

高安全知能集積システム用画像認識プロセッサの構成

An Image Recognition VLSI Processor for Highly-Safe Intelligent Integrated Systems

○張山昌論, 来山信康, 亀山充隆

○Masanori Hariyama, Nobuyasu Kitayama, Michitaka Kameyama

東北大学大学院情報科学研究科

Graduate School of Information Sciences, Tohoku University

キーワード : 瞬時出力生成(immediate-output generation), 空間的並列処理(spatial-parallel processing) 物体認識(object recognition), HSI変換(HSI transformation), 領域分割(segmentation)

連絡先 : 〒980-8579 宮城県仙台市青葉区荒巻字青葉05 東北大学 大学院情報科学研究科 亀山研究室
張山昌論, Tel.: (022)217-7154, Fax.: (022)263-9401, E-mail: hariyama@kameyama.ecei.tohoku.ac.jp

1. まえがき

高安全自動車、セキュリティシステム、ホームサービスロボットなどのリアルワールド応用の実現のためには、人間に対する安全性を確保するため、人間抽出が重要となる。また、このような応用においては、変化する環境に対応するために処理の高速応答性が要求される。しかしながら、画像から人間を抽出するためには膨大な計算量が要求されるために、その専用VLSIプロセッサ化が望まれる。

本稿では、空間的並列アーキテクチャに基づく人間抽出VLSIプロセッサを提案する。人間の体格には固体差があるため、認識のための特徴量としては固体差に依存しないものが必要となる。そこで、人間の手、足、頭などの接続の仕方を特徴量として、人間かどうかを判断することにより固体差に依存しない認識を行なうアルゴリズムを提案

する。そのアルゴリズムにおいては、接続情報を抽出するために細線化を用いているが、その計算量が全体の90以上を占めるため、細線化のための並列処理アーキテクチャが重要となる。

細線化部は、画像メモリ、演算部、中間結果記憶用バッファから構成される。ある与えられたVLSIプロセッサの面積制約下で処理時間を最小化するためには、中間結果記憶用バッファの容量を最小化し、演算部の並列度を高めることが重要となる。そこで、バッファ容量を最小化するために、共通入力に着目した瞬時出力生成スケジューリングを提案する。これは、共通入力を有する演算はできるだけ同時に行ない、入力データがそろった演算をできるだけ速やかに行なうという考え方である。このスケジューリングに基づき、並列に入力データを供給できる場合には、従来主流であるパイプラインアーキテクチャに比べ空間的並列アーキテクチャが有効となることが示される。

2. 人間抽出アルゴリズム

2.1 接続情報に着目した人間検出

図1に人間抽出処理のフローチャートを示す。人間の体格には個体差があるため、人間抽出のための特徴量としては、個体差を吸収できるものが必要とされる。そこで、本アルゴリズムでは、人間かどうかを判断する特徴量として接続情報を用いる。すなわち、画像から頭、手、足などの人間の各部と推測される部分の接続の仕方により、人間かどうかの判断を行っている。

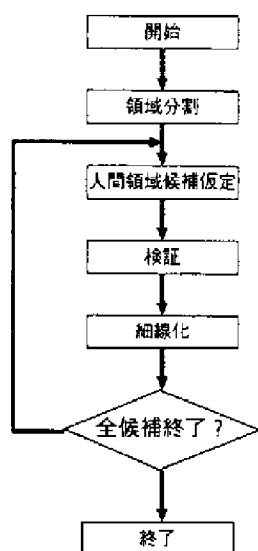


Fig. 1 人間抽出アルゴリズムのフローチャート

2.2 HSI色表現を用いた領域分割

領域分割の基本的な方法として、色による領域分割を採用する。領域を組み合わせる人間領域候補を生成するので、領域数が増加するにしたがって人間領域候補数は指数関数的に増大する。そこで、できるだけ領域数を増加させない領域分割の方法が必要となる。代表的な手法として、RGB色表現を用いた領域分割があるが、RGB色表現では本来同じ色の物体でも、明るさの影響でことなる領域になってしまう。そこで、HSI色表現を用いた領域分割を採用する¹⁾。HSI色表現はRGB色表現を式(1)により変換することで得られる。ここで、R,G,Bは正規化されている、すなわち、 $0 \leq R, G, B \leq 1$

とする。H, S, Iはそれぞれ、色相、彩度、明度と呼ばれる値である。HSI色表現では、人間が同じ色と感ずる色は、同じ色相と彩度の値を取るように変換される。このように、色と明るさを分離できるために、HSI空間でH,Sに着目し領域分割を行えば、明るさが異なる領域であっても、人間が同じ色と感ずる領域を同じ領域とすることができる。図2にRGB色表現とHSI色表現を用いた場合の例を示す。

$$\begin{aligned}
 H &= \frac{\tan^{-1}(G-B)}{2R-G-B} \\
 S &= \frac{(B-R)^2 + (R-G)^2 + (G-B)^2}{3} \\
 I &= R + G + B
 \end{aligned} \quad (1)$$

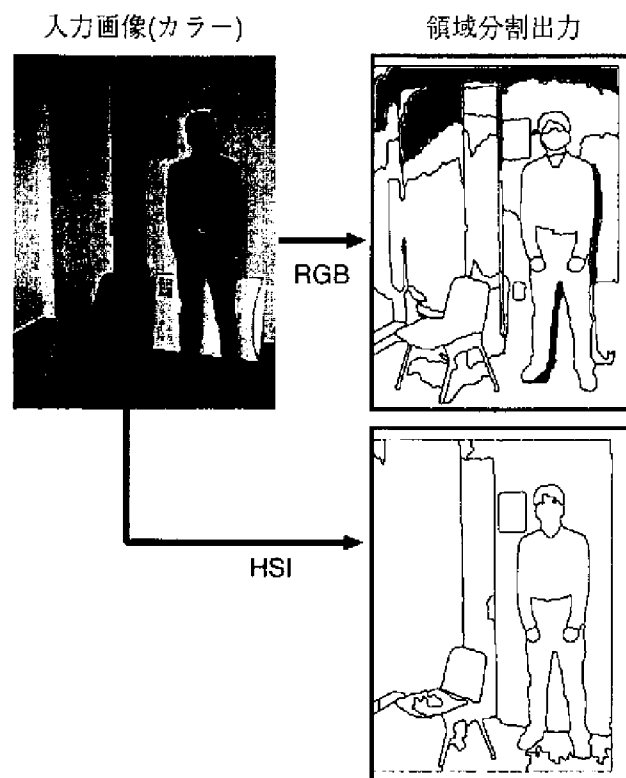


Fig. 2 HSI色表現を用いた領域分割

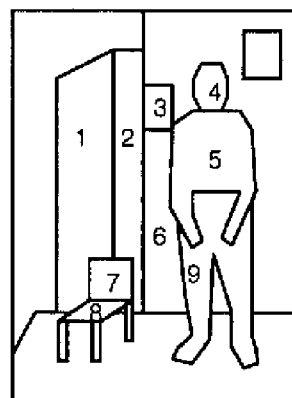


Fig. 3 領域分割結果の例

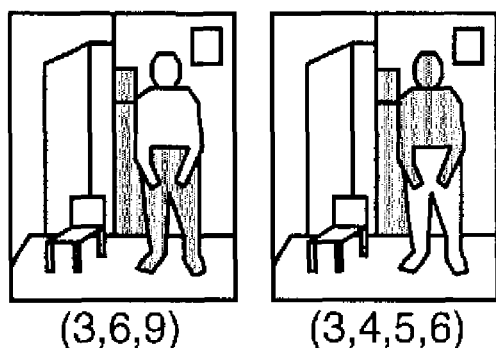


Fig. 4 人間領域候補の仮定

2.3 人間領域候補の仮定

領域を組み合わせ、人間領域の候補を仮定する。図3に示す領域を組み合わせることにより、図4に示すような人間領域候補が得られる。領域数を n とすると、 2^n 通りの人間領域候補が存在する。そこで、本稿では人間の頭部領域に着目し、候補数を減少している。頭部領域に着目する理由は、頭部は他の部分に比べ、姿勢に依存せず安定して抽出できるからである。頭部領域候補を抽出するために、次のような仮定を導入する。

- 画像内に人間の全身が写っている。
- 頭部領域は他の領域とは分けられている。

このように仮定すると、人間の頭部は人間領域候補の20%以内の大きさとなるため、大きすぎる領域は頭部領域とは成り得ない。

通常人間の体は頭部より下にある。したがって、頭部領域が限定されれば、頭部領域より下にある、頭部領域に連結している領域だけを組み合わせ、人間領域候補を生成できるため、組み合わせ数を減少できる。

2.4 並列処理向き細線化アルゴリズム

本稿では人間領域候補のシルエットに対して細線化を行い、その結果の画像の端点、交点を用いて接続情報を抽出する。細線化は図1に示す処理の中で、最も計算量がかかる処理である。一例として、人間領域候補数が13000個の場合には、SUN

SPARC Ultra 10で40時間以上もの時間がかかり、これは全体の処理時間の99.6%を占める処理時間である。このように専用プロセッサによる並列処理が最も望まれる部分ため、細線化アルゴリズムとしては、並列処理向き細線化アルゴリズムを採用する²⁾。

図5に細線化処理の手順を示す。左右上下の4方向から1画素縮小する処理を繰り返し、線の幅が1になったら終了する。通常、細線化では図6に示すような3x3の大きさの論理フィルタが用いられる。このアルゴリズムでは、各Stepにおいて、3x3論理フィルタを全画素に対して並列にかけることが可能である。以下のStepを線幅が1になるまで、繰り返すことで細線化が行われる。

Step 1.(左方向から縮小) 図7(a)のテンプレートとマッチし、かつ、図8(a)~(i)に示すテンプレートのどれかひとつとマッチする場合には、その画素に対する論理フィルタの出力は1とする。それ以外の場合には0にする。

Step 2.(上方向から縮小) 図7(b)のテンプレートとマッチし、かつ、図8(a)~(i)に示すテンプレートのどれかひとつとマッチする場合には、その画素に対する論理フィルタの出力は1とする。それ以外の場合には0にする。

Step 3.(右方向から縮小) 図7(c)のテンプレートとマッチし、かつ、図8(a)~(i)に示すテンプレートのどれかひとつとマッチする場合には、その画素に対する論理フィルタの出力は1とする。それ以外の場合には0にする。

Step 4.(下方向から縮小) 図7(d)のテンプレートとマッチし、かつ、図8(a)~(i)に示すテンプレートのどれかひとつとマッチする場合には、その画素に対する論理フィルタの出力は1とする。それ以外の場合には0にする。

図9に人間領域候補に対する細線化結果の例を示す。

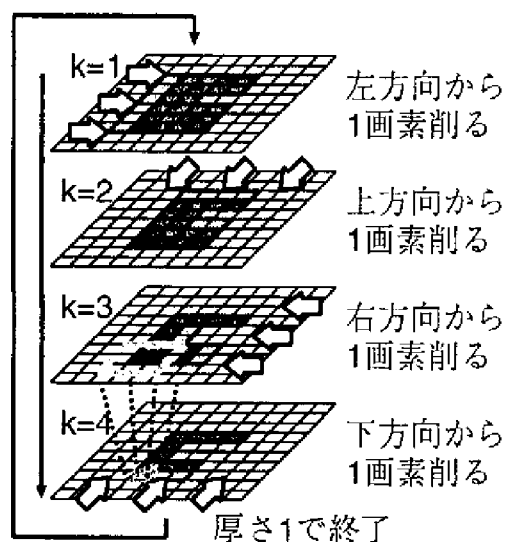


Fig. 5 細線化処理の手順

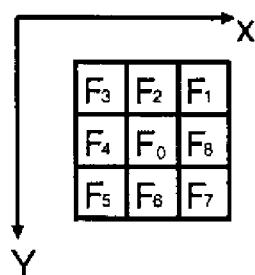


Fig. 6 3x3 ウィンドウ

2.5 接続情報に基づく人間領域候補の検証

図10に示すように、細線化の結果から接続情報を抽出し、人間のモデルとグラフマッチングを行うことで、人間かどうかの判断を行う。人間候補領域が、本当に人間領域である場合には、足領域、手領域、頭領域の大きさが、伸長に対して決められる。そこで、各領域に含まれる端点を、それぞれ、足、手、頭とみなす。この仮定に基づき、人間かどうかをチェックする。例えば、手、または、足が3本以上ある場合には人間ではないと判断する。さらに、手、足、頭の接続の仕方により、人間であれば、数種類のグラフでモデル化できる。この人間のモデルとマッチングすることで、より詳細にチェックを行える。また、手、足の伸長に対する太さなどを用いてさらに詳細にチェックを行える。

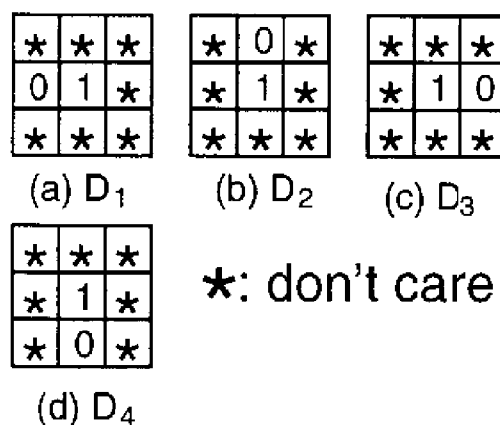
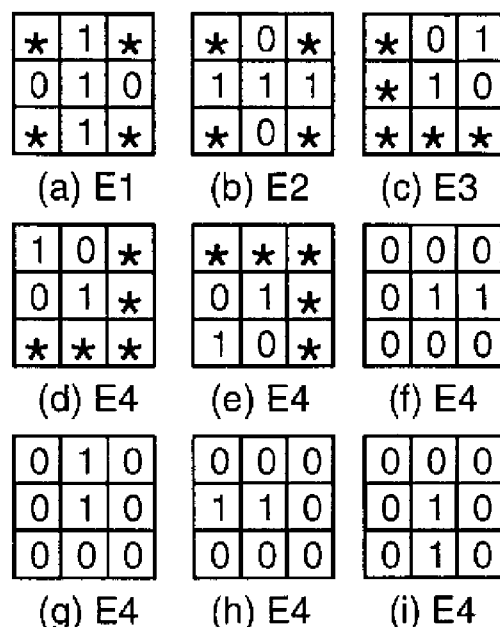


Fig. 7 処理方向を決定するためのテンプレート



*: don't care

Fig. 8 抽出点を決定するためのテンプレート

3. プロセッサアーキテクチャ

図11に人間抽出VLSIプロセッサの全体の構成を示す。領域分割部、人間領域候補生成部、細線化部、検証部から構成される。人間領域候補数が13000回の場合のシミュレーション例では、細線化が全体の処理時間の99.6%を占めている。従って、細線化の並列化が重要となる。図12に細線化部の構成を示す。細線化部は、人間領域候補のシルエット画像を記憶する画像メモリ、演算部、中間結果を記憶するバッファから構成される。今回採用している細線化アルゴリズムは、完全並列処理が可能であるため、画像メモリ、および、バッ

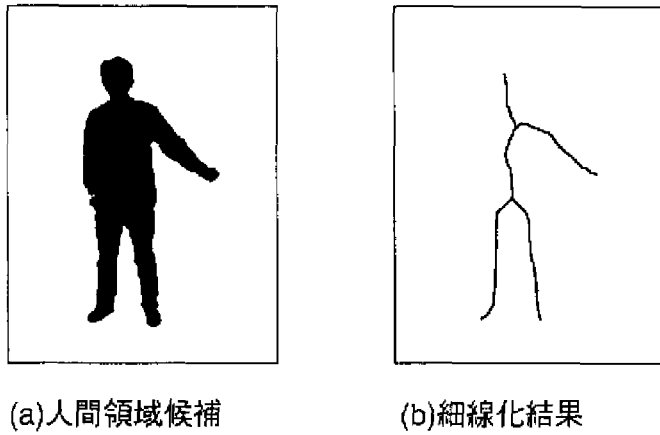


Fig. 9 人間候補領域に対する細線化結果例

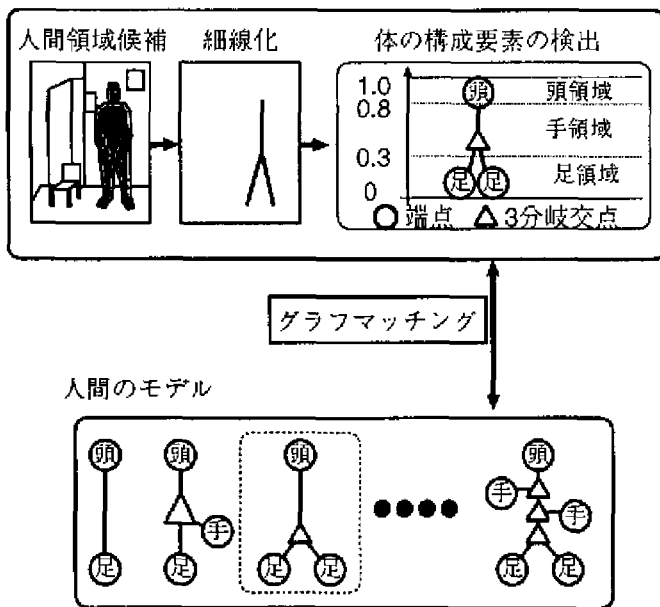


Fig. 10 接続情報に基づく人間領域候補の検証

ファからの入力データを、演算部に並列に供給できれば、演算器の稼働率を100%にできる。すなわち、演算器数に反比例して処理時間を減少できる。画像メモリの容量は固定であるため、与えられたチップ面積制約下で処理時間を最小化するためには、中間結果記憶用のバッファの容量を最小化し、演算器数を最大すればよい。

バッファ容量を最小化するために、共通入力に着目した瞬時出力生成スケジューリングを提案する。この考え方は以下のようなものである。

- 共通の入力データを有する演算をできるだけ同時に行なうようにする。
- 入力データがそろった演算をできるだけ速や

かに行なう。

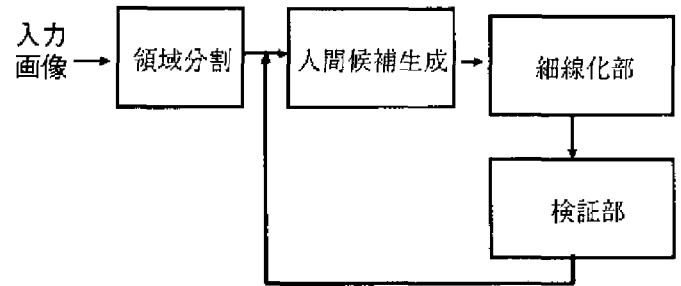


Fig. 11 人間抽出VLSIプロセッサの構成

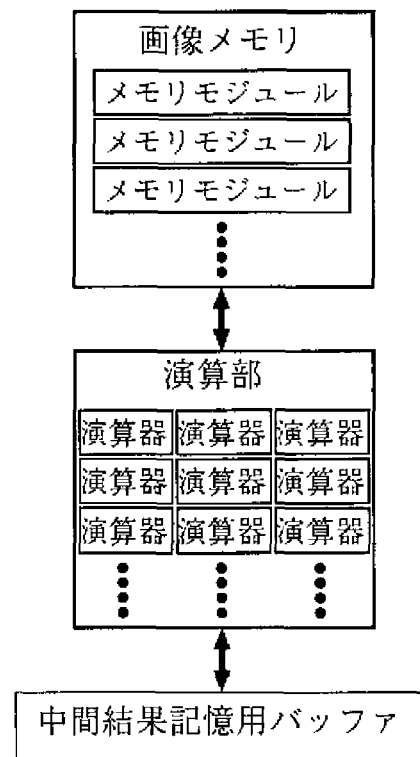


Fig. 12 細線化部の構成

簡単のために図13に示す1x3のウィンドウの例で考える。演算部の並列度に応じて、入力データを並列に供給できるとする³⁾。図14が共通入力に着目した瞬時出力生成に基づくスケジューリングである。共通の入力データ、例えば、D1, D2を用いる演算を同時に、かつ、できるだけ早い時点で行なうことにより、記憶を不要にしている。この例では、バッファは1個でよい。一方、図15に示すように、パイプラインを用いると、共通の入力が異なるステップで使われるために、記憶の必要が生じている。結果として、バッファは3個となる。また、図16に示すように、空間的並列処理のスケ

ジュールにおいても、共通入力に対する演算を同時に行なわなかった場合には、後のステップでその入力を使用するために、バッファに記憶する必要がある。結果として、バッファ数が3個となる。

図17に共通入力に着目した瞬時出力生成スケジューリングに基づく細線化部の構成例を示す。一例として、演算器数を10240個とした場合には、ワークステーションでの処理と比べ、20万倍の高速化が達成される。また、パイプラインアーキテクチャと比較すると、同一トランジスタ数で約10%の性能向上が達成される。

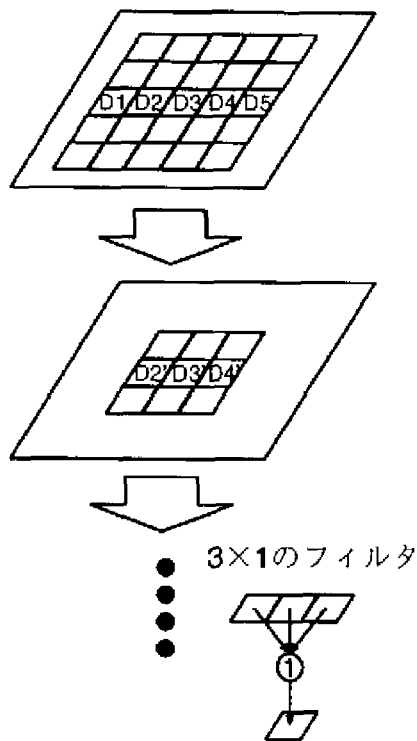


Fig. 13 1x3ウィンドウの例

3.1 むすび

画像から人間抽出するために、人間の体格の面体差に依存しない特徴量として、接続情報に着目したアルゴリズムを提案した。また、そのアルゴリズムにおいて計算量が膨大となる、細線化を高並列に行なうVLSIプロセッサを提案した。並列データ供給が可能な条件下では、瞬時出力生成スケジューリングを用いることで、従来主流であるパイプラインアーキテクチャに比べて、同一面積で処理時

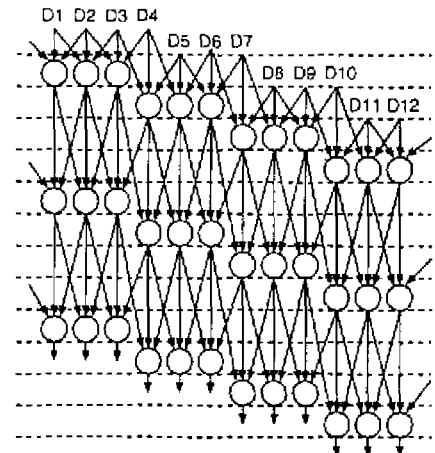


Fig. 14 共通入力に着目した瞬時出力生成スケジューリング

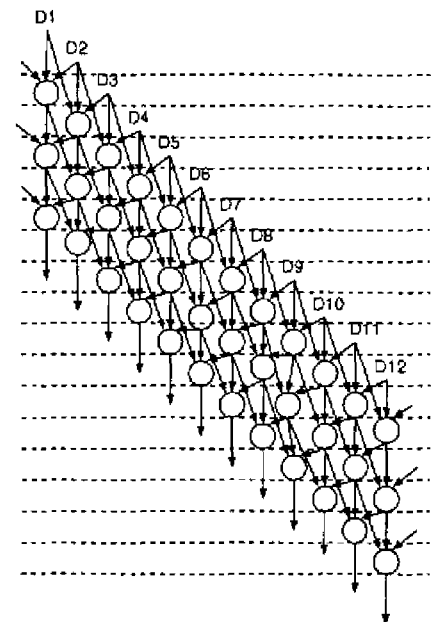


Fig. 15 パイプライン処理のためのスケジューリング

間を減少できることを示した。

今後、他の特徴量、例えば、赤外線などと組み合わせることで、より高精度な人間抽出が可能になるものと期待される。

参考文献

- 1) 境田 慎一, 鹿喰 善明, 田中 豊, 湯山 一郎: K平均アルゴリズムの初期値依存性を利用した統合処理による画像の領域分割法
- 2) 田村 秀行: 細線化法についての諸考察, 信学技報, PRL75-66(1975)
- 3) 張山 昌論, 李 昇桓, 亀山 充隆: 最適メモリアロケーションとモーションステレオVLSIプロセッサへの応用, 第2回システムLSI琵琶湖ワークショップ(1998)

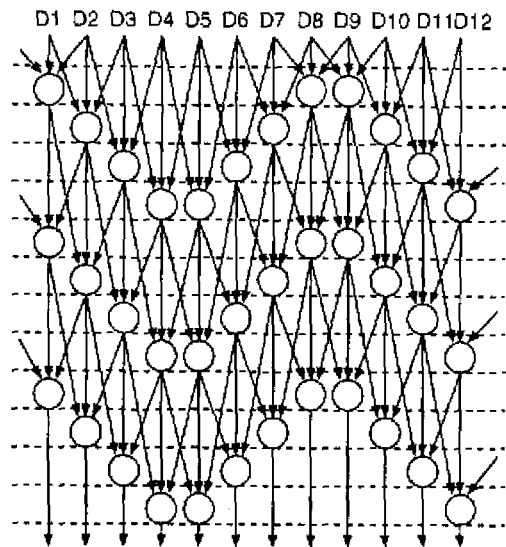


Fig. 16 共通入力に着目しない瞬時出力生成スケジューリング

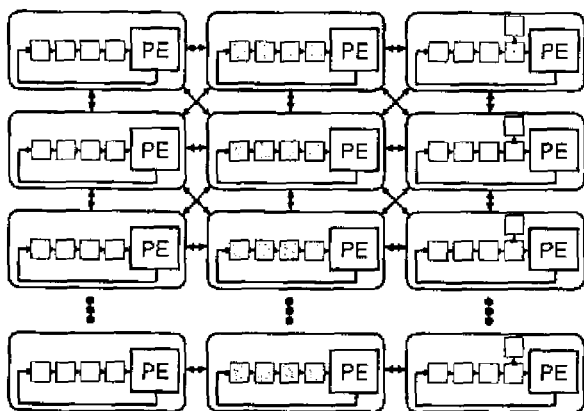


Fig. 17 共通入力に着目した瞬時出力生成スケジューリングに基づく細線化部の構成