

2次元デジタルフィルタの動的データ駆動型並列実現法

7D-3

水野 正之 犬束 敏信 Ali M. Alhaj 岩田 誠 寺田 浩詔

大阪大学 工学部

1 はじめに

2000 × 2000 画素を越えるような高精細な動画画像を実時間あるいはそれを越える速度で扱う複雑な処理を行う場合、その処理には、数十 GFLOPS から TFLOPS を越える処理能力が必要だといわれている。

また、信号処理は粒度の非常に細かい処理を要求するため、そのほとんどがシストリック・アレイあるいはPLA(Programmable Logic Array)によるものであり、そのため各機能機構ごとに専用のプロセッサを用意する必要があった。

本研究では、データ駆動型処理方式が、自然な形で複数の入力セットをパイプライン的に、かつ安全に処理できること、また、非常に粒度の細かい並列処理、およびそれにとまなう通信制御などのコンテキストの切り換えが頻繁に起こるような処理でも、パイプラインの乱れがないことなど [1][2] に着目し、このような細粒度の処理を動的データ駆動型マルチプロセッサで実現する研究の一環として、信号処理に必要不可欠な2次元デジタルフィルタについて、TFLOPSを越える処理要求にも対応できる実現法を提案する。

2 2次元デジタルフィルタ実現での課題

動的データ駆動型マルチプロセッサで2次元デジタルフィルタを実現する場合の主な課題として、(1) 動的データ駆動型プロセッサでの実現に固有の問題 (i),(ii),(iii)、(2) 2次元デジタルフィルタの並列実現に固有の問題 (iii) - (vii) がある。

世代識別用タグ資源の制限

2次元の入力面を持つため、各データの識別子(タグ)に対し工夫が必要となる。すなわち、(i) 動的データ駆動型処理方式では、データをタグを用いて識別するため、その識別用タグ資源の有限性に対する配慮、(ii) 2次元データを1次元の識別子で認識するための配慮、が必要となる。

外部記憶域(EDS)のアクセスの遅延

(iii) データが2次元入力面に対しラスタ・スキャン形式で入力される場合、一時記憶するデータが非常に多いため、そのための外部記憶域(EDS)を効率的かつ安全に利用するための配慮が要求される。

IIRフィルタ構造による出力帰還ループの存在

(iv) 帰還するデータの遅延が全体の動作速度に影響を与えない工夫と(v)それがプロセッサ・トポロジの変化に対して柔軟に対応できる必要がある。

サンプル周波数の増加に対する対応

(vi) サンプル周波数の増加が、そのままプロセッサの性能向上を要求するのではなく、要素プロセッサ数の増加によって対応し、要素プロセッサに要求される処理能力を極力小さくする必要がある。

プロセッサ間通信の影響

(vii) 特殊なプロセッサ・トポロジでなく、一般的なプロセッサ・トポロジでもプロセッサ間通信がボトルネックとならない必要がある。

3 2次元デジタルフィルタの実現

本研究では、先の課題で示した動的データ駆動型プロセッサ特有の問題を2次元デジタルフィルタの実現法を変形することで隠蔽し、データ駆動型プロセッサの特徴が有効に働くような方法(3.1,3.2,3.3節)を提案する。

3.1 2次元デジタルフィルタの低次元化

多次元のデジタルフィルタを、低次元のフィルタの多段分離型に分解する方法として、S.Treitel, J.L.ShanksによるSV分解、A.N.VenetsanopoulosらによるJD分解およびLU分解 [3] などがある。ここでは、(i),(ii),(iii)の問題を解決するため、次のように2次元 FIR フィルタを分解する。

いま、FIR フィルタ

$$H(z_1, z_2) = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} h(i, j) z_1^{-i} z_2^{-j} \quad (3.1)$$

について、 $h(i, j)$ から作られる $m \times n$ の単位インパルス応答行列を $H = [h(i, j)]$ とする。この行列を列方向に行ベクトルの一時独立性から縮退し、

$$H(z_1, z_2) = \sum_{j=0}^{n-1} \left\{ \sum_{i=0}^{m-1} \tilde{h}(i, \tilde{p}(j)) z_1^{-i} \right\} c(j) z_2^{-j} \quad (3.2)$$

の多段分離型並列構造(図1)にする。ここで、 $p = \text{rank}(H)$ で $\tilde{p}(\cdot)$ は縮退後の行位置の、縮退前の行位置への関数である。

行方向の演算を行ってから、その出力値に対して列方向の演算を行なうため、(ii)が解決できる。さらに、列方向の演算がFIFOを含む単純な演算のため、外部記憶域の利用による効率的な一時記憶(iii)が実現できる。

3.2 IIRフィルタ部の再帰ループ内の遅延

図2(a)に示すように2次元IIRフィルタの出力値を同時並列に演算できるデータ群は、通常2次元平面の対角線上のデータである。ここで、2次元IIRフィルタ、

$$H(z_1, z_2) = \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} a(i, j) z_1^{-i} z_2^{-j}}{1 - \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} b(i, j) z_1^{-i} z_2^{-j}} \quad (3.3)$$

A Parallel Dynamic-Data-Driven Implementation of 2-D Digital Filters

Masayuki MIZUNO, Toshinobu INUZUKA, Ali Mohammad Alhaj, Makoto IWATA, Hiroaki TERADA

Faculty of Engineering, OSAKA University

に対して、列方向に参照される出力値を M 段シフトするため、式 (3.3) を再帰的に使用し、

$$H(z_1, z_2) = \frac{\sum_{i=0}^{m-1} \sum_{j=0}^{n+M-1} \tilde{a}(i, j) z_1^{-i} z_2^{-j}}{1 - \sum_{i=0}^{m-1} \sum_{j=M}^{n+M-1} \tilde{b}(i, j) z_1^{-i} z_2^{-j}} \quad (3.4)$$

を得る。M = 1 の例を図 2(b) に示す。この場合、行方向のデータ群が同時並列に演算できるため、行方向に複数の要素プロセッサを用いれば、(i) に対応できる。また、このように、出力帰還ループ内に遅延を持たせれば [4]、出力値の伝搬遅延が全体の速度を規律しない (iv), (v)。

3.3 FIR フィルタ部のインターリーブ化

FIR フィルタを実現する場合、帰還ループ構造を含まないため、パイプライン処理により並列性が得られる。しかし、サンプリング周波数の増加によって、パイプラインステージ間に許容される遅延時間の縮小化にともなうデータ転送の安全性が問題になる (vi)。ここでは、この問題に対処するため、以下の手法を取る。すなわち、FIR フィルタ

$$H(z_1, z_2) = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} h(i, j) z_1^{-i} z_2^{-j} \quad (3.5)$$

を、次の 4 つのポリフェーズ・フィルタに分解して、図 3 に示すように接続することで所望の FIR フィルタを構成する。

$$\begin{aligned} 1) & \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} h(2i, 2j) z_1^{-i} z_2^{-j} & 2) & \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} h(2i, 2j+1) z_1^{-i} z_2^{-j} \\ 3) & \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} h(2i+1, 2j) z_1^{-i} z_2^{-j} & 4) & \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} h(2i+1, 2j+1) z_1^{-i} z_2^{-j} \end{aligned}$$

ここで、 $z_1^{-i} \triangleq z_1^{-2i}$ 、 $z_2^{-j} \triangleq z_2^{-2j}$ である。この場合、(1) 各ポリフェーズ・フィルタをインターリーブ化すれば、各フィルタに要求されるサンプル周波数は、全体のサンプル周波数に比べて 1/4 に低減化される。また、(2) 入出力を除いてポリフェーズ・フィルタ間でのデータ依存性がない。さらに、(3) 各ポリフェーズ・フィルタを図 1 の構造にすれば、その FIFO が有効である限り入出力データの伝搬が全体の速度を規律しない。また、(4) 図 3 はメッシュ型のプロセッサ・トポロジに容易にマッピングできる (vii)。

プロセッサ資源に余裕がある場合は、 $z_1^{-i} \triangleq z_1^{-ki}$ 、 $z_2^{-j} \triangleq z_2^{-kj}$ とすれば、さらに高速のサンプル周波数にも対応できる。

4 おわりに

本研究では、動的データ駆動型プロセッサで 2 次元フィルタを実現する場合の問題点を明らかにし、大規模な画像データの長時間処理に十分耐え得るような実現法の検討を行った。なお、シミュレーションによる本実現法の評価については、当日発表する予定である。

なお、本研究で用いた動的データ駆動型プロセッサ [1], [2] は、シャープ (株) ならびに三菱電機 (株) から提供を受けた。両者の関係各位に厚く御礼申し上げる。

参考文献

[1] 金倉 広志, 宮田 宗一: “動的データ駆動型プロセッサによる並列処理方式の検討”, マイクロコンピュータ・アーキテクチャ・シンポジウム, 平成 3 年 11 月。

[2] Komori, S., Tamura, T., Asai, F., Shima, K., Nishikawa, H., Terada, H.: “A 50MFLOPS Superpipelined Data-Driven Microprocessor”, Proc. ISSCC, pp.284-286 (1991).

[3] Nikias, C.L., Chrysafis, A.P., Venetsanopoulos, A.N.: “The LU Decomposition Theorem and Its Implications to the Realization of Two-Dimensional Digital Filters”, IEEE Trans. on ASSP, 33, 3, pp.694-711 (June 1985).

[4] Parhi, K.K., Messerschmitt, D.G.: “Pipeline Interleaving and Parallelism in Recursive Digital Filters — Part I: Pipelining Using Scattered Look-Ahead and Decomposition”, IEEE Trans. on ASSP, 37, 7, pp.1099-1117 (July 1989).

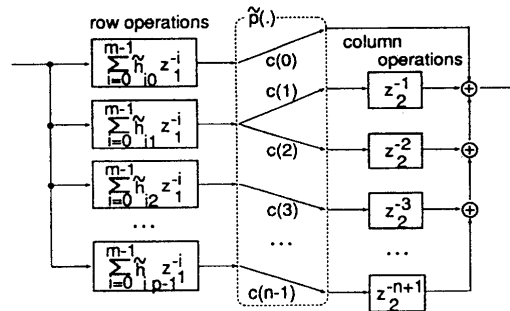


図 1: 2次元デジタルフィルタの低次元化

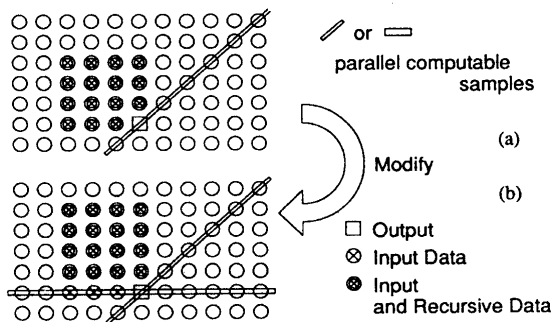


図 2: IIR フィルタ部の再帰ループ内の遅延

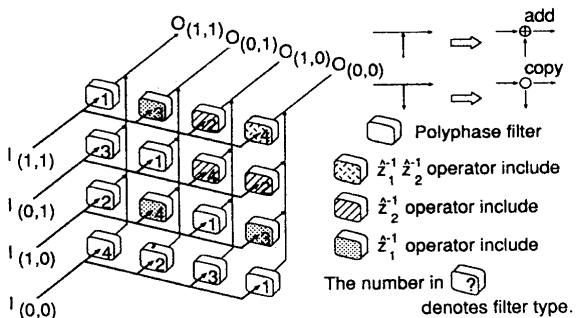


図 3: FIR フィルタ部のインターリーブ化

(例えば、 $I_{(1,0)}$ は、入力面 $I(x, y)$ に対して、 x および y の 2 に対する剰余が、それぞれ 1 および 0 のサンプルを示す。)