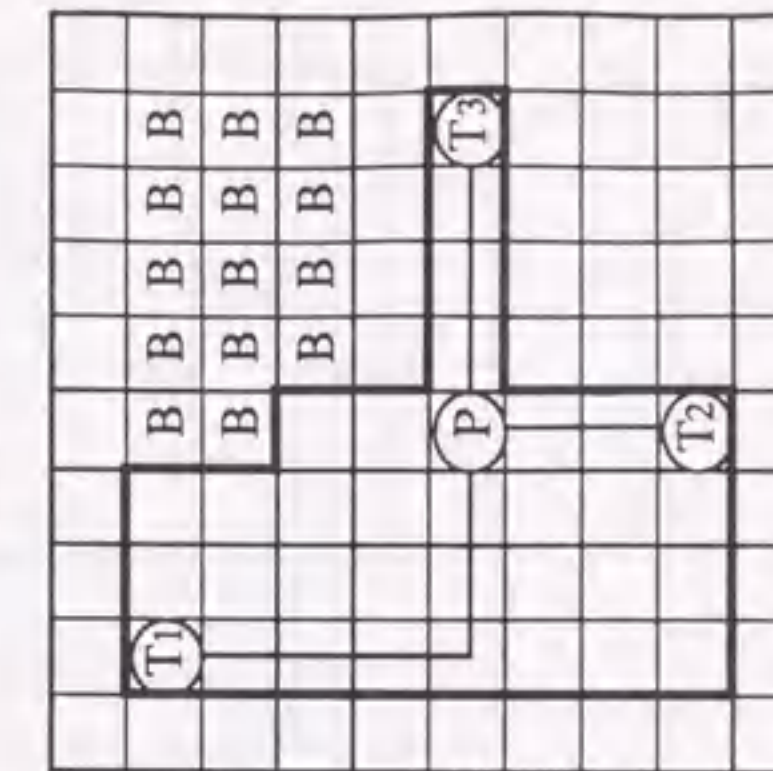
(c) T1, T2間の最短経路領域(網点部)の抽出



(d) 網点部からT3に向かう波面の拡張



(e) T3との間の最短経路領域の抽出



(f) 網点部内でのライン探索アルゴリズムによる配線

図7.3 新アルゴリズムの配線手順

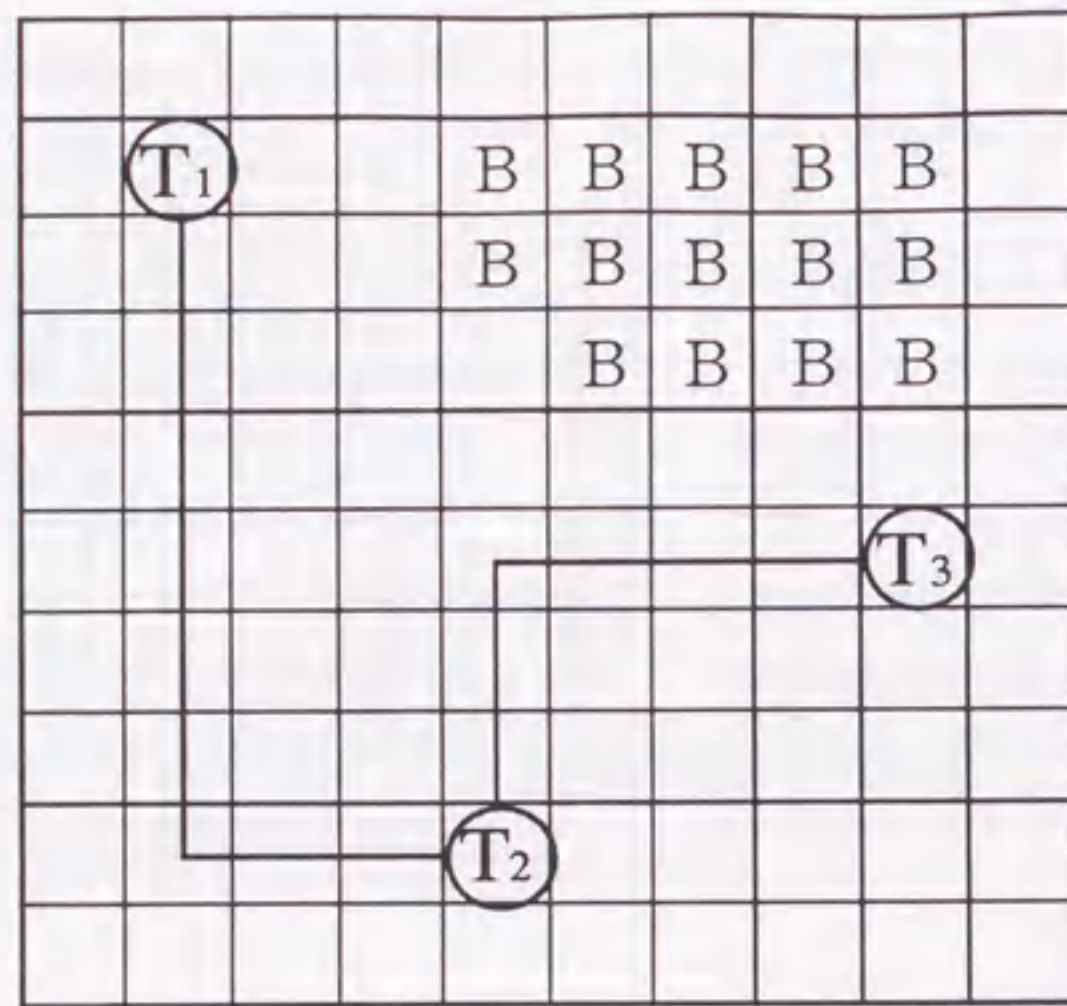


図7.4 ライン探索アルゴリズムによって引かれた結線

ライン探索アルゴリズムは、折れ曲がり数最小の経路を見つける能力を持っている。ハッチ内は長さ最小の経路の集りであることから、図7.3(f)の配線は、長さが最小でかつ折れ曲がり数最小の経路で引かれている。すなわち、配線パターンとしては最も望ましいと考えられている直線状のスタイナー⁵⁹⁾の配線経路が得られている。配線経路の品質の高さは、ライン探索のアルゴリズムのみで求めた場合(図7.4)と比べるとはっきりする。この配線経路では、図7.3で得られた経路とは異なり、経路長も折れ曲り数も最小ではない。なお、新アルゴリズムは、 n 個の端点間の配線に対しても、最短経路の集りからなる領域の抽出とそれに続くライン探索による配線を繰り返すことによって、3点間の場合同様に品質の高い経路を見つけることができる。

7.3.2 AAP1の配線処理速度評価

AAPでは、7.3.1で示した新アルゴリズムを高速に処理できる。これは、迷路法の並列に波面を広げる処理をPE間のシフト転送とビットプレーン間で、ライン探索法のライン単位の拡張を伝搬演算で、それぞれ効率良く処理できるからである。

図7.5はAAP1による迷路法の処理速度を、1MIPSの汎用コンピュータと比較して示した結果である。AAP1のクロック周波数は0.67MHzと低いにもかかわらず、 256×256 の配線領域に対し、100倍以上の処理速度が得られている。

AAP1の新配線アルゴリズムに対する処理速度を、迷路法の処理速度と比較して示す。新アルゴリズムは、迷路法より2倍程度高速に処理されている。これは、新アルゴ

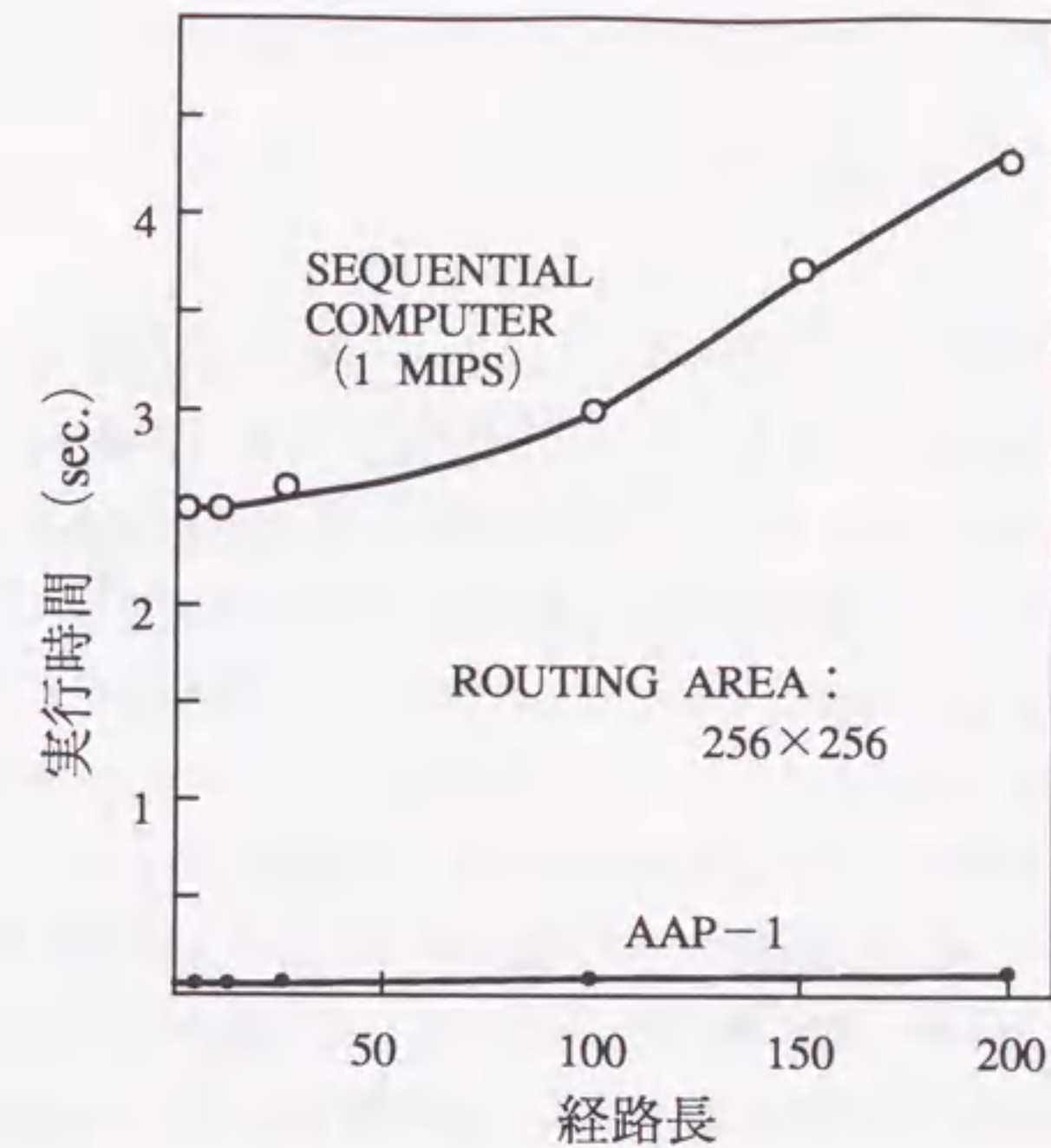


図7.5 迷路法による配線時間

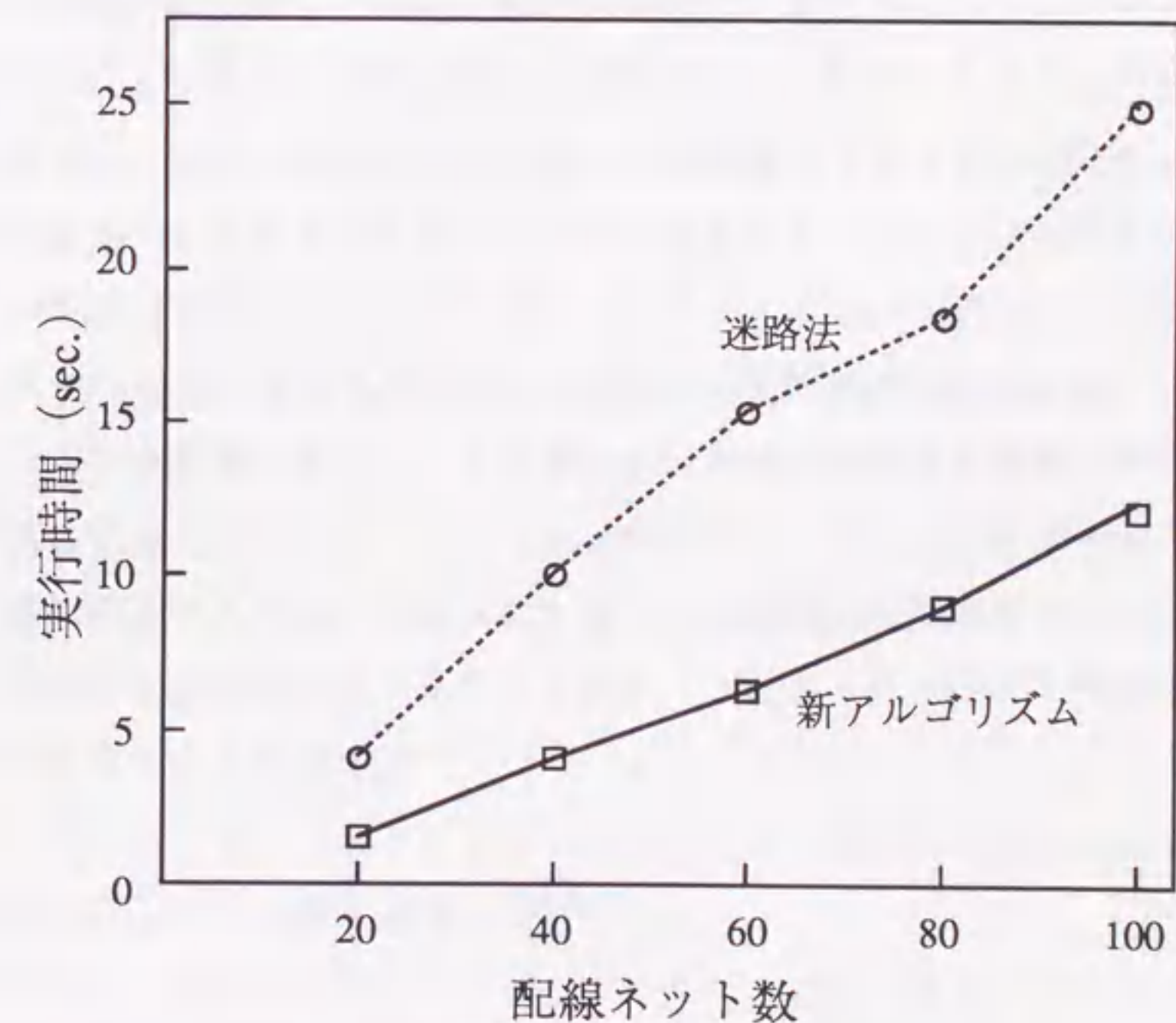


図7.6 AAP1による多端子配線の実行時間(3端子/ネット)

リズムの方が、バックトレース時に必要となる逐次的な演算の量が少なくなるからである。

7.4 むすび

LSI-CADの中でも汎用コンピュータの能力では処理速度や処理品質の要求に答えることが難しい論理シミュレーションとマスクパターン設計における自動配線のAAP1での処理方法を明らかにし、実機を用いてその処理性能を評価した。

論理シミュレーションについては、3入力までの任意の論理ゲートを1個のPEで、論理ゲート間の結線を伝搬転送で、それぞれエミュレートするシミュレーション方法を明らかにし、実機上で汎用コンピュータの5~50倍(クロックサイクルが同一の条件では50~500倍以上)の性能の得られることを確認した。

自動配線については、迷路法にライン探索アルゴリズムを組合わせたセルラ配列型SIMDプロセッサ向きの新アルゴリズムを示した。次いで、その新アルゴリズムがPE間のシフト転送と伝搬演算の活用により効率良く実行され、迷路法が1MIPSの汎用コンピュータの100倍程度の速度で実行されるのに対し、さらに2倍程度高速に実行できることを実機により確認した。

AAP2については、このような実機での性能評価は行っていない。しかし、PEの構成を改良したこと、クロックサイクルが5倍速くなっていること、外付けの局所メモリをPE当たり8Kビット付加したこと等によって、AAP1より1桁以上高速に処理できる見通しが得られている⁵²⁾。論理シミュレーションについては、新たな処理アルゴリズムの検討も行われており、それを用いることでさらに1桁以上の高速化が可能になる³²⁾。

このようにAAPはその超並列構成を活かすことによって、自動配線、論理シミュレーション等の処理を高速に実行することができる。しかし、本章で行った評価は、それぞれ処理の核となる部分に対してのみである。周辺の処理には、処理量自体は多くなくとも、煩雑でかつ並列化困難な処理が存在する。実用的なシステムを構築するには、これらの周辺処理をどのように実行し、全体として高い性能を得るかを解決して行く必要がある。

第8章 小型セルラ配列型SIMDプロセッサ LISCAR1の文字認識処理への応用

8.1 はじめに

OAの進歩と共に重要性が益々高まっている文書画像処理、文字認識処理、文書検索処理等は、2次元あるいは1次元の大規模な配列データを扱う割合の高い処理であり、汎用プロセッサのみでは実用的な処理性能の達成が難しい。このため、布線論理の組合わせを主体とする専用のハードウェア構成により処理装置(例えば光学的文字読取り装置[OCR]^{60,61)})の性能向上が、はかられてきた。しかし、この方法では、ハードウェア構成を特定のアルゴリズムに適合化しているために

- ①アルゴリズムの変更に対応できず、異なるアルゴリズムの搭載が困難である、
- ②ハードウェアの開発を、アルゴリズムと独立には進められない

等の欠点があり、装置開発におけるコスト低減、期間短縮等の障害となっていた。

専用ハードウェアによらぬ高速化手法としては、複数のマイクロプロセッサによる並列処理を利用する方法があり、この分野でもいくつかの装置が開発されている。この中には数個の画像処理プロセッサによる汎用的構成の装置の報告もある⁶²⁾ものの、大半は高性能化のためにプロセッサ間でパイプラインを組む専用の構成をとっており、搭載アルゴリズムの可変性は小さい^{63,64)}。

これに対し、小型のセルラ配列型SIMDプロセッサであるLISCARは、

- ①文書画像処理、文字認識処理で大半を占める8ビット以下の短語長の配列データ処理を得意とする、

②汎用的な構成でも構成プロセッサ数を増やすことによって性能向上が可能である等の特性を備えており、小型・高速性と搭載アルゴリズムの可変性を両立できる可能性がある。このような観点からLISCARへの認識アルゴリズムの異なる複数の文字認識処理の搭載を試みた。本章では、これらの処理をLISCAR上に搭載する際の具体的な処理方法を示すとともに、処理性能を評価し、問題点を明らかにして行く。

8.2 手書き英数カナ認識処理⁶⁵⁾

専用イメージスキャナやファクシミリを入力手段とする手書き文字の認識処理システムへの適用をめざし、次の要求条件を設定した。

- ①ファクシミリ入力の文字パターンデータに対し、汎用スキャナ入力の文字認識装置並みの認識精度を達成すること。すなわち、不特定多数が丁寧に筆記した数字に対

して99%程度以上、カナに対して95%程度以上の正読率を達成すること。

②認識速度は、ファクシミリから帳票を入力する際の実行的な文字入力速度を上回る10文字/秒以上であること。

要求条件の①は、解像度が低い上に白黒判定のしきい値を調整できないために、専用スキャナに比べ画質がかなり劣るファクシミリ入力の文字認識装置を実用に供するための条件であり、極めて厳しく、数ある認識アルゴリズムの中でもこの要求にこたえられるものはごく少ない。そこで、これまでの専用ハードウェアに搭載した実績から唯一見通しの立つ輪郭特徴を組合せた位相構造化法⁶⁰⁾に、ファクシミリ対応の前処理を付け加えた認識アルゴリズムを用いることにした。ただし、このアルゴリズムは、専用ハードウェアによる逐次的な処理機構の使用を前提に開発されてきたものであり、LISCARとの適合性は高いとはいえず、今度は要求条件の②をクリアすることが難しくなる。だからといって、アルゴリズムをLISCAR向きに修正することも許されない。これまでにかかりの期間をかけて作成してきた認識辞書に再度手を加えなければならなくなるからである。結局、アルゴリズムに内在する並列化可能な処理を極力LISCARの並列処理部で実行することで、処理効率の向上をはかることが必要になる。

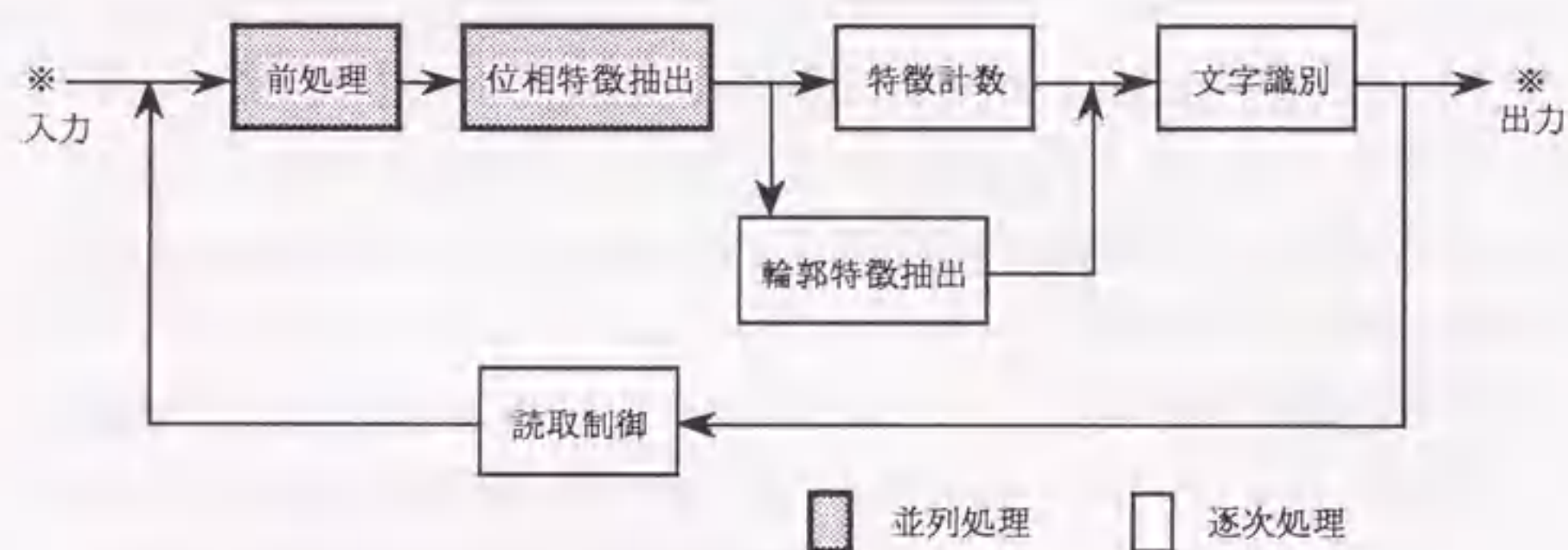


図8.1 位相構造化法による認識処理の構成

8.2.1 認識処理の構成

輪郭特徴を組合せた位相構造化法の処理構成を、図8.1に示す。この図で前処理は、入力の文字パターンに対するノイズ除去、縮小等の処理である。位相特徴抽出処理は、文字パターン上の全画素について、黒点である場合には、その黒点が輪郭点でないか、輪郭点であればその傾きが縦、横、斜のいずれであるか、白点である場合には、上下左右のいずれの方向に黒点が存在するかを求める処理である。特徴計数処理は、得られた位相特徴を分類計数する処理である。輪郭特徴抽出処理は、文字ストロークの輪郭

部の屈曲度と屈曲方向の分布を求める処理である。文字識別部は、得られた特徴を辞書と照合し、認識結果を得る処理である。

8.2.2 並列処理と逐次処理の切り分け

前処理、位相特徴抽出処理は、文字パターンをビットプレーンとして直接処理することから、LISCARのライン単位の並列処理で効率良く処理することが可能と考えられる。

一方、特徴計数、輪郭特徴抽出、文字識別については、並列処理の適用は容易ではない。特に文字識別については、場合分けを繰り返し、辞書を順次たぐることで正解を探す全くの逐次的な照合アルゴリズムを用いている⁶¹⁾ため、SIMD型の並列処理による高速化は期待できない。これに対し、特徴計数、輪郭特徴抽出の一部については、並列処理適用の可能性はあっても、規則性が低い、処理が複雑化する等の理由により、高速化の効果はそれほど期待できないことが明らかとなった。

そこで、特徴計数処理以降については、LISCARのアドレス生成部と制御スカラ演算部を利用した逐次処理で実行することとした。

8.2.3 並列処理の内容

(1) 前処理

前処理で行う主要な処理は、ノイズ除去と縮小である。いずれも文字パターンをビットライン(要素の語長が1ビットの1次元配列)の集りとして配列データメモリAMに格納し、ビットライン単位の並列演算の繰り返しにより実現可能なLISCAR向きの処理である。実際、ノイズ除去は3x3のマスキ演算であり、ビットラインの左右シフトとビットライン間の論理演算の組合わせで実行できる。縮小についても、その要求が文字パターンの高さ方向のみであることから、文字パターンの構成ビットラインを適当な間隔で間引くことによって対応できる。ただし、画素1個分の厚みしかない細線が消失しないように、間引き対象の前後のビットラインを論理演算により加工する。

(2) 位相特徴抽出

位相特徴抽出は、1次位相特徴抽出、統合、高次化の順で行う。1次位相特徴を数字の5を例に説明する(図8.2)。この図で、Qは白点、Pは黒点を示している。白点、黒点はそれぞれ次のような特徴を持つ。白点の特徴は、例えば□が、その白点の上下と左側に黒点が存在することを、└がその白点の下側と左側に黒点が存在することを表す。また、黒点の特徴は、Iがその黒点が文字線の内点であることを、V、H、

Sがその黒点が文字線の輪郭点でかつ傾斜が、それぞれ垂直、水平、斜めの方向であることを、表す。

この1次位相特徴の抽出はLISCARでは次のように行う。すなわち、白点の特徴は、左右方向については黒点の有無を各黒点を始点とする伝搬演算により、上下方向の黒点の有無をビットライン間の論理演算により実現される上下方向の伝搬演算により、

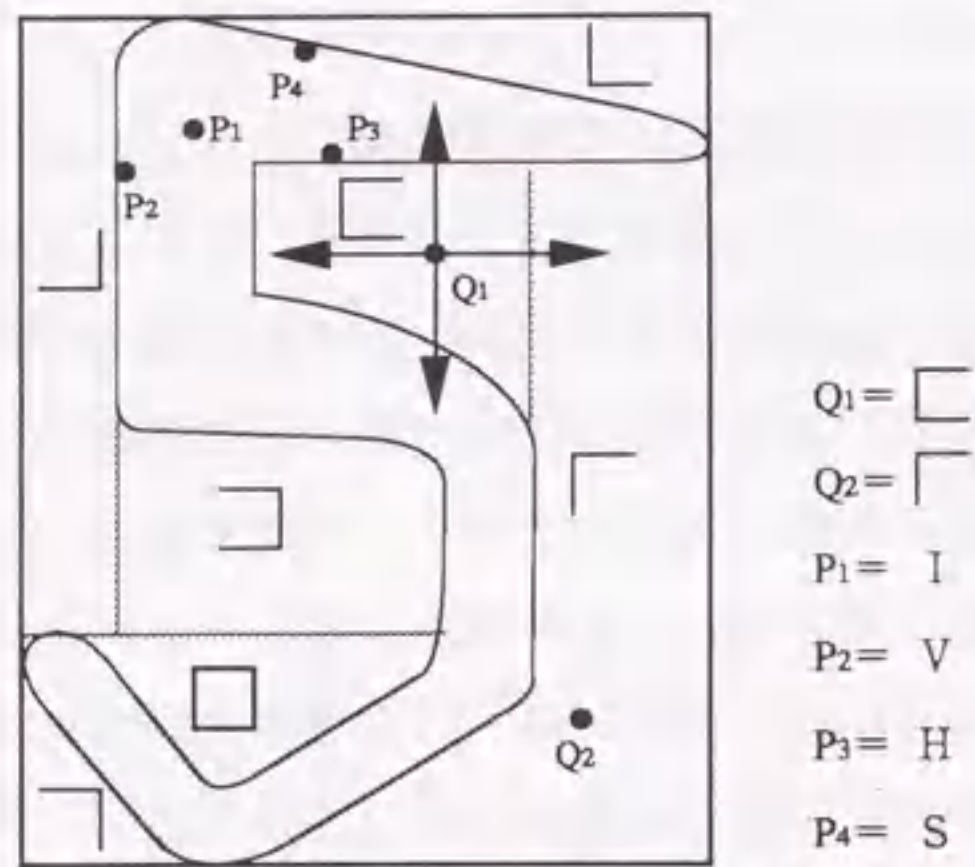


図8.2 1次位相特徴

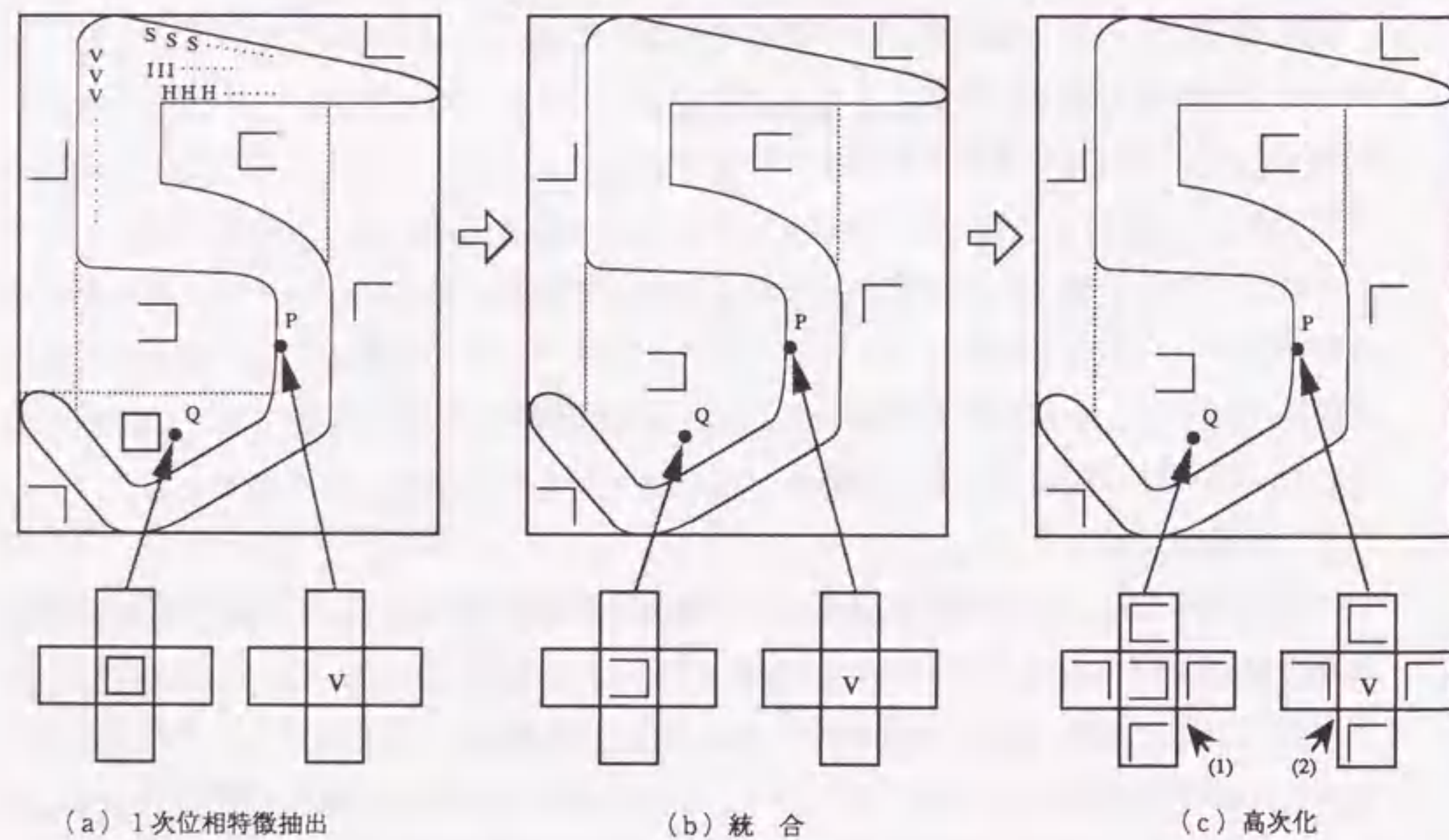


図8.3 位相構造化法による特徴抽出

各白点に通知することによって、黒点の特徴は3 x 3のマスキ演算を繰り返すことによって、それぞれ抽出する。

統合処理は、図8.3(b)に示すように、隣接する白点の特徴を統合する処理であり、LISCARでは上下左右方向の伝搬演算を繰り返して行う。

高次化処理は、各画素の上下左右方向の1次位相特徴を集積する処理である。図8.3(c)に示すように、黒点に関しては、その点から上下左右方向を眺め、最初に見出される各白点特徴をその黒点特徴に付加して高次化特徴とする。また、白点については、その点から上下左右方向を眺め、文字線を越えて最初に見出された白点特徴を、その白点特徴に付加して高次化特徴とする。この高次化処理についても、LISCARでは上下左右方向の伝搬演算を繰り返して行う。

8.3 印刷漢字読み取り

既存の印刷文書を専用のイメージスキャナを介して短時間で読み取って、文書データベースを効率良く作成することを目標に、日本語文書に現れる文字（以後印刷漢字と呼ぶ）に対し、認識精度99%以上、認識速度20~30字/秒以上を目指した。印刷漢字では、漢字のカテゴリ数が3,000~4,000にもなるため、認識率の面で望ましい特徴ベクトルの全数照合を辞書との間で行おうとすると処理量が膨大になってしまう。従って、この膨大な照合処理の高速化と認識精度の向上の両立を、LISCARの並列処理活用により、いかに実現するかがポイントになる。

8.3.1 認識処理の構成

一般に漢字認識のアルゴリズムは、演算量の大半を占める照合処理で、認識対象の文字毎に抽出する特徴ベクトルと辞書内の標準ベクトルとの間の識別計算（距離計算、類似度計算等）を繰り返して行うことから、並列処理向きと言える。それでも、カテゴリ数が大きく、複雑な識別関数を選ぶと目標の認識速度は達成できなくなる。さらに、全体の高速化には、前処理、特徴抽出にも並列処理がうまく適用できることが必要になる。以上の高速化の条件から、特徴としては、LISCARのビットライン単位の並列処理により比較的容易に抽出処理可能な傾斜別とペリフェラル⁶⁶⁾を、識別関数としては演算量の少ないシティブロック距離を選択した。この印刷漢字の認識処理系の構成を図8.4に示す。なお、前処理は、12本/mmのページ型の専用イメージスキャナからのページ入力に対応できるものと、16本/mmのハンドスキャナからの帯状入力に対応できるものを用意する。

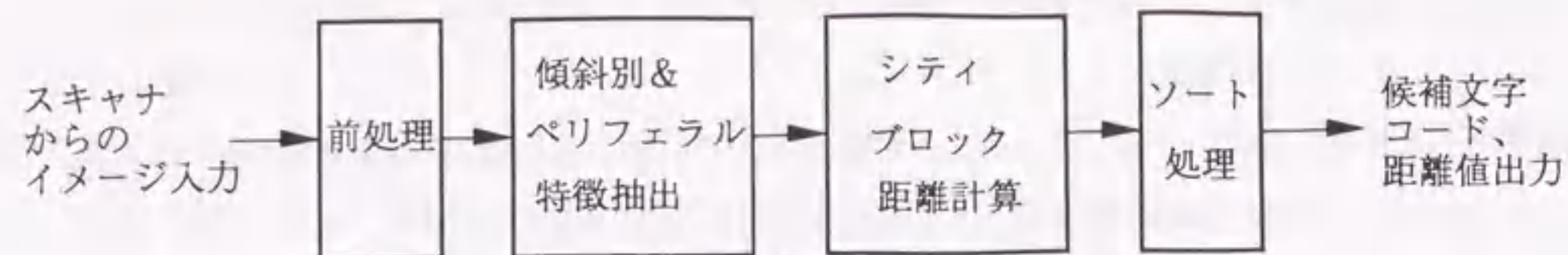


図8.4 印刷漢字認識処理系の構成

8.3.2 並列処理の概要

(1) 前処理

ページ入力の場合は、文書イメージデータの縦方向、および横方向の周辺分布を求め、周辺分布の周期性から文字列を抽出する。次に、文字列と垂直方向の周辺分布を求め、黒の部分と白の部分の位置関係をもとに文字を検出して切出す⁶⁷⁾。ここでは、長さが256のビットライン単位の並列処理によって高速化を図っている。横方向の周辺分布はLISCARの90度回転機能を用い、ページ全体を90度回転したイメージデータに対して行っている。

ハンスキャナの帯状入力の場合は、帯状のイメージ領域を複数の区間に分割し、その区間毎に、文字列方向に黒画素の周辺分布をとる。周辺分布で一定値以上の値を持つ領域を切出し領域の候補として抽出する。次に、その先頭の区間の周辺分布でセンサ視野中心部に最も近い部分の切出し領域候補を文字列の切出し領域の先頭と見なし、その

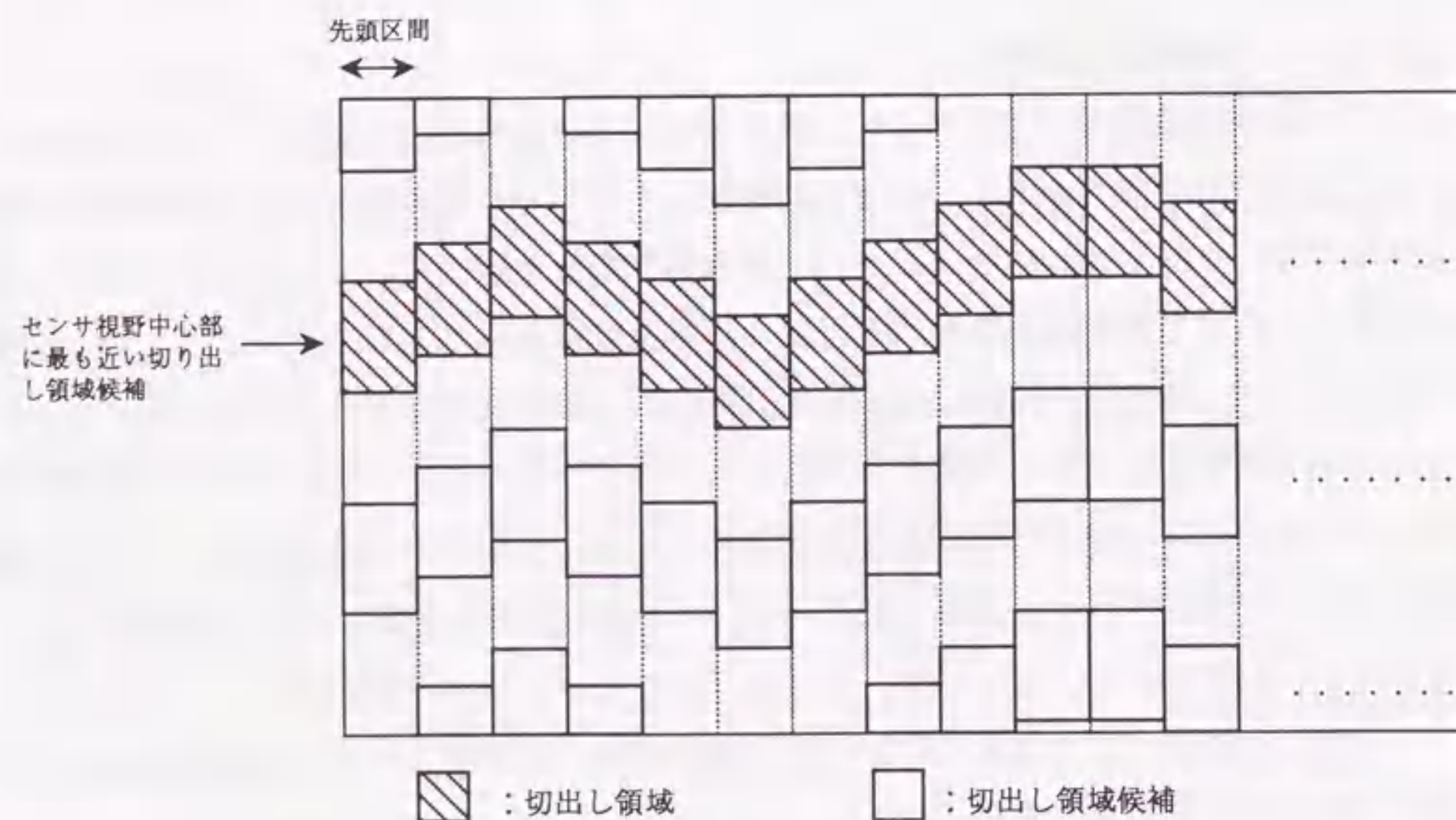


図8.5 ハンスキャナ入力の文字列切出し

切出し領域が区間毎に順に移動して行くのを、隣接する区間の間では切出し領域が必ず重なりを持つとして、切出し領域候補を追跡する。LISCARではこの追跡処理を、伝搬演算を利用して効率良く実行できる。

切出した文字は、3x3のマスク演算によりノイズ除去(平滑化)し、64x64のサイズに正規化する。演算方法は、手書き英数カナ文字の場合と同様に長さが256のビットライン単位の並列処理で実行する。

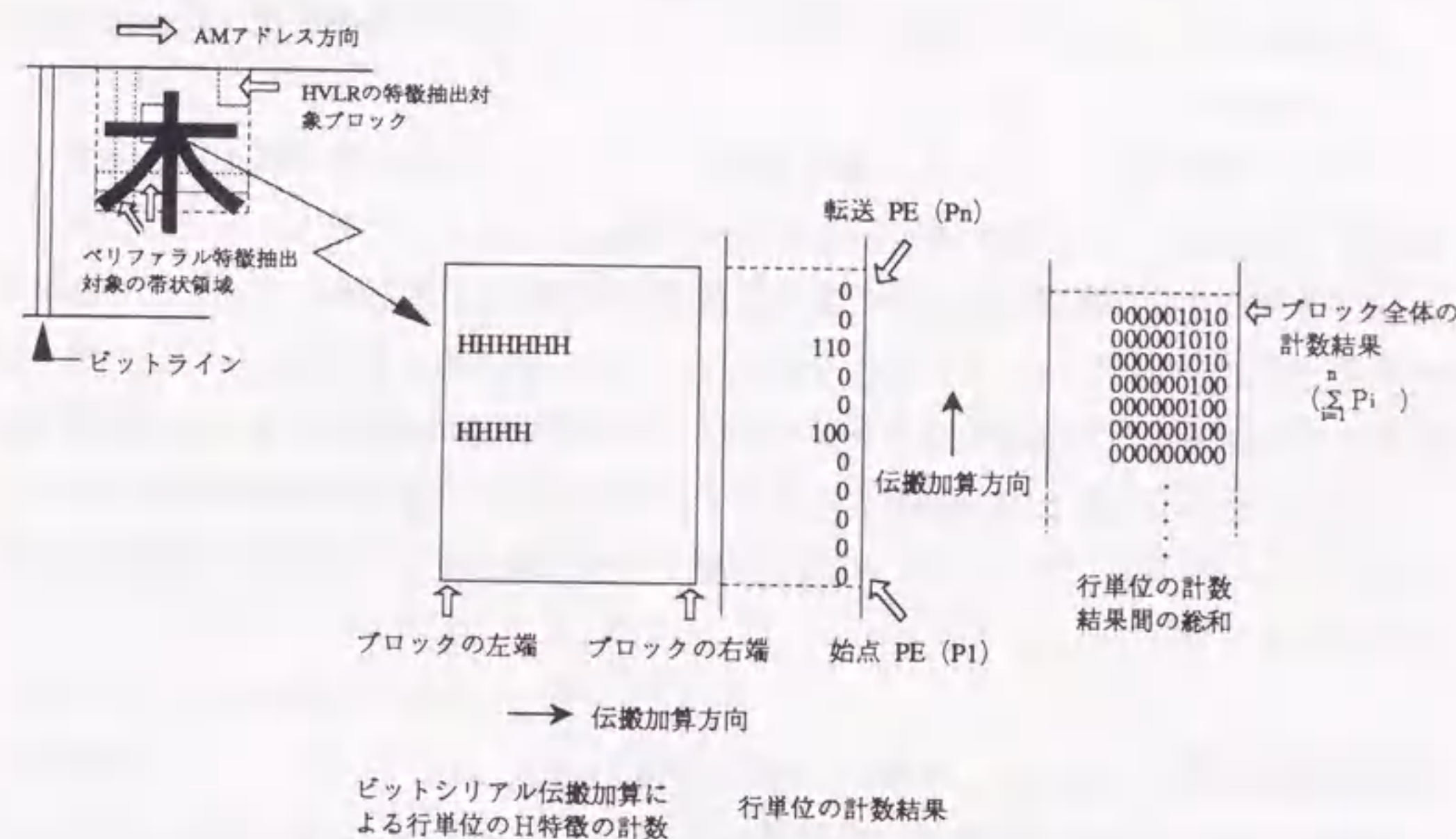


図8.6 特徴抽出とその計数処理のイメージ

(2) 特徴抽出

傾斜別特徴は、文字の輪郭部の黒点を連結方向別に4種類(水平:H、垂直:V、左上がり:L、右上がり:R)に分類し、64x64の文字パターンを4x4に等分割したそれぞれの方形ブロック毎で集計したものである。LISCARによるこの傾斜別特徴抽出の方法を、図8.6に文字「木」に対するH特徴抽出を例に示す。この特徴抽出の手順は、図3.6の総和演算法に基づいたもので、以下の通りである。

- ① AM内に格納している文字パターンに対し、ビットライン単位の演算により、輪郭を抽出する。
- ② 輪郭に対し、3x3のマスク演算によりH特徴の黒点を抽出しAM内に格納する(図中央左より参照)。
- ③ H特徴をブロックごとに行単位でビット直列伝搬加算により計数する。その結果は

図中央右の枠内に示すように得られる。

④その計数結果を、ビット並列データに変換した後、ブロック単位で並列に総和をとる。(図右側参照)。

この他のV, L, Rの特徴についても同様に求めている。

一方、ペリフェラル特徴は、縦横各々から外接枠内を8本の帯領域に等分割し、その各帯領域毎に外接枠と1本目の文字線間あるいは2本目の文字線間に存在する白点数を計数した値である。LISCARによるこの特徴の抽出は、

- ①各带状領域ごとの計数対象白点領域をビットライン単位の論理演算と伝搬演算により求める、
- ②その計数領域をビットライン単位で読みだし、PE配列内のグレイスケールラインに加え込む、
- ③その結果を伝搬加算により带状領域毎に集計する

の手順で行っている。

以上の特徴抽出の結果得られる特徴の数は、傾斜別が各傾斜毎に16で4方向分を合せて64、ペリフェラルが方向毎に1本目と2本目の分の16で4方向分を合せて64、さらに単純にブロックごとの黒点の数を計数するメッシュ特徴16が加わり、全体では144となる。

(3) 距離計算

距離計算はJIS第1水準の漢字を含む3,328カテゴリ分の標準特徴ベクトルからなる辞書に対して行う。具体的には、特徴抽出によって得られた特徴ベクトルを放送により各PEの局所メモリである配列データメモリ(AM)に格納し、あらかじめ各PEの局所メモリに分散配置しておく13カテゴリ分の標準の特徴ベクトルとの間で、各PEごとに並列に距離計算を行い距離値の2次元配列を得る。

(4) ソート処理

各カテゴリを距離の小さい順に並べ直すソート処理は次に示す近似的な方法を用いる。

- ①各PEの局所メモリに得られている距離値の最小値とその次に小さい距離値を選び出す。
- ②各PEの最小値の配列(グレイスケールライン)の全体の中から、最小値検索により小さい順に所定の数だけを選び出す。
- ③各PEの最小値の次に小さい距離値の配列の全体の中から、小さい順に所定の数だけ

け選び出し、その中に②で選び出した配列より小さい値があれば、それを②の配列に組入れ、それを最終的な候補距離値配列とする。ただし、②の配列への組入れがあった場合には、そのことがわかるようにフラグをたてる。

このソート処理は高速ではあるものの、処理手順から明らかなように、③で入れ替えが起る場合には、1位以外は、不正確である可能性が残る。しかし、類似した文字の標準の特徴ベクトルがPE間で分散配置されるように、辞書をAMに適切に割付けることによって、ソート処理誤りの低減が可能であり、これによる認識率低下は無視できる程度に小さくなるものと考えられる。

8.4 手書き漢字読み取り

英数カナに加え、漢字をも扱える帳票読み取り装置実現を目指し、手書き漢字認識機能のLISCARへの搭載を試みた。手書き漢字は、印刷漢字に比べると字形の変形の程度が激しく、実用的な認識精度を得るには高度の認識アルゴリズムの搭載が要求される。このため処理の規則性が低下し、SIMDタイプの並列処理による高速処理が困難になる。従って、この問題をいかに克服して実用的な認識速度と精度を確保するかが、漢字認識機能実現上のポイントとなる。

8.4.1 認識処理の構成

認識アルゴリズムとしては、比較的規則性の高い処理のみで、高い認識率の得られる外郭方向寄与度特徴⁽⁶⁸⁾を次元圧縮した後、シテブロック距離計算あるいはユークリッド距離計算で識別する方法を採用する。処理系の構成を図8.7に示す。

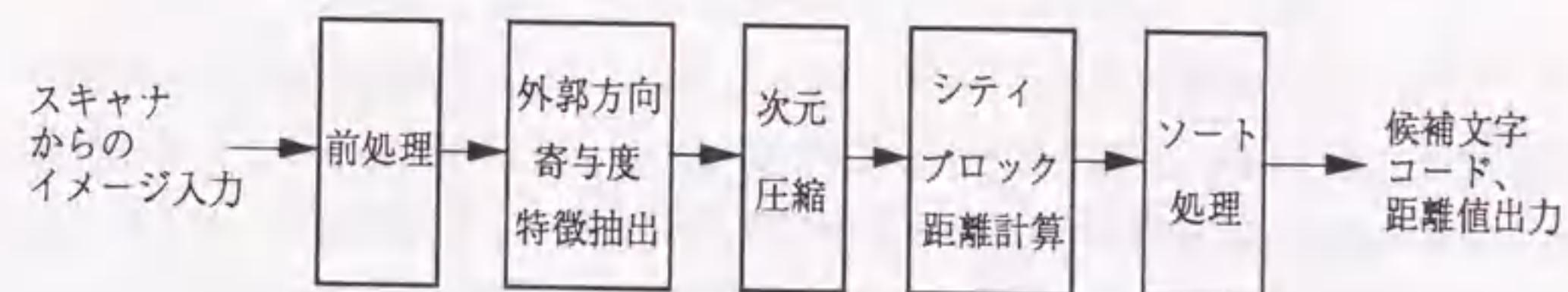


図8.7 処理系の構成

8.4.2 並列処理の概要

前処理とソート処理の内容は、基本的に前節の印刷日本語文書認識処理と同一である。従って、ここでは、特徴抽出、次元圧縮、距離計算について説明する。

$$\begin{aligned}
 L_1 &= 6 & t_1 &= 0.49 \\
 L_2 &= 3 & t_2 &= 0.24 \\
 L_3 &= 7\sqrt{2} & t_3 &= 0.81 \\
 L_4 &= 2\sqrt{2} & t_4 &= 0.23
 \end{aligned}$$

$$\sqrt{\sum_{m=1}^4 L_m^2} = 12.3$$

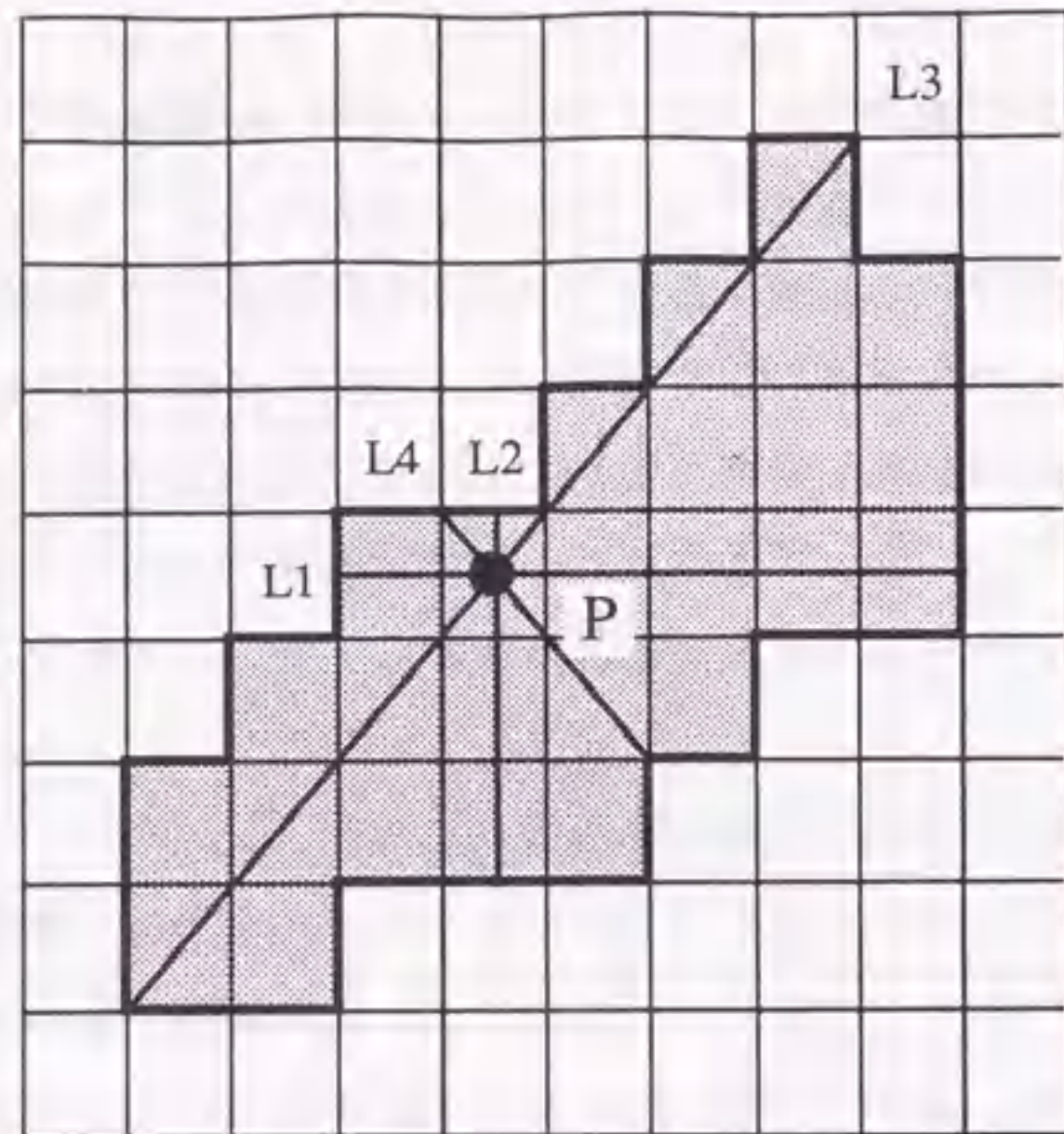


図 8. 8 文字パターンと 4 方向寄与度との関係

(1) 外郭方向寄与度特徴抽出

処理内容の説明に入る前に外郭方向寄与度特徴について説明する。文字線の外郭点の方向寄与度 $t = (t_1, t_2, t_3, t_4)$ は、次式で定義される。

$$t_m = L_m / \sqrt{\sum_{m=1}^4 L_m^2} \quad (8.1)$$

ここで、 L_1, L_2, L_3, L_4 は、図 8. 8 に示されるように対象とする外郭点の所属する 4 方向の黒点連結の長さである。ただし、斜め方向については実際の黒点連結の長さに $\sqrt{2}$ を乗じた値としている。外郭方向寄与度特徴とは、周囲の 8 方向から見て、白から黒に変わる最初の外郭点 (1 次外郭点)、2 番目の外郭点 (2 次外郭点)、3 番目の外郭点 (3 次外郭点) の 4 方向寄与度を、それぞれ 1 次から 3 次の投影軸に取り出し、この投影軸を 8 区間に分割し、その区間毎の方向寄与度の総和をとったものである。従って、この特徴抽出によって得られるベクトル (原特徴ベクトル) の要素の数は、8 投影方向、4 方向寄与度、3 次外郭、8 区間がかかって、768 となる。

この外郭方向寄与度特徴の LISCAR1 による抽出処理は次のように行う。はじめに次の手順で黒点連結の長さの投影軸グレイスケールラインを求める。

①文字パターンとそれを 90 度回転したパターンに対し、ビットライン単位で上下方

向と斜め方向に走査し、4 方向の黒点連結の長さのテーブルを AM 内に作成する。

②それぞれの黒点連結の長さのテーブルを 90 度回転し、回転無しの黒点連結の長さのテーブルの組と回転した黒点連結の長さのテーブルの組を AM 内に作成する。

③文字パターンとそれを 90 度回転したパターンに対し、上から下方向および斜め左上から右下方向にビットライン単位で走査し、1~3 次の外郭点の検出処理を実行する。この処理により外郭点が見つかった場合に、その点の 4 方向の黒点連結の長さを AM 内から読み出し、対応する次数の投影軸グレイスケールラインにコピーする。ここで、斜め方向走査は、 i 番目の文字パターンのビットラインの外郭点の検出処理で用いる $i-1$ 番目の文字パターンと投影軸のグレイスケールラインを左右方向に 1 ビットづつシフトすることで等価的に実現する。

次に外郭別に黒点連結の長さの投影軸グレイスケールライン間の算術演算を行うことで、方向寄与度のグレイスケールラインを算出する。続いて、このグレイスケールラインを 8 等分し、その区間毎で伝搬加算により方向寄与度を累積し、最終的な外郭方向寄与度特徴を求める。

(2) 次元圧縮

係数行列と特徴ベクトルとの乗算となる次元圧縮処理は、以下の①、②の処理を特徴ベクトルを構成する全ての要素に対して繰り返すことにより行う。

①特徴抽出によって得られた特徴ベクトルの要素を全 PE に放送する。

②放送した要素と、それに対応する係数行列のグレイスケールラインとの間で乗算を実行し、その結果を累算用のグレイスケールラインに加え込む。

ただし、係数行列は、あらかじめグレイスケールラインの本数が特徴ベクトルの要素数に等しくなる方向で格納しておく。

8. 5 速度性能の改善

高速化前の 143 ns 版 LISCAR1 に搭載した各認識処理の精度と速度性能を表 8. 1 に示す。認識精度については、高水準の認識アルゴリズムを搭載しただけあって、従来の専用ハードウェア構成の OCR と同等かそれ以上の性能が得られている。しかし、認識速度については、専用ハードウェア構成の 20~30 字/秒以上の性能に比べ見劣りするばかりか、当初目標も満たせていない。そこで、本節では処理方法の見直しによる速度性能の改善を試みる。

表 8. 1 各文字認識処理の性能

		認識精度		認識速度 (字/秒)
		正読率	エラー率	
手書き 英数カナ	数字	99.66	0.06	7.7
	英字	98.25	0.14	—
	カナ	96.23	0.46	—
印刷漢字*1		98.05	0.16	25.5
手書き漢字*2		98.0	—	4.6

*1 カテゴリ数：3,328 *2 カテゴリ数：4,096

8. 5. 1 高速化手法

実際に試みた3つの高速化法について述べる。

(1) 複数文字の同時並列処理による高速化

LISCARの構成プロセッサ数は256で、ビットラインの長さは256である。これに対し、認識対象の文字パターンのサイズは64 x 64以下であり、文字パターンをビットライン単位の順次走査により直接処理する前処理や特徴抽出処理では、PE配列の3/4が無駄になってしまう。この無駄をなくすには、同時に複数の文字を並列に処理するのが有効である。並列処理可能な文字数としては、文字パターンを無駄なく割り付けるとすると、最大4まで可能である。ただし、手書き認識処理では外郭方向寄与度特徴抽出においては、もともと文字パターンのサイズのエリアの2倍の処理幅の128を用いて処理しているため、2文字並列に制限される。

(2) 語長低減による距離計算の高速化

LISCAR Iでは、シティブロック距離計算を各PEのビット直列演算で行う。このため、演算時間はほぼ特徴ベクトルの要素の語長に比例することになる。従って、要素の語長を低減できれば、その分演算時間を短縮できる。

実際、傾斜別特徴では、16 x 16の方形ブロックを横切る線の数がせいぜい2本程度であり、現れる特徴値は最大64程度であると考えられることから、8ビットの語長を確保するのはもともと精度過剰といえる。逆に、与えられた語長分を完全に利用するようにすれば、認識率の低下なく、要素の語長が短縮できるものと考えられる。このような考えに基づき、次式の定比例的なスケール変換による4ビットまでの語長短縮を試み、印刷漢字認識処理の読み取り精度をシミュレーションにより評価した。

$$Y_i = a + \{[X_i - \text{MIN}(X_i)] / [\text{MAX}(X_i) - \text{MIN}(X_i)]\} \times (b-a) \quad (8. 2)$$

ここで、 Y_i は変換後の*i*番目の特徴値、 X_i は*i*番目の変換前の特徴値である。MIN(X_i)、MAX(X_i)は特徴値 X_i の最小値、最大値をそのまま用いるのではなく、それぞれ1.02、0.98を乗じた値を用いる。

評価結果を、単純に4ビット足切りした場合と比較して、表8. 2に示す。この結果より、定比例的なスケール変換によれば語長を4ビットまで短縮しても、認識精度がほとんど低下しないことがわかる。

表 8. 2 特徴値の語長短縮の認識率への影響

	語長	語長低減手法	正読率
506 5号 未学習データ	8	—	99.76
	4	定比例法	99.76
		足切り	99.63
701 7ポイント 未学習データ	8	—	99.30
	4	定比例法	99.27
		足切り	99.02

(3) 逐次処理の低減

以上の(1)、(2)では、主に並列処理部の高速化法である。しかし、もともとDSP部やアドレス生成部での逐次的な処理の割合の大きい手書き英数かな認識処理においては、並列処理部が高速化すると、今度は逐次処理の部分が目立ってくる。そこで、手書き英数カナ文字の認識アルゴリズムを見直し、逐次処理部分の処理量を低減した。具体的には、位相特徴のみで識別可能かどうかを判断し、可能な場合には輪郭特徴抽出処理をバイパスするようにして高速化をはかった。

8. 5. 2 高速化の結果と考察

それぞれの認識処理の高速化の結果を処理時間の内訳と合せて表8. 3～8. 5に示す。8. 5. 1節の高速化法の適用と一部処理方法の見直しだけで、当初の1.8倍～2.1倍の高速化に成功している。手書き数字認識と手書き漢字認識の結果は、並行して行った高速化のためのマイナチェンジの結果である80ns版のLISCAR Iで測定したものである。これらについては、143ns版からハードウェアが実行的に

1.6倍程度高速化されたことも加わり、表8.1に示した認識速度に比べ、全体で3倍程度高速化され、専用ハードウェアに匹敵する性能が達成されている。印刷漢字認識については、143nS版での性能を示しているが、80nS版では同様に1.6倍程度高速化され、70文字程度の認識速度が見込まれる。

表8.3 手書き数字認識の高速化前後の処理時間の内訳と認識速度

処理内容		処理時間 (ms)		高速化法
		高速化前	高速化後	
前処理	文字パターン入力	1.2	同左	4文字並列
	ノイズ除去	2.9	0.74	
位相特徴抽出	位相特徴抽出	32.2	7.6	4文字並列
	シリパラ変換	5.0	5.0	
輪郭特徴抽出		12.8	1.3	逐次処理バイパス
位相特徴係数		16.0	同左	
文字識別		10.8	同左	
その他		1.7	同左	
全処理時間		82.6	44.3	
認識速度 (文字/秒)		12.1	22.6	

表8.4 印刷漢字認識の高速化前後の処理時間の内訳と認識速度

処理内容		処理時間 (ms)		高速化法
		高速化前	高速化後	
前処理		8.9	6.7	4文字並列
位相特徴抽出		4.4	1.1	4文字並列
識別	距離計算 (シティブロック)	24.4	12.5	定比例法による語長短縮
	ソート処理	1.5	1.5	
全処理時間		39.2	21.8	
認識速度 (文字/秒)		25.5	45.9	

表8.5 手書き漢字認識の高速化前後の処理時間の内訳と認識速度

処理内容	処理時間 (ms)		高速化法	
	高速化前	高速化後		
前処理	8.4	4.3	2文字並列	
特徴抽出	黒点連結の長さ	87.2	29.6	2文字並列、処理法見直し
	方向寄与度	8.5	8.5	
	区間内累積	1.3	1.3	
次元圧縮処理	11.3	11.3		
識別	距離計算 (シティブロック)	18.8	18.8	
	ソート処理	1.5	1.5	
全処理時間	137.0	75.3		
認識速度 (文字/秒)	7.3	13.3		

しかし、高速化の内容を個別に見てみると、並列処理が必ずしも有効に機能していない部分があり、これが高速化のネックになっていることがわかる。その典型は、手書き英数カナ文字の処理時間内訳に見られ、並列化困難な、位相特徴係数と識別の処理時間が全体の60%も占めている。また、一見高速化がうまくいっていると思われる印刷漢字、手書き漢字認識処理においても、SIMD型の並列処理がうまく機能しないことから、十分な高速化が達成できていないところがある。例えば、手書き認識処理の特徴抽出では2文字並列化をはかった上に処理の最適化をはかっても、その処理時間は全体の52%も占めている。これは、外郭方向寄与度特徴の抽出処理では、大半の処理を文字線の輪郭部分に対してのみ行えばよいのに、LISCAR1の単純なSIMD処理ではこのような選択的な処理が困難で、文字パターン全体をくまなく処理しているためである。また、全体に占める処理時間自体は大きくはないが、文字パターンの正規化処理では文字毎の拡大縮小率が異なるため、文字単位の並列処理が利用できず、LISCAR1の256プロセッサの3/4が無駄になってしまう。印刷漢字認識、手書き漢字認識処理では、実はさらにもう一つ大きな問題がある。それは、認識精度の点でシティブロック距離より優れているといわれるユークリッド距離がLISCAR1では、シティブロック距離計算の4倍程度の処理時間を要するために、利用できないことである。速度が大きく低下する原因は、LISCAR1ではユークリッド距離計算に表れる乗算もビット直列に処理せねばならず、一回の乗算に語長の2乗のオーダーのサイクル数を要するためである。

8.6 むすび

本章ではL I S C A R 1の文字認識処理への応用について述べた。

はじめに、手書き英数カナ文字、印刷漢字、手書き漢字の3種の代表的な文字認識処理に対し、前処理、特徴抽出、識別のそれぞれをどのように実行するかを示した。具体的には、手書き英数カナ文字認識処理については、並列処理に不向きな特徴計数処理以降はL I S C A Rの逐次処理機能を利用して実行することとし、残りの前処理と位相特徴抽出をビットライン間の論理演算と上下左右方向の伝搬演算によって実行する方法を示した。印刷漢字認識処理については、文字切出しを含む前処理から、傾斜別&ペリフェラル特徴抽出、シティブロック距離計算、ソート処理までの主な文字認識処理の全てをビットライン単位あるいはグレイスケールライン単位の算術論理演算あるいは伝搬演算によって実行する方法を示した。また、手書き漢字認識処理については、印刷漢字認識処理とは異なる処理である外郭方向寄与度特徴抽出をL I S C A R 1の単純なS I M D処理機能によって実行する方法と、抽出した特徴ベクトルの次元圧縮処理のための行列とベクトルの乗算の方法を示した。

次いで、これらの処理方法による実機での認識精度と認識速度の測定から、認識精度は96~99%と汎用コンピュータ上でアルゴリズムを評価した通りの値が得られることを示した。しかし、認識速度は5~20文字/秒程度と、当初目標を達成できていないことから、速度性能の改善を試みた。具体的には、高速化法として複数文字の同時並列処理、距離計算における被演算データの語長低減、一部逐次処理の条件に応じたスキップの3つの方法を各認識処理に可能な限り組込むことを試みるとともに、L I S C A R 1をクロックサイクル143nsの版から80nsの高速版に置き換えた。その結果、当初に比べ3倍程度の高速化が達成され、専用ハードウェアに匹敵する13文字/秒~46文字/秒の認識速度が得られることを確認した。

最後に、高速化後の処理時間の内訳を分析し、

- ①手書き英数カナ認識処理では、並列処理部分の高速化に伴い、逐次処理部分が全処理時間の60%を占め高速化の障害になっていること、
- ②逐次処理の占める割合を数パーセント以下にまで低減可能な手書き漢字認識処理においても、外郭方向寄与度の特徴抽出処理が、L I S C A R 1のS I M D処理の制約により高速処理困難なため、全体の52パーセントの処理時間を占めていること、
- ③認識精度向上の観点から、要求の強いユークリッド距離計算が、L I S C A R 1の乗算性能の低さから利用できないこと

等を明らかにした。

第9章 L I S C A R 2の漢字認識処理への応用

9.1 はじめに

前章では、L I S C A R 1により最高水準の認識精度の文字認識処理アルゴリズムが専用ハードウェア並みの速度で実行できることを示した。しかし、最高水準の認識精度と言え、人間に比べるとまだまだ低く、印刷漢字や手書き漢字では、文字品質によっては確認修正の手間が大きくなりすぎ使いものにならない場合さえある。このため認識精度向上の要求は強く、より高い正読率の得られる高度の認識アルゴリズムの検討が続けられている。

残念ながらこれらの高精度の認識アルゴリズムは、L I S C A R 1の得意とする傾斜別特徴やシティブロック距離をベースとした単純な処理のみでは構成できない。例えば、最近行われた文字認識アルゴリズムのコンテストで上位を占めたアルゴリズムは、特徴として輪郭特徴を用いたり、識別関数としては積和演算を多用する高度なものが利用されている⁶⁹⁾。このような認識処理をL I S C A R 1のシンプルなS I M D処理でカバーするのはかなりの無理があり、十分な速度性能が得られる見通しはたない。

このような状況のもとに新たに改良試作したのがL I S C A R 2である。本章では、このL I S C A R 2上に、8方向の外郭方向寄与度特徴に投影距離⁷⁰⁾を組み合わせた最新の手書き漢字認識アルゴリズムを組込む際に、高速化に工夫を凝らした特徴抽出処理、次元圧縮処理、大分類処理、識別処理の処理方法を明らかにし、速度性能を評価する。

9.2 認識処理の構成

図9.1にL I S C A R 2に搭載する手書き漢字認識処理の手順を示す。図8.7のL I S C A R 1の手書き漢字認識処理との違いは、外郭方向寄与度特徴の方向寄与度を4方向から8方向に拡張している点、識別処理を大分類と識別に分け大分類でユークリッド距離による全数照合を、識別で大分類で上がった上位候補に対する投影距離による照合を行う点である。

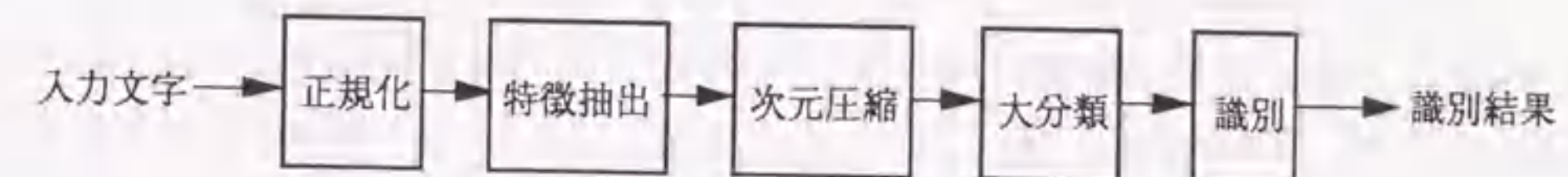


図9.1 手書き漢字認識処理の構成

9.3 並列処理の概要

9.3.1 特徴抽出処理

(1) 高速化の方針

LISCAR1では、外郭方向寄与度特徴の抽出時間が長く、これが手書き漢字認識全体の速度を低下させる主因になっている。2文字並列や処理の最適化を行っても特徴抽出処理の高速化が十分達成できないのは、外郭部の黒点連結の長さを投影軸のグレイスケールラインに投影する処理が、LISCAR1の単純なSIMD制御では効率よく実行できないからである。具体的には、次の3つの処理を文字パターンを構成する各ビットラインごとに毎回繰り返し実行しなければならないからである。

① 1~3次の外郭部の検出、

② 検出された外郭部黒点連結の長さの1~3次、4方向成分(計3x4本)の投影軸グレイスケールライン(要素のデータ語長6ビット)へのコピー、

③ 斜め方向走査の場合の全投影軸グレイスケールラインの1ビット左右方向シフト。

これらの3つの処理の中で②と③の処理が特に重い。ビット直列処理ベースのLISCAR1では、総計72ビットのコピーあるいはシフトを、ビット処理に分解して実行しなければならないからである。

これに対してLISCAR2ではコピーあるいはシフトの単位が8ビットあるいは16ビットと大きく②と③の処理の負担はかなり軽減される。それでも、LISCAR2では配列サイズが1/4の64になっている上、前節で説明したように方向寄与度の方向を4から8に拡大すると、コピーあるいはシフトの対象の総計ビット数が拡大前の2倍の144ビットにもなるため、LISCAR1と同一の処理アルゴリズムを用いるのでは大幅な処理速度の向上は望めない。そこで、LISCAR2では新たに組込んだ転送回転型メモリアクセス機構による間接アドレッシング機能を駆使することによって、文字パターンを構成するビットライン毎の②と③の処理を不要として、特徴抽出処理の高速化をはかる。

(2) 8方向寄与度への拡張

文字線の外郭点の方向寄与度 $t = (t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8)$ を、次式で示すように8方向に拡張する。

$$t_m = L_m / \sqrt{\sum_{m=1}^8 L_m^2} \quad (9.1)$$

ここで、 L_1, L_2, \dots, L_8 は、図9.2に示すように対象とする外郭点から8方向に向かう黒点連結の長さである。斜め方向については実際の黒点連結長に $\sqrt{2}$ を乗じた値とす

$L_1 = 5$	$t_1 = 0.47$
$L_2 = 2$	$t_2 = 0.19$
$L_3 = 4\sqrt{2}$	$t_3 = 0.53$
$L_4 = 4\sqrt{2}$	$t_4 = 0.53$
$L_5 = 1$	$t_5 = 0.09$
$L_6 = 3$	$t_6 = 0.28$
$L_7 = 1\sqrt{2}$	$t_7 = 0.13$
$L_8 = 2\sqrt{2}$	$t_8 = 0.27$

$$\sqrt{\sum_{m=1}^8 L_m^2} = 10.6$$

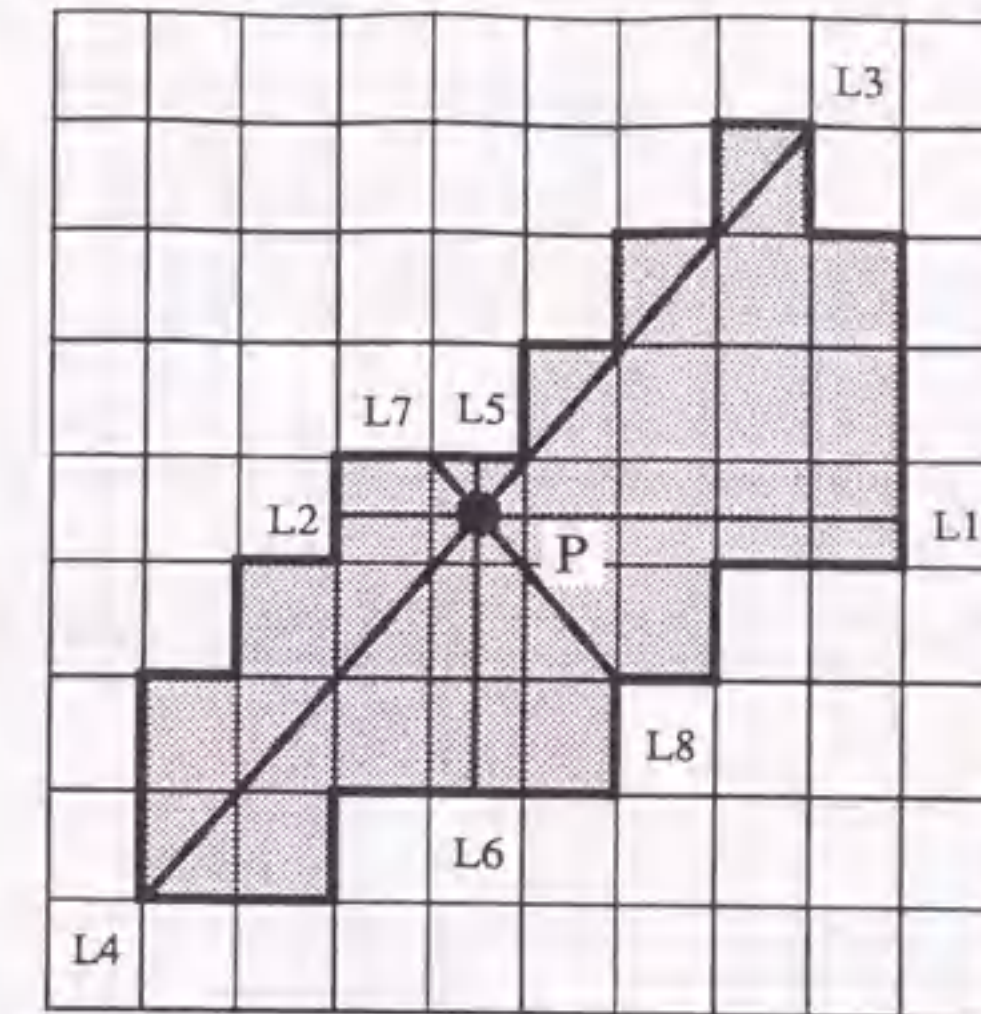


図9.2 文字パターンと8方向寄与度との関係

るのは、4方向寄与度の場合と同じである。方向寄与度が4方向から8方向に拡張される点を除くと、外郭方向寄与度特徴抽出の考え方は、LISCAR1の場合と全く同一である。従って、この特徴抽出によって得られるベクトル(原特徴ベクトル)の要素の数は、8投影方向、8方向寄与度、3次外郭、8区間がかかって、1,536となる。

(3) 特徴抽出処理の内容

高速化のポイントは、AM内に形成する黒点連結の長さのテーブルを、間接アドレッシング機能を利用して、文字パターン走査より得られる外郭アドレス値で読み出すことにより、投影軸グレイスケールラインを繰り返し処理なしで効率良く求めることにある。以下、この間接アドレッシングによる投影軸グレイスケールラインの算出を実現するための、黒点連結の長さのテーブルの作成法と1次から3次の外郭点アドレスグレイスケールラインの算出法を中心に説明する。

図9.3は、あらかじめ演算器配列の上側に示すレジスタファイルに格納しておく文字パターンを、上側から下に向かってビットライン単位に走査し、黒点連結の長さを計算し、その結果を配列データメモリ(AM)内に格納して順次黒点連結の長さのテーブルを作成して行く際のレジスタファイルとAM内のデータの割付けを示している。この図から明らかなように、 $L_1 \sim L_8$ の黒点連結の長さは連続して格納している。これは、間接アドレッシングにより黒点連結の長さ一式をまとめてアクセスできるようにするためである。

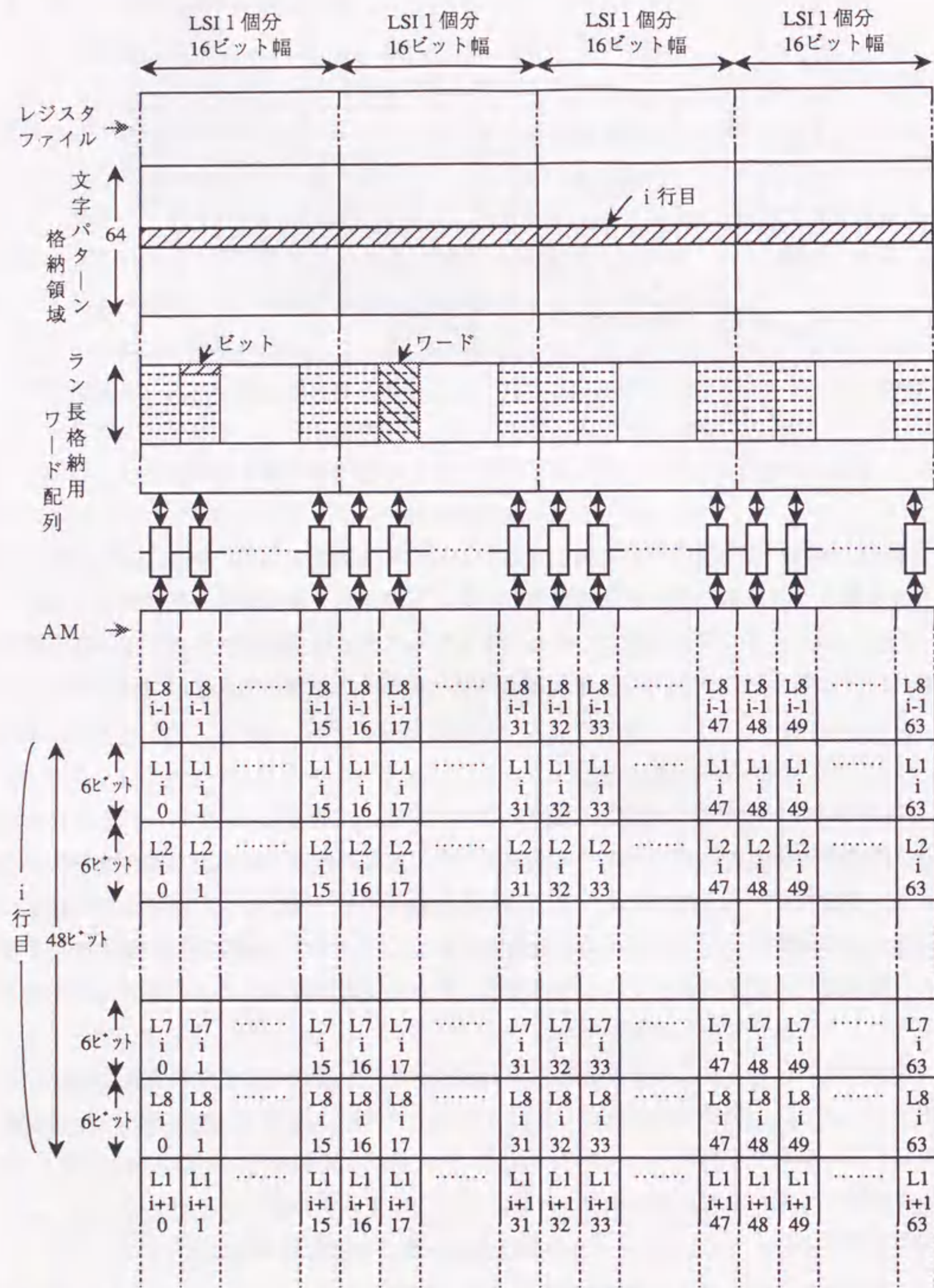


図9.3 文字パターン走査による黒点連結の長さのテーブル作成時のメモリ割付け

図9.4は、LISCAR2の間接アドレッシングによるデータアクセス機能を利用して外部部の黒点連結の長さをアクセスできるように、黒点連結の長さのテーブルの格納形式をビット直列型からビットパラレル型に変換した様子を示している。この直並列変換は、LISCAR2の転送回転型メモリアccess機構をLSI単位で動作させることで実現する。

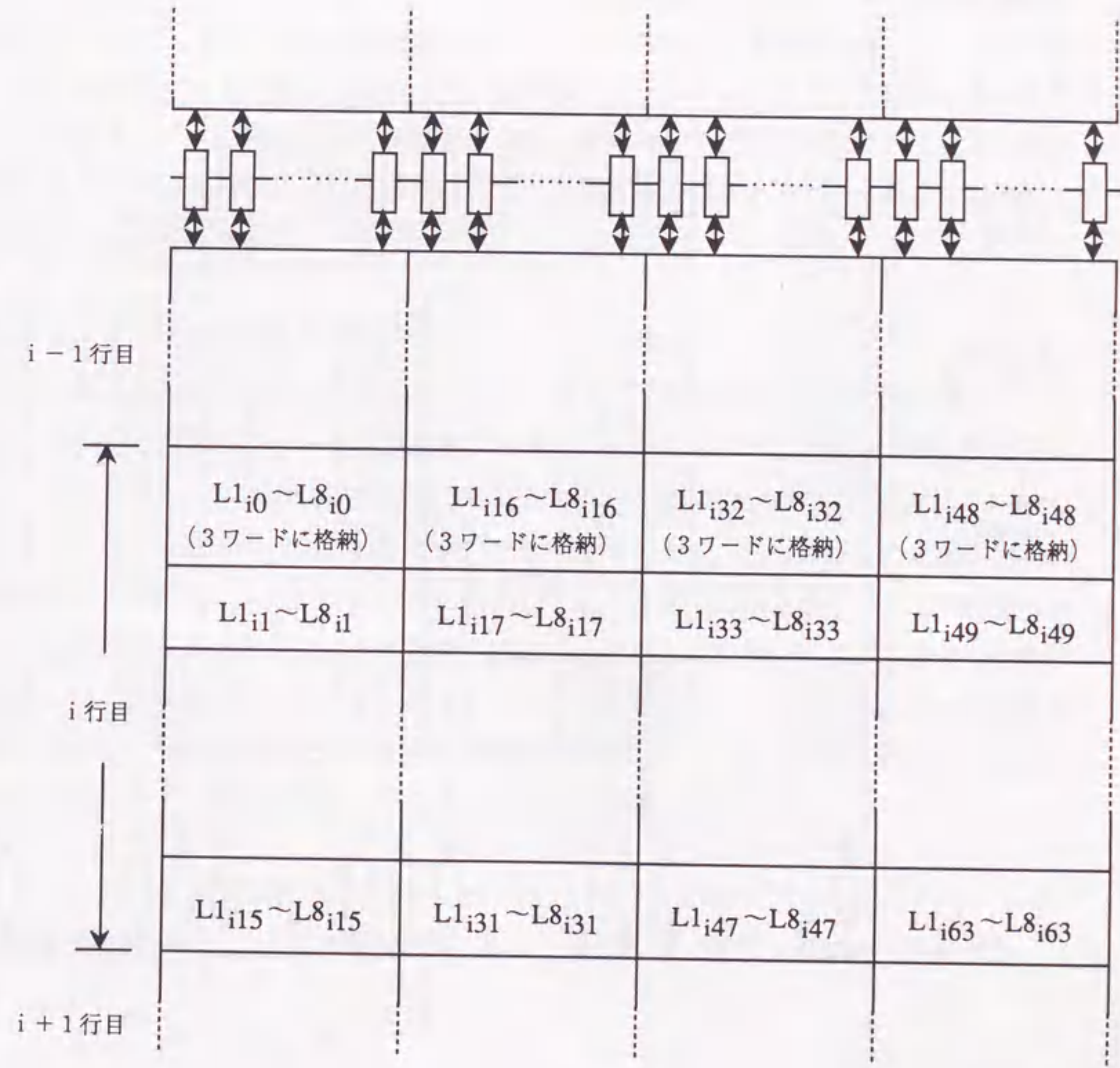


図9.4 直並列変換後の黒点連結の長さのテーブル

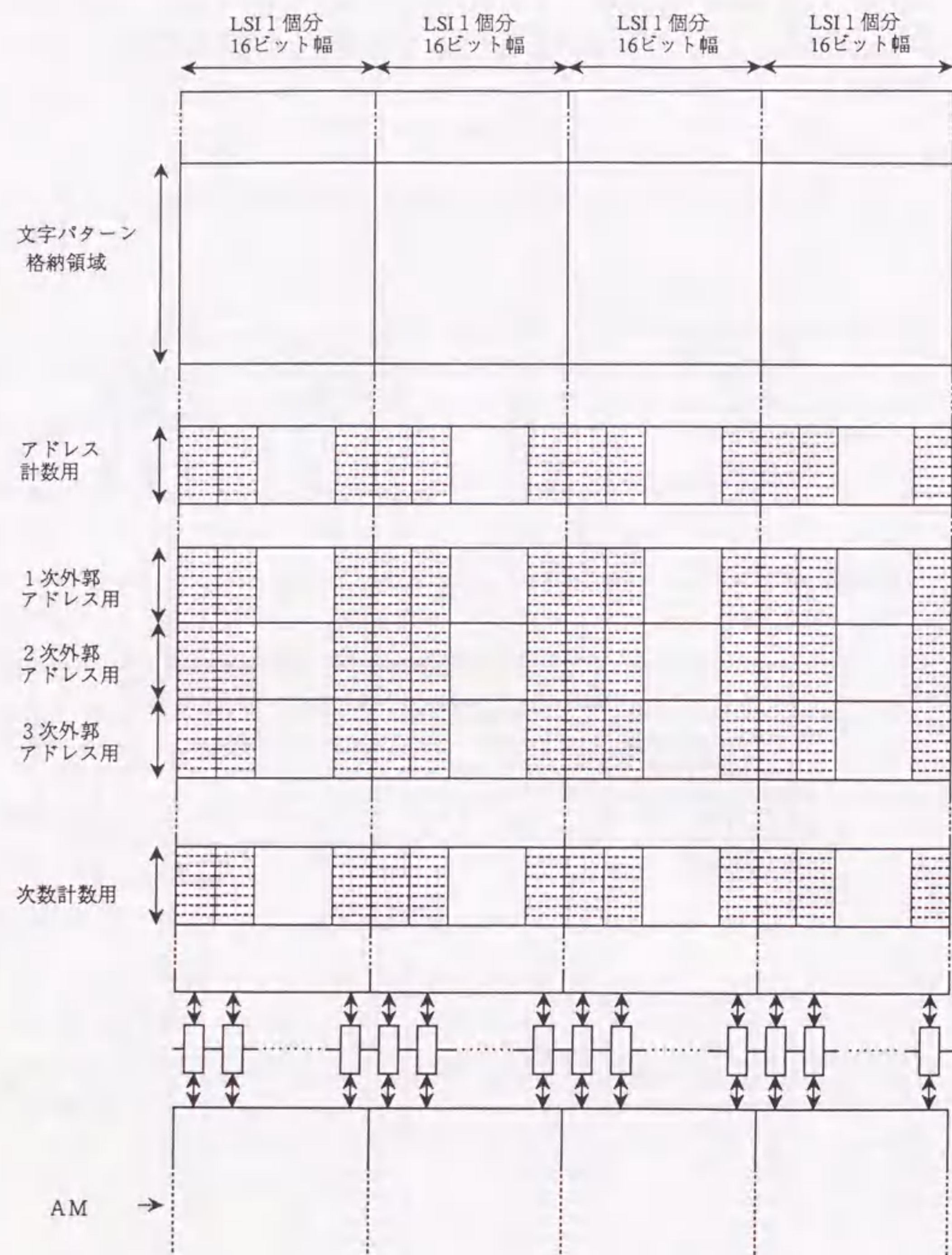


図9.5 外部アドレス抽出処理におけるメモリ割付け

図9.5は、1次から3次の外郭アドレスを求める際のメモリへのデータ割付けを示している。この処理は、文字パターンをビットライン単位で走査し、外郭部を検出したならば、並行して求めておく走査中のビットラインのアドレス値を、外郭アドレスのグレイスケールラインに、やはり並行して求めておく外郭次数値を参照して選択的にコピーすることによって実現する。これによって得られたアドレス値は、黒点連結の長さのテーブルのアドレス値そのものではない。黒点連結の長さが3ワード単位でまとまっていることから、3倍するとともにベースアドレスとして黒点連結の長さのテーブルの先頭アドレス値を加える必要がある。

この最終的なアドレス値のグレイスケールラインを用いて図9.3に示される黒点連結の長さのテーブルを、転送回転型メモリアクセス機構による間接アドレッシングを利用して、直接アクセスすることによって、投影軸グレイスケールラインを一気に求めることができる。これ以降の方向寄与度算出処理、区間内の累積処理はLISCAR1と同様に実行する。もちろん、LISCAR2ではこれらの処理を8ビット単位で処理する分だけ、高速化される。

9.3.2 次元圧縮処理

1536次元の原特徴ベクトルと1,536 x 256の係数行列との積をとり、256次元の特徴ベクトルを求める処理である。LISCAR1の場合と異なりシストリックアルゴリズム⁷⁾を用いる。図9.6は、その処理手順の概要を、3次元の特徴ベクトルと3 x 3の係数行列との積を例に示したものである。あらかじめ、係数行列を紙面の上下方向にねじった形でAM内に格納しておき、要素の語長が8ビットの特徴ベクトルを転送回転型メモリアクセス機構の転送パスを介し左方向にシフトしながら係数行列との間で乗算を行い、結果をレジスタファイルで累算することにより、次元圧縮処理を行う。LISCAR2では、AMからの係数行列のアクセスと、乗算あるいは特徴ベクトルのシフトを並列に行えるので極めて高い処理効率を得られる。

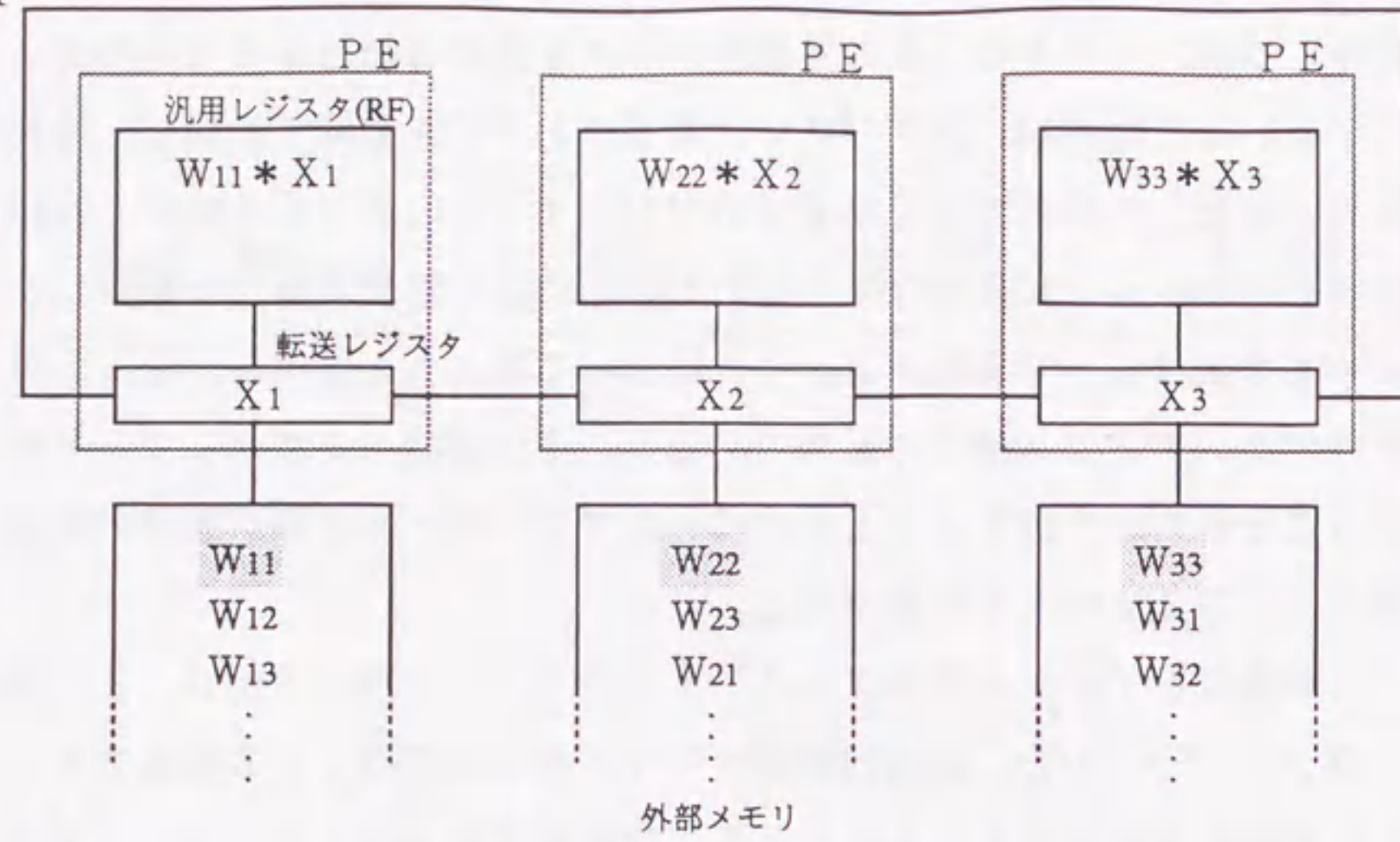
9.3.3 大分類処理

辞書側の標準ベクトル \bar{x} と特徴ベクトル x との間のユークリッド距離の2乗Dを

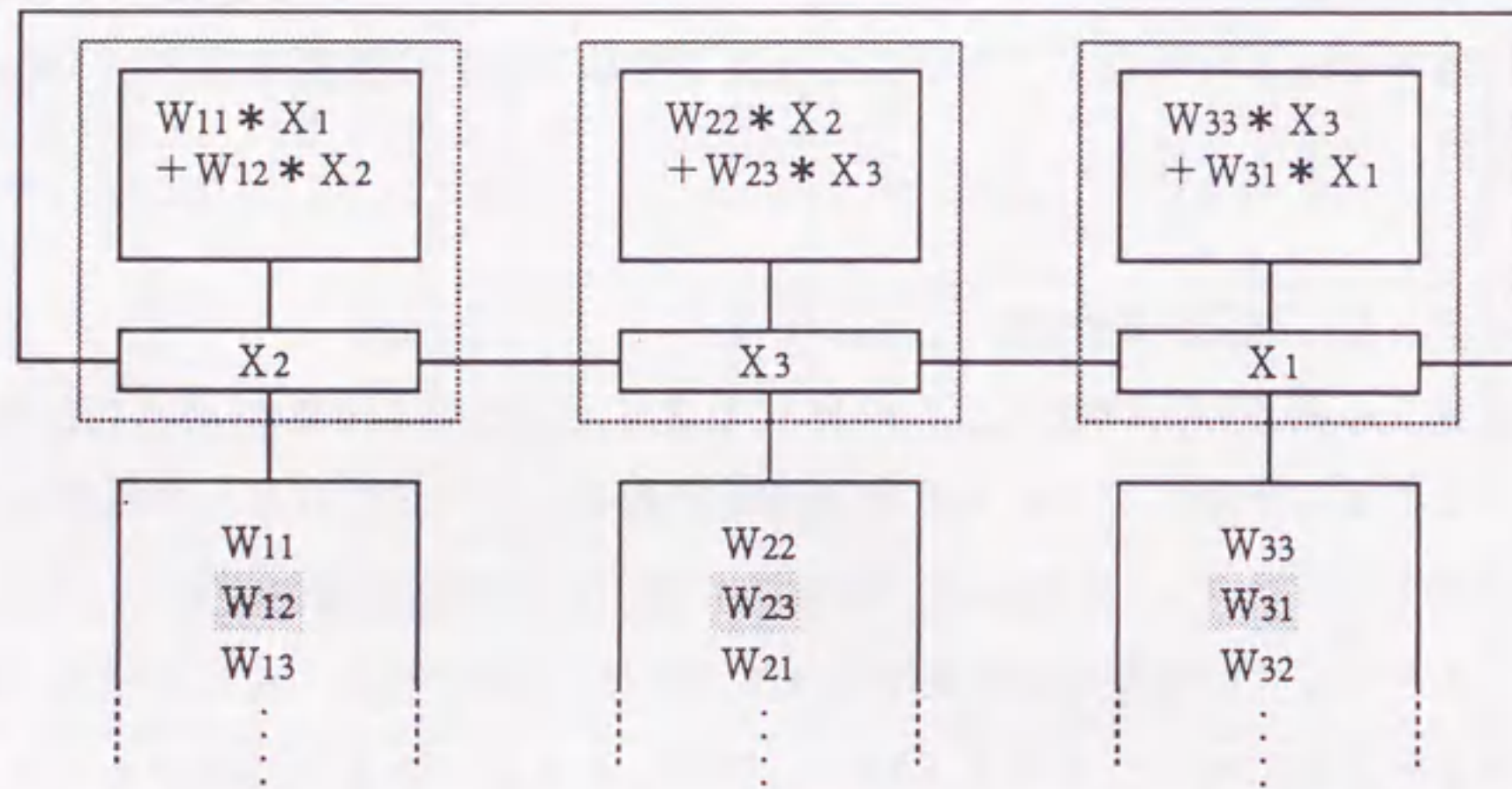
$$D = \sum_{i=1}^{256} (x_i - \bar{x}_i)^2 \quad (9.1)$$

を求める処理であり、次元圧縮処理の場合同様に実行する。すなわち、辞書側の標準ベクトルをあらかじめ左右方向にねじった形で格納しておき、次元圧縮処理で得られた

ステップ 1



ステップ 2



ステップ 3

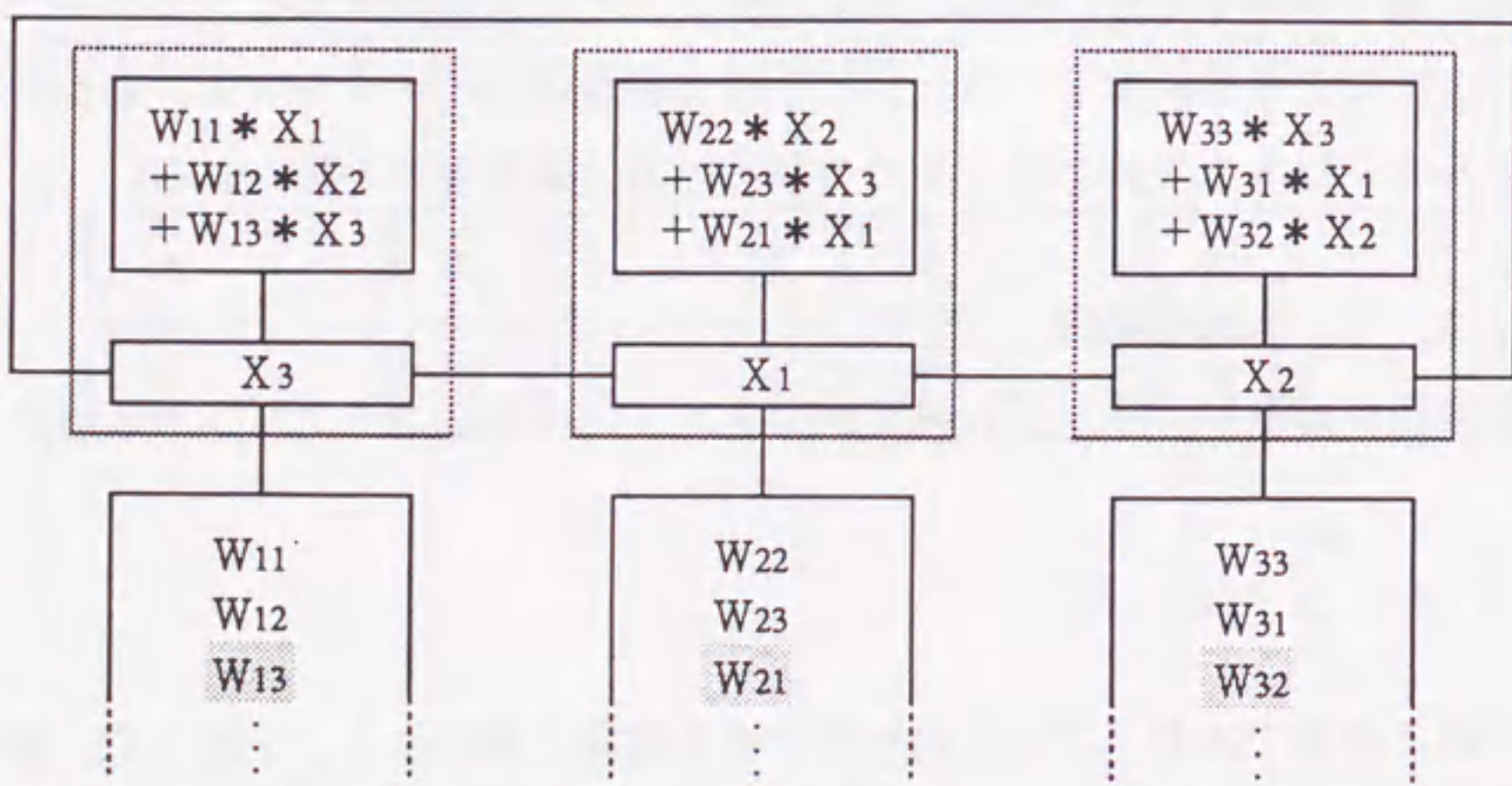


図 9. 6 シストリックアルゴリズム処理における次元圧縮処理

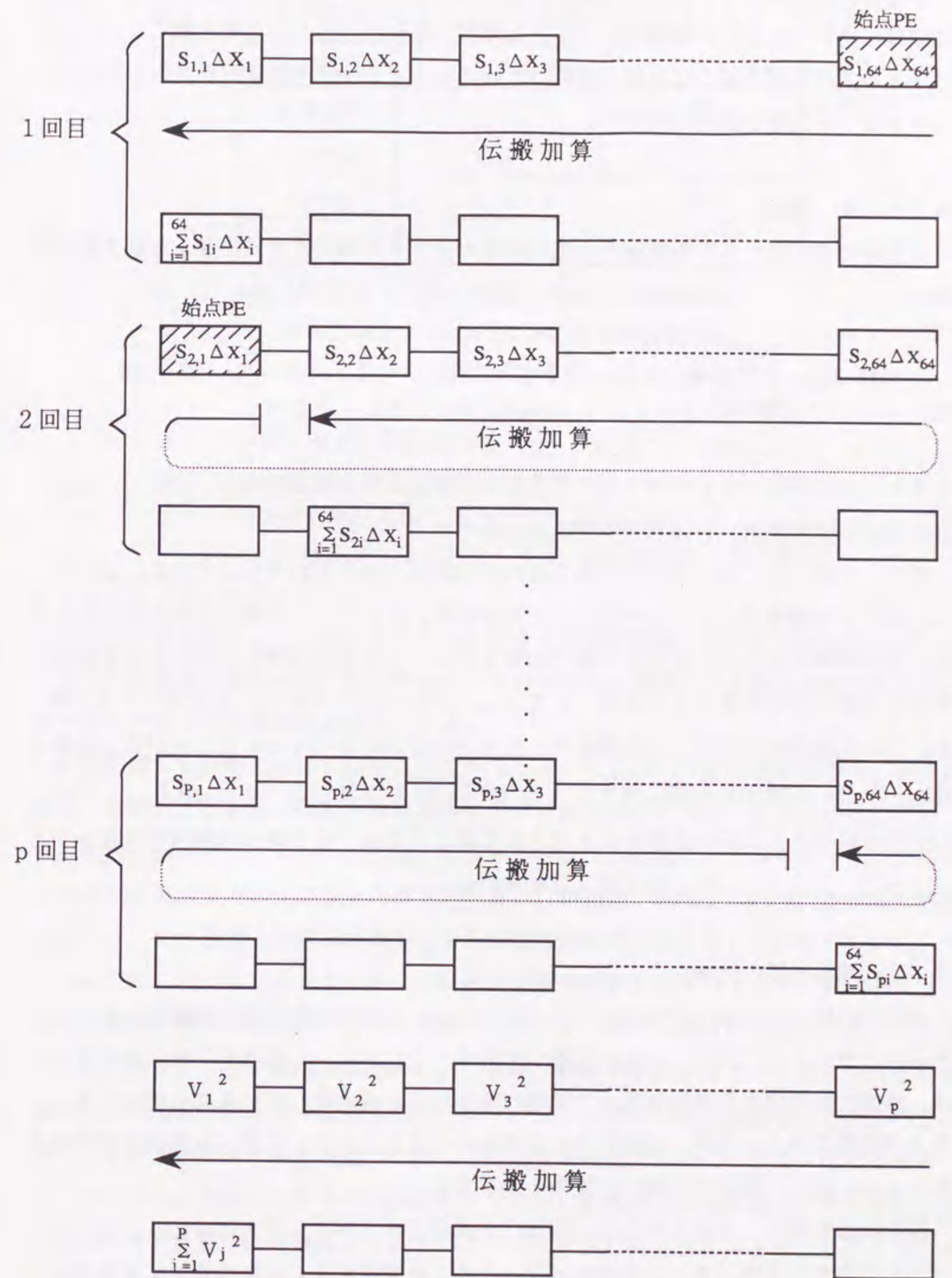


図 9. 7 投影距離の演算法

256次元の特徴ベクトルをPE配列の右から左にシフトしながら、AMの各PE対応の領域に格納されている標準ベクトルとの間で、差分をとったあとで2乗し、レジスタファイル内に累算することを繰り返す。これによって64個の標準ベクトルとの間のユークリッド距離を並列に求める。

9.3.4 識別

大分類処理で得られた距離値の小さい候補カテゴリに対し、次式で与えられる投影距離H

$$H = D - \sum_{j=1}^p \left| \sum_{i=1}^{256} s_{jic} (x_i - \bar{x}_i) \right|^2 \quad (9.2)$$

を求め、その距離の小さいカテゴリを最終的な候補とする処理である。ここで、 s_{jic} は部分空間行列の要素、Pは部分空間の次元数である。

図9.7は(9.2)式の右辺第2項の積和演算の実行方法を示している。はじめに、各PEで標準ベクトルとの差分 Δx を求める。次いで、その値とAM内に格納しておく部分空間行列Sの1行目との間で乗算を行い、その乗算結果を右端のPEを始点とする左方向の伝搬演算により累積し、 $\sum s_{lic} (x_i - \bar{x}_i) [=V_i]$ を左端のPEに得る。この伝搬演算を利用した乗累算を、始点を右方向に1つずつずらしながらp回繰り返し、PEの左端から右側に向う $\sum s_{lic} (x_i - \bar{x}_i) \sim \sum s_{pic} (x_i - \bar{x}_i) [=V_p]$ の配列を得る。これらの累算結果をさらに2乗してから、左方向の伝搬演算を行うことによって、(9.2)式の右辺第2項を求める。

9.4 速度性能の評価

前節の処理方法に基づいたプログラムを作成し、各処理の所要時間を測定した。その測定結果をLISCAR1の値と比較して表9.1に示す。この表から明らかなように、特徴抽出処理から識別処理まで処理内容はかなり高度化しているにも関わらず、個々の処理では3~9倍、全体では5倍程度のLISCAR1に比した高速化が達成されている。以下、個別に高速化の要因を分析する。

特徴抽出処理は、9倍とかなりの高速化が達成されている。これには、9.3.1

(1)で述べた文字パターンの構成ビットラインを走査することに繰り返す必要のあった②と③の処理が、転送回転型メモリアクセス機構を用いた間接アドレッシングにより黒点連結の長さのテーブルをアクセスすることによってほとんど一回分の処理にまで短縮

表9.1 漢字認識に必要な主要な処理の所要時間(ms)の比較

処理内容	LISCAR1	LISCAR2
特徴抽出	39.4 * 1	4.5 * 2
次元圧縮	11.3 * 3	2.2 * 4
大分類	18.8 * 5	5.9 * 6
識別	—	0.7 * 7

- * 1 方向寄与度が4方向の外郭方向寄与度特徴抽出
- * 2 方向寄与度が8方向の外郭方向寄与度特徴抽出
- * 3 768 x 256の係数行列と768次元の特徴ベクトルとの積
- * 4 1,536 x 256の係数行列と1,536次元の特徴ベクトルとの積
- * 5 シティブロック距離計算、カテゴリ数：4,096
- * 6 ユークリッド距離計算、カテゴリ数：4,096
- * 7 大分類で識別困難な入力文字の2割の文字に対してのみ行う。
詳細識別対象文字数：64、部分空間次元数：8

できたのが最も寄与している。もちろん8ビット単位の演算、AMに対するロード、ストアと内部演算の並列実行、転送回転型メモリアクセス機構による黒点連結の長さのテーブルの直並列変換等も寄与している。

次元圧縮、大分類については演算単位を8ビットに拡張し、かつ乗算器を各PEに内蔵し、乗算とAMからの乗数、被乗数のロードを並行実行可能としたことが効いている。転送回転型メモリアクセス機構の隣接PE間シフト機能によりシストリックアルゴリズムが利用できるようになった効果も大きい。識別は、乗算とロードの並行実行機能に加え、ツリー構成の伝搬演算機構により伝搬加算が高速化されたのが効いている。

ところで、LISCAR2で新たに組込んだ転送回転型メモリアクセス機構とツリー構成伝搬演算機構の高速化の効果を比較してみると、ツリー構成の伝搬演算機構の効果が以外に小さいことが明らかになった。これは、転送回転型メモリアクセス機構が特徴抽出処理の高速化に大きく貢献したのに対し、ツリー構成伝搬演算機構は方向寄与度特徴の区間内累積処理と識別処理の一部に伝搬加算が利用できたにとどまったためである。このように利用できる処理が少なかったのは、総和あるいは累算について1ビット単位の実行が必要な伝搬加算より、8ビット単位で実行可能なシストリックアルゴリズムが有利と判断し、極力シストリックアルゴリズムを利用するようにしたからである。しかし、前節で示した並列処理法の検討の過程で、シストリックアルゴリズムのLISCAR2への組み込みは、伝搬加算に比べるとかなり難しいことも判明した。こ

れは、シストリックアルゴリズムでは、係数行列をねじった形式のスキュード配列形式でAM内に格納する必要があるだけでなく、配列の任意の区間での総和が困難なためである。このシストリックアルゴリズムの問題を解決するには、バイパス型の伝搬演算機構を用いて伝搬加算の単位を8ビットあるいはそれ以上に拡張するのが、最も効果的であると考えられる。

9.5 むすび

本章では、前章の手書き認識処理に対し、

- ①方向寄与度を4方向から8方向に拡張する、
- ②識別を大分類、詳細識別の2段階で実行するように変更する、
- ③大分類の距離計算をシティブロック距離からユークリッド距離に変更する、
- ④詳細識別に投影距離を採用する

等の改良を加えた新しい手書き認識処理を、LISCAR2で効率良く実行する処理法を明らかにするとともに、その処理速度を実機を用いて評価した。

はじめに、LISCAR1での高速化のネックが、外郭方向寄与度特徴抽出において、外郭部の黒点連結の長さを投影軸のグレイスケールラインに投影する処理が効率良く実行できないことにあることを指摘し、LISCAR2ではこの処理を転送回転機構による間接アドレッシング機能を用いて高速化する方針を示した。続いて、この方針に基づく、文字パターンの走査によりあらかじめ求めておく外郭点のアドレス値で、配列データメモリ内に形成する黒点連結の長さのテーブルを間接アドレッシングすることにより直接アクセスする投影処理法を示した。

さらに、次元圧縮の行列とベクトルの乗算をシストリックアルゴリズムを用い転送のオーバーヘッドなく実行する方法、LISCAR2の各PEの乗累算機能とツリー構成伝搬演算機構の高速の伝搬演算機能を組合わせて投影距離を効率良く求める方法を示した。

最後に、LISCAR2にこれらの処理を搭載し処理速度を評価し、方向寄与度が8方向に拡張されたり、シティブロック距離からユークリッド距離に変更されたりして、処理量がかなり増加しているにも関わらず、LISCAR1の速度に比べ、それぞれの処理で3~9倍、全体の平均では5倍程度の高速化が達成されていることを確認した。

第10章 結論

10.1 研究の内容と主な成果

L SIの集積能力をほぼ完全に引出せる単純かつ規則的なセルラ配列構成を前提としたSIMDプロセッサを取上げ、そのPE配列部構成法、L SIおよびシステム構成法、演算処理法等を追究するとともに、L SI-CADおよび文字認識処理への応用を試みた。その結果、次の成果を得ることができた。

- ①隣接結合ネットワークの遠隔PE間の転送性能を補強する2種類のツリー構成の伝搬演算機構を提案したこと。
- ②効率的な90度回転、PE単位の自律的なアドレッシング等を可能とする転送回転型メモリアクセス機構を提案したこと。
- ③PE単位で個別に設ける制御レジスタにより実現される可変のPE配列構造とそれによって実現される各種の機能を示したこと。
- ④1ビット構成と8ビット構成の2種類のPE配列部用L SIの構成法、設計法を明らかにし、実チップを実現したこと。
- ⑤3台のセルラ配列型SIMDプロセッサ（小型機：LISCAR1，LISCAR2，大型機：AAP2）の構成・設計法を明らかにし、試作機を実現したこと。
- ⑥AAP1のL SI-CADへの応用で、クロックサイクル同一の条件で、汎用コンピュータの50~500倍の速度性能が得られることを実証したこと。
- ⑦LISCAR1の文字認識処理への応用で、主要な認識処理アルゴリズムを専用ハードウェアに匹敵する13~46文字/秒の速度で実行できることを実証したこと。さらに、LISCAR2により、より高度な手書き漢字認識処理がLISCAR1の手書き漢字認識処理に比べ5倍程度高速化されることを示したこと。

以下、これらの本研究で得られた成果の内容を詳しく述べる。

(1) ツリー構成伝搬演算機構の提案

隣接結合ベースのネットワークの遠隔PE間の転送性能を改善する手段として、バイパス、セレクトティブ型の伝搬演算機構を多段に階層化することから得られるN進ツリー構成の伝搬演算機構を提案した。伝搬演算機構は、送信PEから受信PEまでの各PEが同一の演算を行い、その結果を順次伝搬させて受信PEに最終的な演算結果を得る機

構であり、隣接結合のみのネットワークでは最終的な結果を得るまでに送信PEから受信PEまでの距離Mに比例した時間がかかる。これに対し、本ツリー構成伝搬演算機構では、伝搬演算時間が、 $O(N \log_N M)$ と高速化され、かつ、Nを大きくすることによってハードウェア規模が低減されることを示した。さらに、バイパス構成とセレクトティブ構成の比較から、セレクトティブ構成が、算術演算のビット幅については1ビットに限定されるものの、同程度のハードウェア規模で2倍の速度性能の得られることを示した。

(2) PE配列の可変構造とその演算アルゴリズムの提案

PE配列の稼働率の向上を狙って、複数PEの協調的動作により構成される演算ユニット(PU:Processing Unit)を問題向きに再構成する適応的(Adaptive)な可変構造を提案した。また、PUの再構成は、全PE共通に配る命令を個別に修飾するPE毎の制御レジスタを必要に応じて書き換えて、各PEの役割分担を変更することによって実現され、送信、受信PEの任意設定が可能な伝搬演算機構と組み合わせることで、PUごとの総和演算、行列計算等が効率的に実行されることを示した。

(3) 転送回転型メモリアクセス機構の提案

PEごとに1個ずつ設けるルーティングレジスタの並びを、本来のPE間転送パスとしてだけでなく、一定個数のPE(PE部分配列)毎に、局所メモリ構成用のメモリユニットに対するアドレス生成器、あるいは部分ビットプレーンの90度回転器として利用するデータアクセス機構を提案した。そして、それらのデータ転送機能、アドレス生成機能、90度回転機能を組み合わせることで、SIMDプロセッサの汎用化に有効な多値プレーンの90度回転、2次元アクセス、間接アドレッシング等を効率よく実効できることを示した。また、セルラ配列型SIMDプロセッサで本来必要とする隣接PE間転送パスがほぼそのまま流用できるため、わずかの端子数、ゲート数の増加でPE配列に組込めることも示した。

(4) 大規模PE配列LSIチップの実現

外付けのメモリICとの対を規則的に並べるだけでPE配列の構成が可能な単純で規則性の高いPE配列LSI(AAP2-LSI)の構成、設計法を明らかにし、 $2\mu\text{m}$ のCMOS技術を用いて1ビットPE64個からなるチップを実現した。また、その構成、設計法である以下の3つの技術、

①トランジスタ当たりの機能の高い出力レベル補償型nMOSパストランジスタ論理

の多用、

- ② 8×8 の論理的なPE配列を縦長PEの 4×16 の配列とする高密度レイアウト、
- ③単純かつ規則的なセルラ配列構成によって可能となったPE部マスクパターンの人手設計

により、スタンダードセルベースのフルカスタムLSIに比べ、4~5倍の高集積度を實現し、目論見通りLSIの集積能力を十分に活かせることを実証した。

さらに、1ビット構成のAAP2-LSIの限界を克服すべく、 $0.8\mu\text{m}$ のCMOSゲートアレイ技術を用いて、8ビット構成PEの配列を内蔵するPE配列LSI(AAP3-LSI)を試作した。各PEのレジスタファイルを機能制限し、ハードウェアマクロのRAMを8分割することで構成する、演算能力を外付けの標準DRAMのアクセス速度に合わせる等の設計手法により、8ビット乗算器、ツリー構成伝搬演算機構、転送回転型メモリアクセス機構等を組み込みながら、1チップに16PEの搭載に成功している。

(5) 大規模システムおよび小形経済的システムの実現

8ビットの固定小数点加減算で7,355MOPSものピーク性能が得られる配列サイズが 256×256 の超並列構成セルラ配列型SIMDプロセッサ(AAP2)を試作した。この中で、高密度の両面実装技術を開発し、1,024個のAAP2-LSIと8,192個の64KビットSRAMからなるPE配列をわずか 0.2m^3 の容積に実装できることを示した。また、新しいボードおよびラック間の接続構成技術を開発し、均一な接続距離を保ったままで配列サイズを拡張できることを示した。

ビットライン単位の算術論理演算に対し2,860MOPSものビット演算性能が得られる小型のセルラ配列型SIMDプロセッサ(LISCAR1)も試作した。このプロセッサは、4個のAAP2-LSIに汎用のDSP LSIと標準DRAMを組み合わせることで、B4サイズのボード1枚に実装することに成功している。

さらに、同様の構成の1ボードプロセッサ(LISCAR2)をAAP3-LSI4個を用いて試作した。LISCAR1に比べ、配列サイズが全体で 8×8 と $1/4$ になっているにも関わらず、乗算を用いるユークリッド距離計算で2.5倍、乗算を用いないシティブロック距離計算、90度回転で、それぞれ2倍および4倍以上の性能が得られることを示した。

(6) AAP1のLSI-CADへの応用での高速性の実証

AAPの応用先として超高速処理が必要でかつSIMD処理向きのLSI-CADで

ある論理シミュレーションとマスクパターン設計の自動配線を取上げ、それぞれの処理方法を明らかにし、AAP1を用いてその処理性能を評価した。

その結果、LSI-CADについては、論理シミュレーションが1ゲートを1PEで、ゲート間の結線を伝搬転送でエミュレートするシミュレーション法を用いることによって、クロックサイクル同一の条件で汎用コンピュータの50~500倍の速度性能が得られることを明らかにした。また、自動配線が、迷路法にライン探索アルゴリズムを組み合わせるセルラ配列型SIMDプロセッサ向きアルゴリズムを、PE間のシフト転送と伝搬演算で活用することにより、クロックサイクル同一の条件で汎用コンピュータの100倍以上の速度で実行されることを明らかにした。

(7) LISCARの文字認識処理への応用での高速性の実証

小型経済性の要求される文字認識装置への組込をねらい、従来の専用ハードウェアでは困難であった複数の認識アルゴリズムのLISCAR1への搭載を試みた。その結果、前処理、特徴抽出、識別の各処理の並列処理方法に加え、2ないし4文字並列処理、距離計算の語長短縮、逐次処理バイパス等の高速化手法を明らかにし、処理内容の大きく異なる手書き英数カナ文字、印刷漢字、手書き漢字に関するこれまで開発してきた主要な認識アルゴリズムの全処理の実行が、それぞれ23、46、13文字/秒の速度で実現されることを実機上で確認した。

さらに、処理速度の劣る手書き漢字認識処理に対し、高速化と一層の認識精度向上の両立を目指し、LISCAR2の適用を試みた。外郭方向寄与度特徴を4方向寄与度から8方向寄与度に、識別処理をシティブロック距離単独からユークリッド距離による大分類と投影距離による詳細識別の2段階構成へと認識処理を高度化しているにも関わらず、LISCAR1に比べ5倍程度の高速化が可能になることを実機上で確認した。

10.2 今後の研究の方向と課題

本研究は開始時よりはば15年が経過した。その間、AAP1, 2は試作レベルに留ったのに対し、LISCAR1は、その小型経済性とプログラムの変更により種々の認識処理が搭載できることが評価され、文字認識装置の認識処理エンジンとして実用化され、一定の成功を納めるにいたった。

これに対して、逐次処理プロセッサの進歩は予想以上で、本研究の開始と期を同じくして、やはりLSIの設計ネック回避を背景に登場した汎用処理指向のRISCプロセッサが著しい進歩をとげ、CISCプロセッサも巻き込む形で、大幅な性能向上を果たしている。この逐次処理プロセッサの飛躍的発展には、登場の背景からは想像もつか

ぬ程、複雑化したRISCプロセッサを短期間に実現可能とするまでの設計ツールや設計技術の進歩が大きく貢献している。この結果、現在では機器組込み用のRISCプロセッサが、各種の専用プロセッサを駆逐したばかりか、専用の逐次処理プロセッサともいべき信号処理用プロセッサ(DSP)の領分まで脅かす存在となっている。

LISCARに対する現状も同様であり、RISCプロセッサに対する性能上の優位性は急速に失われ、今後の発展は有りえないようにさえ見える。

もちろん、このLISCARの競争力の低下の原因は主に開発力の差であって、セルラ配列型SIMDプロセッサ自体の本質的な優位性が失われたからではない。しかし、汎用性やプログラミングの容易性で劣り、巨額の開発投資に耐えるだけの需要が見込めない以上、これまでと同じようなアプローチ、同じような応用を指向しているのでは、RISCプロセッサに太刀打ちできなくなるのは明らかである。以下では、このような状況のもとで今後とすべき研究の方向と解決すべき課題についてまとめる。

(1) 新たな応用分野の開拓

これまでLISCARの有力な応用先であった文字認識処理の分野においても、一部の高精度の認識アルゴリズムを除き単独のマイクロプロセッサで必要な速度と精度を満たせるようになって来ている。現在では、100MIPSを越えるローコストのRISCマイクロプロセッサが利用できるようになったからである。こうなるとコストで劣るLISCARのような専用プロセッサを文字認識処理にわざわざ使う必要はなくなってくる。

しかし、これを持って、専用プロセッサの応用分野がしだいに狭まって来ていると考えるのは早計である。マルチメディア処理のようにSIMDプロセッサ向きの定型的な処理を大量に必要とする分野が新たに出現しているからである。事実、TI社で開発された1次元構成のセルラ配列型SIMDプロセッサ³⁴⁾は、テレビ受像機の前処理用LSIとして大きな成功を治めている。また、HPやSUNに用いられているRISCプロセッサには、画像圧縮の符号化、複合化に対応するためにSIMD型の処理機構が組込まれている⁷⁾さらに、最近では、SIMD処理機構を組込むことでマルチメディア処理全般への応用を目指したメディアプロセッサなる専用プロセッサLSIが出現している。今後は、このような分野にすばやくアプローチし、必要に応じてアーキテクチャを適合化することで応用領域を広げていく努力がますます重要になると考えられる。

(2) 技術の進歩に応じて、逐次的に高性能化が可能なアーキテクチャの追究

汎用プロセッサの世界では、一度作ったアーキテクチャを10年20年と発展させて行くのは常道であり、アーキテクチャの開発時にその発展性に十分な考慮が払われている。しかし、LISCAR1では、アーキテクチャを維持したままで小型化、高性能化をはかることが困難で、1から2への移行時に全面的なソフトウェアの変更を余儀なくされた。これは製品開発の側からすると受入れがたいものであった。専用ハードウェアに対し、バージョンアップや新製品開発のコストが低いという専用プロセッサの利点が半減されてしまうからである。

このような状況に至った原因を考えると、文字認識処理の特性に対するアーキテクチャ上の考慮が不十分であっただけでなく、開発当初に数年後のバージョンアップのビジョンまで描き切れていなかったことによるところが大きい。もっとも、後者の問題については、いまもって解決の方向が見えているわけではない。応用やアーキテクチャが変れば小型化、高性能化のための対応も大きく変ってくるからである。従って、今後、アーキテクチャの大幅な見直しをはかったり、異なる応用を目指す際には、互換性を維持したまま小型化、高性能化できるようにすることは、セルラ配列型SIMDプロセッサの専用プロセッサの開発を進める上でも基本的な課題といえる。

(3) セルラ配列型PE配列LSI構成の見直し

セルラ配列構成LSIの基本的な考え方は、規則性の高いPE配列のみ搭載するというものである。これは、本研究開始時には設計が単純化される、容易に高い集積度が得られる、LSIを規則的に並べるだけで任意のサイズのPE配列を構成できる等大きな利点があると考えられた。しかし、現在ではこれらはあまり利点とは言えなくなっている。なぜならば、設計の生産性が上がり設計の単純性に固執する必要がなくなってきたり、小型のプロセッサを目指す場合には集積度が上がり必要な数十程度の並列度は十分1チップに入るようになってきているからである。また、並列度自体を大きくできたとしても、文字認識処理における2文字並列や4文字並列のような手法を使わないと、その能力を引出せないのならば、プログラミングを難しくするばかりで、利点が少ないことも明らかになってきたからである。これは、換言すると、プログラミングの観点からは、高速化の程度が同程度ならば、並列度を大きくするよりもPE自体の性能改善をはかる方が利点大きいといえる。

こうなると1チップ化可能な範囲の並列度とし、周辺回路を極力取込んで、1チップマイクロプロセッサ同様な使い方ができるようにする方が、使いやすさや経済性の点で利点大きいと考えられる。今後はこういった方向でのアーキテクチャの追究が成功のポイントとなる。

(4) ソフトウェア開発に対する拒否反応の払拭

本論文では専用プロセッサ指向と言うことで、ソフトウェアの開発容易性の点には、触れないようにしてきたが、実はかなり深刻な問題である。何が深刻かといえば、マクロアセンブリレベルの言語に対する周囲の拒否反応が強いことである。確かに、このような専用言語をマスタするには、1、2ヶ月の期間を要する。また、習熟してもCやFORTRANのような言語のように抽象的な記述ができるわけでもない。しかし、ひとたび高速あるいはスタティックステップ数の小さいプログラムを作ろうとした場合には汎用プロセッサと比べそれほど生産性に違いがでるわけではない。逐次型の汎用プロセッサでは、自由度が高い分、最適な処理法を見つけるのに試行錯誤を繰り返す必要があるからである。同一内容のプログラムの高速化の経験からすれば、生産性は、マクロが整備されている条件では、高級言語による汎用プロセッサのプログラミングに比べ、高々2、3分の1に低下するだけであった。もちろん、これには、インタラクティブなデバッガを利用すれば逐次型の汎用プロセッサと同様な感覚でデバッグできるSIMDプロセッサのデバッグ容易性も寄与している。

いずれにしても、生涯に記述されるプログラムの総ステップ数が数十Kにも及ばない可能性のある専用プロセッサに対して、高性能の高級言語コンパイラを用意するわけにはいかない。専用言語ベースのプログラム開発環境を提供するだけでもかなりの負担である。今後はより開発コストの低いプログラム開発環境の構築方法を明らかにするとともに、習熟期間の短い専用言語の仕様を追究し、専用言語に対する拒否反応を払拭していく必要がある。

謝 辞

本論文をまとめるに当たり、懇切なご指導とご教示を賜った名古屋大学工学部稲垣康善情報工学科教授、鳥脇純一郎情報工学科教授、鳥田俊夫電気工学科教授、ならびに高木直史情報工学科助教授に厚く感謝いたします。また、筆者が名古屋大学大学院在学中の心温まる御指導とその後の変らぬ励ましにより本論文の執筆へと導いて下さった故赤尾保男元名古屋大学工学部電気工学科教授に深く感謝いたします。

本論文は、筆者が1979年から1993年にかけてNTT武蔵野電気通信研究所集積回路研究部、厚木電気通信研究所カスタム化技術研究部、横須賀電気通信研究所宅内機器研究部、LSI研究所LSI開発研究部、設計システム研究部、ヒューマンインタフェース研究所マルチメディア処理研究部、第5プロジェクトチームにおいて行った研究の成果をまとめたものであります。これらの各所で本研究の機会を与えて頂いた須藤常太元カスタム化技術研究部部長、小森和昭NTTインテリジェントテクノロジー株式会社社長（元宅内機器研究部統括役）、酒井保良LSI研究所長（元設計システム研究部長）、徳永幸生映像処理研究部部長（元マルチメディア処理研究部グループリーダー）、北村正元第5プロジェクトチームリーダー、武田和光元設計システム研究部グループリーダー、中川亨元マルチメディア処理研究部グループリーダーに感謝いたします。

本研究を進めるに当たり、直接ご指導いただきました中島孝利NTTエレクトロニクステクノロジー営業本部副本部長（元集積回路研究部研究室長）、青木誠元交換方式研究部主幹研究員、宮原末治メディア応用システム研究部主幹研究員、多田俊吉関連企業本部担当部長（元言語メディア処理研究部主任研究員）、川谷隆彦元第5プロジェクトチーム主幹研究員、杉山吉関連企業本部担当部長（元カスタム化技術研究部主幹研究員）に心からお礼申し上げます。

また、本研究は多くの方々の支援、協力のもとで行ってきたものであります。特に、大規模セルラ配列型SIMDプロセッサAAPの試作とLSI-CADへの応用に関しては、渡辺琢美高速LSI研究部主幹研究員、北村美宏広帯域研究部主幹研究員、および土屋敏雄関連企業本部担当課長から、小型高並列プロセッサLISCAR1, 2の試作と文字認識への応用に関しては、宮原末治メディア応用システム研究部主幹研究員、多田俊吉関連企業本部担当部長、木村義政メディア応用システム研究部主任研究員、宮本信夫メディア応用システム研究部主任研究員、下原勝憲画像通信処理研究部グループリーダー、および曾根原登映像処理研究部グループリーダーから、多大の支援と協力をいただいたことに深く感謝いたします。

最後に、本論文執筆の機会を与えて下さった唐津高機能LSI研究部部長、笠井高機能LSI研究部グループリーダーに感謝いたします。

参考文献

- (1) S. H. Unger: "A computer oriented toward spatia problem", Proc. IRE. Vol.46, No.10, pp.1744-1750(1958)
- (2) B. H. McCormick: "The Illinois pattern recognition computer ILLIAC III", IEEE Trans. Electronic Computers, EC-12, No.12, pp.791-813 (1963)
- (3) G. H. Barnes, et al.: "The ILLIAC IV Computer", IEEE Trans. computers, Vol.C-17, No.8, pp.746-757 (1968)
- (4) 加藤, 苗村: "並列処理計算機", オーム社(1976)
- (5) M.Griffin, G.Tahara, K.Knorpp, R.Pinkham, B.Riley: "An 11-Million Transistor Neural Network Execution Engine", ISSCC'91 Digest of Technical Papers, pp.180-181(1991)
- (6) D. A. Patterson and C.H.Sequin: "Design consideration for single chip computer of the future", IEEE Trans. C-29, No.2, pp.108-115(1980)
- (7) H. T. Kung: "Why systolic architecture?", IEEE Computer, Vol.15, No.1, pp.37-46(1982)
- (8) M. J. B. Duff and D. M. Watson: "The cellular logic array image processor", Comp. Journal, Vol.20, pp.68-72(1977)
- (9) 宮田裕行: "高速画像処理向きセルラ・アレイ・プロセッサ", 情処学会計算機アーキテクチャ研究会資料, CA53-6(1984)
- (10) R. Brooks and R. Lum: "Yes an SIMD mashine can be used for AI", Proc. of IJCAI, pp.73-79(1985)
- (11) K. E. Batcher: "STARAN parallel processor system hardware", Proc. of National Computer Conference, pp.405-410(1974)
- (12) P. M. Flanders, D. J. Hunt, S. F. Reddaway, D. Parkinson: "Efficient high speed computing with the distributed array processor", in High Speed Computer and Algorithm Organization, New York:Academic, pp.113-128(1977).
- (13) M. J. B. Duff et al.: "Review of the CLIP image processing system", Proc. of National Computer Conference, pp.1055-1060(1978)
- (14) K. E. Batcher: "Design of a massively parallel processor", IEEE Trans.Computers, C-29, No.9, pp.836-840(1980)
- (15) K. E. Batcher: "Bit-serial parallel processing systems", IEEE Trans.Computers, Vol.C-31, No.5, pp.374-384(1982)
- (16) W. Daniel hillis: "The connection machine", An ACM Distinguiushed Dissertation, The MIT Press, London(1985)

- (17) T. Blank: "The MasPar MP-1 architecture", CompCon Spring 90, San Francisco, pp.20-24 (1990)
- (18) H. J. Siegel, J. D. Armstrong and D. W. Watson: "Mapping computer-vision-related tasks onto reconfigurable parallel-processing systems", IEEE Computer, vol.25, No.2, pp.54-63 (1992)
- (19) A. H. Karp, D. Heller and H. Simon: "1993 Gorden bell prize winners", IEEE Computer, vol.27, No.1, pp.69-75(1994)
- (20) J. Beetem, M. Denneau and D. Weingarten: "The GF11 supercomputer", 12th symposium on Computer Architechture Conf. Proc., pp.108-115(1985)
- (21) T. Feng: "A survey of interconnection network", IEEE Computer, Vol.14, No12, pp.12-30 (1981)
- (22) A. P. Reeves: "A Systematically designed binary array processor", IEEE Trans. Computers, C-29, No.4, pp.278-287(1980)
- (23) G. E. Blelloch: "Scans as primitive parallel operations", Proc., Int.Conf.on Parallel Processing, pp.355-362(1987)
- (24) G. E. Blelloch and J. J. Little: "Parallel solutions to geometric problems on the scan model of computation", Proc., Int. Conf. on Parallel Processing, pp.218-222(1988)
- (25) J. J. Little, G. E. Blelloch and T. A. Cass: "Algorithmic techniques for computer vision on a fine-grained parallel machine", IEEE Trans. Pattern Anal. & Mach. Intell., Vol.11, No.3, pp.244-257(1989)
- (26) "Connection machine model CM-2 technical summary", Thinking Machines Corporation Version 5.1(May 1989)
- (27) K. Hwang(堀越監訳):"コンピュータの高速演算方式", 近代科学社, pp82-88(1979)
- (28) J. D. Ullman: "Computational aspects of VLSI", Computer Science Press (1984)
- (29) 岡田, 田島, 森: "新しい可変構造プロセッサアレイ - E T L A N M A -", 情処学会研報, CA-24, pp.95-104(1977)
- (30) R.W. Hockney, C.R. Jeshope(奥川, 黒住訳): "並列計算機", 共立出版, pp.141- 156 (1984)
- (31) L. Snyder: "Introduction to the configurable, highly parallel computer", IEEE COMPUTER Vol.15, No.1, pp.47-56(1982)
- (32) Y. Kitamura, T. Watanabe and K. Matsuhiro: "CAD oriented 2-dimensional array processor - part2:routing and logic simulation", IFIP Workshop on CAD Engines(1987)

- (33) 藤田善弘: "画像演算メモリ", 情処学会計算機アーキテクチャ研究会資料, 91-ARC89-1, pp.1-8(1991)
- (34) J. Childers, P. Reinecke, H. Miyaguchi, Y. Takahashi, Y. Yaguchi and M. Takeyasu: "SVP: serial video processor", IEEE 1990 CICC, 17.3.1(1990)
- (35) N. Yamashita et al.: "A 3.84GIPS integrated memory array processor LSI with 64 processing elements and 2Mb SRAM", IEEE ISSCC Digest of Technical Papers, pp.260-261(1993)
- (36) K.E. Batcher: "The Multidimensional access memory in STARAN", IEEE Trans.C-26, No.2, pp.174-177(1977)
- (37) 元岡達, 田中英彦, 上森明, 鈴木達郎: "2次元記憶を用いた連想記憶システム", 信学技報, EC76-80(1976)
- (38) D. Van Voorhis and T. Morrin: "Memory systems for image processing", IEEE Trans. Computers, C-27, Vol.2, pp.113-125(Feb.1978)
- (39) D. Hayes and B. Strawhorns: "The application of multi-dimensional access memories to high performance signal processing systems", Proc. ICASSP, pp.1412-1415(1985)
- (40) 田畑邦晃, 武田晴夫, 町田哲夫: "ラスト走査とテーブル参照による画像回転の高速処理", 信学論(D), J69-D, No.1, pp.80-90(1986)
- (41) E.R. Komen: "Efficient parallelism using indirect addressing in SIMD processor arrays", Pattern Recognition Letters 12, pp.279-289(1991)
- (42) C.L. Kuszmaul: "Rapid transpose methods on massively parallel SIMD Computers", The society for industrial & applied mathematics annual meeting(July 1990)
- (43) 中山晶, 木村文隆, 吉田雄二, 福村晃夫: "パイプライン型並列画像処理アルゴリズムの一提案", 信学論(D), J71-D, No.9, pp.1879-1882(1988)
- (44) T. Sudo, T. Nakashima, M. Aoki and T.Kondo: "An LSI adaptive array processor", IEEE ISSCC Dig. Tech. Papers, pp.122-123(1982)
- (45) T. Kondo, T. Nakashima, M. Aoki and T. Sudo: "An LSI adaptive array processor", IEEE J. Solid-State Circuits, C-18, Vol.2, pp.147-156(1983)
- (46) C. Mead and L. Conway: "Introduction to VLSI systems", Addison-Wesley (1980)
- (47) 佐々木, 矢野, 力野, 関: "パストランジスタ回路集積化技術: L E A P ", 1994年信学会秋期全国大会, A-64, pp.64(1994)
- (48) 近藤: "CMOSトランスファゲート回路の比較検討", 昭和59年信学会総合全国大会 589, pp.2-352(1984)
- (49) W. R. Iversen: "New CMOS chip processes data in parallel", Electronics Week, November 12, pp.17-18(1984)

- (50) 前田: "V L S I 画像処理プロセッサ", 情処学会誌 Vol.31, No.4, pp.480-485 (1990)
- (51) T. Kondo, T. Nakashima, T. Tsuchiya, Y. Sugiyama and T. Sudo: "A Large scale cellular array processor: AAP-1", Proc., 1985 ACM Computer Science Conf., pp.100-111(1985)
- (52) 近藤, 杉山, 中島: "2次元アレープロセッサ(A A P 2)とプログラミング言語", 信学論, Vol.J71-D, No.8, pp.1399-1406(1988)
- (53) 曾根原, 近藤, 木村, 内田, 平岩: "ニューロ処理指向並列プロセッサ (Neurocessor) 及びこれを用いた学習型個人適応システムの開発", 1993年信学会秋期全国大会, D-31, pp.6-33(1993)
- (54) 杉山, 渡辺: "A A P による論理モジュール並列配置", 信学技報 CAS84-12, pp.25-32(1984)
- (55) M. A. Breuer and K. Shamasa: "A hardware router", Journal of Digital Systems, Vol.4, No.4, pp.393-408(1980)
- (56) T. Blank, M. Stefik and W. Vancleemput: "Parallel bit map processor architecture for DA algorithms", IEEE 18th Design Automation Conference, pp.837-845(1981)
- (57) W. Heyns et al.: "A line-expansion problem with a guaranteed solution", Proc. 17th Design Automation Conference, pp.243-249(1980)
- (58) T. Watanabe, H. Kitazawa and Y. Sugiyama: "Parallel adaptable routing algorithm and its implementation on a two dimensional array processor", IEEE Trans. Comput.-Aided Des. Integrated & Syst., CAD-6, Vol.2, pp.241-250(1987)
- (59) M. Hanan: "On steiner's problem with rectilinear distance", SIAM J. Appl.Math.14, pp.255-265(1966)
- (60) 赤松, 塩, 飯田, 川谷: "手書き漢字用文字読取装置", 研実報, Vol.36, No.4, pp.579-587(1987)
- (61) 宮本, 堤田, 中島, 川谷: "英数カナ用高精度文字読取装置", 研実報, Vol.37, No.2(1988)
- (62) 上, 岡本, 天満, 浅井: "階層化判別法とその文字認識システム P C - O C R", 信学技報 PRU-86-76, pp.49-55 (1987)
- (63) 松村: "卓上型手書き漢字 O C R の開発", 信学会総合全国大会1540, pp.6-124 (1985)
- (64) 吉本, 安部, 小森: "高速印刷文字認識装置の開発", 信学技報 PRU88-16, pp.1-7(1988)
- (65) 多田, 宮原, 木村, 川谷: "L I S C A R を用いた手書き帳票読取り", NTT R&D, Vol.39 No.11, pp.1603-1612(1990)

- (66) 目黒, 梅田: "マルチフォント印刷漢字認識装置", 信学論 J67-D, No.8, pp.908-915 (1984)
- (67) 宮原, 木村, 豊田, 宮田: "部分パターンによる可変ピッチ文書からの文字切出しと認識", 信学論 J72-D-II, No.6, pp.846-854(1989)
- (68) 萩田, 内藤, 増田: "外郭方向寄与度特徴による手書き漢字の識別", 信学論 J66-D, No.10, pp.1185-1192(1983)
- (69) 能見他: "手書き数字認識における誤読・リジェクトパターンの分析 第2回文字認識コンテストの実施結果より", PRU93-46, pp.25-32(1989)
- (70) 池田, 田中, 元岡: "手書き文字認識における投影距離法", Vol.24, No.1, pp.106-112 (1983)
- (71) R.B.Lee: "Accelerating Multimedia with Enhanced Microprocessors", IEEE Micro, Vol.15, No.2, pp.22-32 (1995)

研究業績リスト (*は本論分に関連する業績を示す。)

1. 学会誌論文

- (1*) T. Kondo, T. Nakashima, M. Aoki and T. Sudo: "An LSI adaptive array processor", IEEE J.Solid-State Circuits, C-18,2, pp.147-156(1983)
- (2*) 近藤, 杉山, 中島: "2次元アレープロセッサ(AAP2)とプログラミング言語", 信学論(D), Vol.J71-D, No.8, pp.1399-1406(1988)
- (3*) 近藤: "LSI化に適したツリー構成走査演算機構", 信学論(C-II), J74-C-II, No.5, pp.388-397(1991)
- (4*) 近藤, 木村, 曾根原: "1次元SIMDプロセッサ向きデータアクセス機構", 信学論(D-II), J74-D-II, No.6, pp.269-278(1993)
- (5*) T. Kondo, Y. Kimura and N. Sonehara: "Single-Board SIMD Processors using Gate-Array LSIs for Parallel Processing", IEICE Transactions on Electronics, E76-C, No.12, pp.1827-1834(1993)
- (6*) 多田俊吉, 近藤利夫, 宮原末治: "小形高並列プロセッサと文字認識への応用", 信学論(D), J71-D, No.8, pp.1546-1552(1988)
- (7) 宮原, 近藤, 多田: "SIMD型並列プロセッサを用いたフルテキスト検索", 情報学会論文誌, Vol.33, No.3, pp.397-404(1992)

2. 査読付き国際会議論文

- (1*) T. Kondo, T. Nakashima, T. Tsuchiya, Y. Sugiyama and T. Sudo: "A Large Scale Cellular Array Processor: AAP-1", Proc., 1985 ACM Computer Science Conf., pp.100-111(1985)
- (2*) T. Kondo, T. Tsuchiya, Y. Kitamura and Y. Sugiyama: "Pseudo MIMD array processor -AAP2", 13th Symposium on Computer Architecture Conf. Proc., pp.330(1986)
- (3*) T. Kondo, S. Tada and S. Miyahara: "Kanji character recognition unit with hand-scanner using SIMD processor", SPIE Vol.1001 Visual Communication and Image Processing, pp.476-483(1988)
- (4*) T. Sudo, T. Nakashima, M. Aoki and T. Kondo: "An LSI Adaptive Array Processor", ISSCC Digest of Technical Papers, pp.122-123(1982)
- (5) T. Watanabe, Y. Sugiyama, T. Kondo and Y. Kitamura: "Neural Network Simulation on a Massively Parallel Cellular Array Processor: AAP-2", Proc. of Int'l Joint Conf. on Neural Networks(IJCNN-89), pp.II-155-161(June 1989)

5. 解説論文等

- (1*) 近藤, 中島: "A A P (Adaptive Array Processor) とその応用", O plus E, No.55, pp.68-75(1984)
- (2*) 近藤, 中島: "1 ビット 2 次元 S I M D プロセッサ A A P", bit臨時増刊:並列コンピュータアーキテクチャ Vol.21, No.4(1989)
- (3*) 近藤, 多田, 宮原, 杉山: "小型並列プロセッサ L I S C A R の構成", NTT R&D Vol.39, No.11, pp.1585-1592(1990)
- (4*) 宮原, 多田, 鈴木, 近藤: "L I S C A R を用いた印刷文書読取り", NTT R&D Vol.39, No.11, pp.1593-1602(1990)
- (5) 近藤, 木村, 曾根原: "ニューロ処理指向並列プロセッサ: ニューロセッサ", NTT R&D Vol.42, No.3, pp.-(1993)

略語、定義語一覧

A	51
間接アドレッシングの際、各 P E で行うアドレス生成の所要サイクル数。	
A A P [Adaptive Array Processor]	10
本研究で試作した大規模セルラ配列型 S I M D プロセッサの名称。	
A M [Array data Memory]	109
各 P E が個別に有する局所メモリの配列を一つのメモリと見立てたもの。	
A D M [Address Modifier]	91
転送回転型メモリアクセス機構で、全 P E 共通のアドレスを、各 P E で個別に生成するアドレスで修飾するためのハードウェア (図 6. 3 参照)。	
C _{cc}	51
転送回転型メモリアクセス機構において、配列データメモリから 2 次元配列データの 1 列をアクセスするのに要するアクセスサイクル数。	
C _{cp}	54
転送回転型メモリアクセス機構において、短冊状配列データのスキュード配列形式で格納される 2 次元配列データの 1 列を配列データメモリからアクセスするのに要するアクセスサイクル数。	
C _{rc}	51
転送回転型メモリアクセス機構において、配列データメモリから 2 次元配列データの 1 行をアクセスするのに要するアクセスサイクル数。	
C _{rp}	54
転送回転型メモリアクセス機構において、短冊状配列データのスキュード配列形式で格納される 2 次元配列データの 1 行を配列データメモリからアクセスするのに要するアクセスサイクル数。	

I O B [I/O Buffer register]	89
L I S C A R 2 の各 P E に内蔵されるレジスタファイル。データ入出力用のバッファレジスタ、各種演算のワークレジスタとして用いる。	
I E [Interface Element]	25
セレクトティブ方式の伝搬演算機構で、各 P E 対応に局所伝搬演算と広域伝搬演算を行うための演算要素 (図 2. 1 2 参照)。	
L I S C A R [LIne Scannable Cellular ARray processor]	64
本研究で試作した小型セルラ配列型 S I M D プロセッサの名称。	
L M [Local Memory]	86
P E ごとに付加する局所メモリ。	
M	6
配列のサイズ。1 次元の配列データあるいはプロセッサ配列の場合には、その長さを表す。2 次元のプロセッサ配列の場合には、縦あるいは横方向の長さを表す。	
M I M D [Multiple Instruction stream Multiple Data stream]	2
複数プロセッサからなる並列プロセッサの 1 構成方式。構成プロセッサの各々が互いに異なるプログラムで、異なるデータを処理することを特徴とする並列プロセッサ構成方式。	
M D A [Multi-Dimensional Access memory]	6
多次元の配列データに対し、複数の次元方向のデータの並びを一括してアクセスすることのできるメモリ。	
N	12
伝搬演算機構において、バイパスする P E 数。	

O C R [Optical Character Recognition unit]	109
文字認識処理装置。	
P E [Processing Element]	2
セルラ配列型並列プロセッサを構成するプロセッサ。	
P O U [Propagation Operation Unit]	18
伝搬演算機構を構成する伝搬演算ユニット。	
P O E [Propagation Operation Element]	17
伝搬演算ユニット (P O U) を構成する伝搬演算要素。	
P U [Processing Unit]	33
本研究対象のセルラ配列型 S I M D プロセッサにおいて、複数の P E で構成される演算単位。	
P 1	17
伝搬演算エレメント P O E の構成要素の演算器で、局所的な伝搬演算に用いる。	
P 2	18
伝搬演算エレメント P O E の構成要素の演算器で、局所的な伝搬演算結果と直全の伝搬演算ユニットまでの伝搬演算結果とから、広域的伝搬演算結果を得るのに用いる。	
Q F P [Quad Flat Package]	90
表面実装型パッケージの一種。4 つの側面のすべてからリードピンが出ているパッケージ。	
R E [Register Element]	47
ルーティングレジスタ R T R G を構成するレジスタ要素。	

R F [Register File]	89
レジスタファイル。	
R T R G [RouTing ReGister]	45
データ転送用レジスタ (ルーティングレジスタ)。	
S	51
P E 配列内に形成されるルーティングレジスタ配列でのサイクル当たりのシフト量。	
SEL [SELector]	21,25
セレクタ。	
SA	24
セレクトティブ伝搬演算機構の伝搬演算要素 P O E 内の A 系統のパスを形成するセレクタ。	
SB	24
セレクトティブ伝搬演算機構の伝搬演算要素 P O E 内の B 系統のパスを形成するセレクタ。	
S O G [Sea Of Gates]	91
トランジスタ敷詰め型ゲートアレー L S I	
S O P [Small Outline Package]	86
表面実装型パッケージの一種。2つの側面からリードピンが出ているパッケージ。	
S T N [Selective-Tree propagation Network]	89
セレクトティブツリー構成伝搬演算ネットワーク。	

S P U [Selective Propagation operation Unit]	90
セレクトティブ方式の伝搬演算ユニット。	
T R M A P [Path for Transmission, Rotation and Memory Access]	90
転送回転型メモリアクセス機構。	
T A T [Turn Around Time]	8
設計、開発の一巡分の期間。	
T O P	17
伝搬遅延時間。	
t _{pd}	17
バイパス方式の伝搬演算要素 P O E を構成する演算器 P 1 _A 、P 1 _B の伝搬遅延時間。	
t _{pdS}	20
セレクトティブ方式の伝搬演算要素 P O E を構成する演算器 P 1 _A 、P 1 _B の伝搬遅延時間。	
V M E [Versa Module European]	96
16、32、64ビットの非同期バス。IEEE 1014標準。	
W	21
演算幅、演算語長。	
インタフェースエレメント	25
I E 参照。	
オフセット	18
伝搬演算ユニット P O U がカバーする部分伝搬領域の直前までの広域的な伝搬演算結果。	

グレイスケールライン	84
2ビット以上の語長の要素並びからなる1次元配列。	
グレイトーン	102
グレイスケールラインの要素の語長。	
次元圧縮	11,120
固有ベクトル行列による線形変換 (K-L変換 [Karhunen-Loeve transform]) を用いて、ベクトルの次元数を低減すること。文字の特徴ベクトルの次元圧縮は、特徴として機能しない次元成分が取り除かれるので、分類、識別に必要な辞書のサイズ、演算量の低減に有効である。	
スキュード配列	51
行方向あるいは列方向にねじる形で格納された配列 (図4.1参照)。	
大分類	11
文字認識処理において識別を階層的に行う際の第1段階に行う大雑把な分類。	
転送回転型メモリアクセス機構	10,45
TRMAP参照。	
伝搬演算	5,12
データ配列 a_1, a_2, a_3, a_M に対し $a_1, (a_1@ a_2), (a_1@ a_2@ a_3), \dots, (a_1@ a_2@ a_3@ \dots @ a_M)$ を求める演算であり、配列要素がPE配列に分散配置されている場合に、途中の演算結果をPE間で順次伝搬させることによって求められることから、本論分では伝搬演算と呼んでいる。高級言語APLの演算子の走査と同一機能であることから走査演算とも呼ばれる。なお、@は、+, V, \wedge , min [最小値を求める演算], max [最大値を求める演算] 等の演算子である。	
伝搬演算ユニット	18
POU参照。	

2次元アクセス	6
2次元配列データに対し、行単位でも列単位でもアクセスできる機能。	
ビット直列	2,35
1ビット単位で逐次的に扱うこと。	
ビットスライス	36
複数縦続的に接続することで、多ビットの演算器を構成することのできる数ビット幅の演算器構成をいう。	
ビットパラレル	35
複数ビットを単位に扱うこと。	
フリップネットワーク	5,46
セルラ配列型SIMDプロセッサのグッドイヤーSTARANに採用された置換ネットワーク。	
ライン	46
LISCARの処理単位の1次元配列。	
伝搬演算エレメント	17
POE参照。	

