

ステレオ動画像を用いた移動物体抽出とその応用

Moving Object Extraction Using Motion Stereo Images and Its Application

○渡邊郷史, 張山昌論, 亀山充隆

○Satoshi Watanabe, Masanori Hariyama, Michitaka Kameyama

東北大学

Tohoku University

キーワード: 絶対差分和 (sum of absolute differences), オプティカルフロー (optical flow), ステレオビジョン (stereo vision), ロジックインメモリアーキテクチャ (logic-in-memory architecture), 最適アロケーション (optimal allocation)

連絡先: 〒980-8579 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
渡邊郷史, Tel.: (022)217-7155, Fax.: (022)263-9167, E-mail: watanabe@kameyama.ecei.tohoku.ac.jp

1. まえがき

3次元物体の位置・速度の推定は, 高安全自動車, ホームサービスロボット, ホビエロボットなどの実現のための重要な要素技術である。例えば, 高安全自動車衝突警報システムにおいては, 対向車等の位置・移動速度に基づき将来の軌道を予測することにより, 衝突可能性を判定することが可能となり, 危険な場合にはドライバーに警報を与えることができる。

3次元物体の位置・速度の推定を行う方法として, ステレオビジョンとオプティカルフローを用いてステレオ動画像を取得する方法がよく知られている。信頼性が高いステレオ動画像を取得するためには, ステレオビジョン, オプティカルフロー抽出において, 信頼性の高い画像間の対応を求めることが重要となる。本稿では, 画像間の対応を求める方法として, SAD(sum of absolute differences)を用い

た方法を用いる。SADを用いた対応点探索においては, SADのウィンドウサイズは, 曖昧性を除去できる程度に大きく, 透視投影の歪みを除去できる程度に大きいことが望まれる。本稿では, SADグラフの最小値のユニーク性に基づき, 画素毎に適切なウィンドウサイズを決定するような, ウィンドウサイズ可変マッチングアルゴリズムを提案する。

SADを用いた対応点探索においては, 候補ウィンドウ数が膨大となるために, 汎用プロセッサを用いると計算時間が膨大となる問題がある。例えば, 640×480 ドットの2枚の画像間の対応を求める処理を汎用プロセッサ (Pentium III 500MHz) を用いて行うと30分以上の処理時間を要する。そのため, 専用VLSIプロセッサの開発が重要となる。

SADに基づく対応点探索においては, 同じウィンドウ内の各ピクセルに対してAD(Absolute difference)を計算できるというウィンドウレベルの並列

性と、異なる候補ウィンドウに対して並列にSADを計算できるというウィンドウレベルの並列性が存在する。ウィンドウサイズ可変マッチングアルゴリズムでは、ピクセルレベルの並列性がウィンドウサイズに応じて変化するため、1個のウィンドウ内のピクセルレベルの並列性を活用したアーキテクチャでは、演算器の稼働率が低くなる。この問題を解決するために、ウィンドウ並列ピクセル直列アーキテクチャを提案する。本アーキテクチャでは、1個のウィンドウ内のピクセルに対しては直列的にAD演算を行うことにより、ウィンドウサイズの変化を処理ステップ数の変化で対応でき、演算器の稼働率を100%にできる。一方、候補ウィンドウ数はウィンドウサイズの変化によらず一定であることに着目し、ウィンドウレベルの並列性を最大限活用する。このように、固定の並列度だけを活用することにより、演算器の稼働率を100%に保つたまま、メモリ・演算器間の相互結合網の単純化が可能となる。

ステレオビジョンの対応点探索部をGate Arrayの一種である、Module Based Array (Light Speed社、50万ゲート)(MBA)を6個用いて画像サイズ640×480、ウィンドウサイズ3×3～25×25まで演算できるプロセッサを試作したところ、汎用プロセッサの数千倍の高速化が達成された。

また、0.5μm CMOS設計ルールを用いて、画像サイズ160×160、ウィンドウサイズ3×3～25×25まで演算できるオプティカルフロー抽出VLSIプロセッサを設計したところ、汎用プロセッサの数万倍の高速化が達成された。

2. アルゴリズム

2.1 高信頼ステレオビジョンアルゴリズム

ステレオビジョンでは、左右の画像の対応点が求められれば、3角測量の原理に基づいて3次元座標を計算することができるため、対応点を正確に求めることが重要となる。ウィンドウサイズを W とす

ると、対応点は次式の差分絶対和(SAD)を計算することにより求められる。

$$F_W = \sum_{j=-\frac{W-1}{2}}^{\frac{W-1}{2}} \sum_{i=-\frac{W-1}{2}}^{\frac{W-1}{2}} |I_L(X_L+i, Y_L+j) - I_R(X_R+i, Y_R+j)|$$

ここで、 (X_L, Y_L) 、 (X_R, Y_R) はそれぞれ左右画像中のウィンドウの中心座標、 I_L 、 I_R は左右の画像上の画素値である。Fig. 1に示すように、任意の点 $P(X_L, Y_L)$ とエッジローライン上の全ての点 $Q(X_R, Y_R)$ に対してSADを計算する。SADが小さいほどウィンドウ内の画像が類似していることを示す。

SADが最小となる点を対応点とする方法を用いると、ウィンドウサイズが小さい場合には、SADが類似した領域が複数存在すること起因する誤対応の可能性が高くなる。一方、ウィンドウサイズが大きい場合には、中心画素以外の多くの点の情報が含まれるため対応の信頼性が低下する。

この問題を解決するために、類似した点が存在しないできるだけ小さいウィンドウサイズで対応点を決定するアルゴリズムを提案する。ウィンドウサイズが W の場合のSADグラフの最小値のユニーク性 S_W を次式により与える。

$$S_W = (f_1 - f_2)/W^2$$

ここで、 f_1, f_2 は、Fig. 2に示すように、それぞれSADグラフ F_W における最小値、2番目の極小値である。また、 $0 \leq F_W \leq I_{max}$ (I_{max} は最大の画素値)であることに着目し、ウィンドウサイズが異なる場合のユニーク性を比較できるように W^2 で正規化

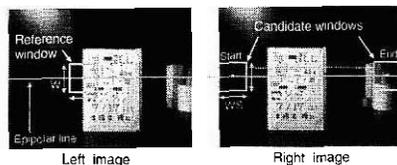


Fig. 1 Search for a corresponding point

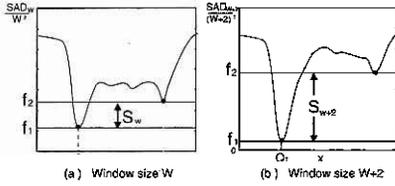


Fig. 2 SAD graphs for Window sizes W and $W+1$.

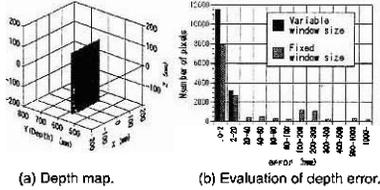


Fig. 3 Simulation result.

している。はじめに, Fig. 2(a)に示すように, 小さいウィンドウサイズ W で SAD を計算し, SAD が最小となる点を Q_1 を仮の対応点とし, ユニーク性を S_w とする. 次に, Fig. 2(b)に示すように, 大きいウィンドウサイズ $W+2$ で SAD を計算し, Q_1 が真の対応点であるかどうかを検証する. SAD グラフが Q_1 で最小値となり, かつ, ウィンドウサイズが W の場合と比較してユニーク性が大きくなる場合 ($S_w \leq S_{w+2}$) にだけ Q_1 を対応点とする. 最大のウィンドウサイズまで処理を行っても対応点が決定されない場合には対応点なしとする. Fig.3に, Fig.1の画像に対するシミュレーション結果を示す. Fig. 3(b)より, ウィンドウサイズ可変アルゴリズムを用いた方が, ウィンドウサイズ固定の場合に比べて, 誤対応点の数を減少できることがわかる. 本アルゴリズムの特徴は, 実験的に求めた閾値を必要とせず, 自動的に適切なウィンドウサイズを決定できることである¹⁾.

2.2 高信頼オプティカルフロー抽出アルゴリズム

オプティカルフローは, Fig. 4に示すように異なる時刻に取得された2枚の画像間の対応を求める

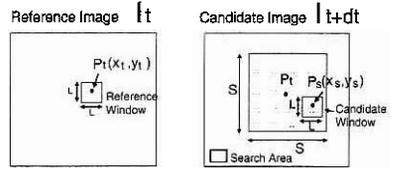


Fig. 4 Optical flow extraction.

ことにより抽出される. ここで, 画像 I_t, I_{t+dt} はそれぞれ時刻 t と $t+dt$ におけるカメラからの取得画像である. まず, 参照画像 I_t 上に点 $P_1(x_1, y_1)$ を中心とする大きさ $L \times L$ のウィンドウ(参照ウィンドウ)を設定する. 次に, 候補画像 I_{t+dt} 上に, 点 P_2 を中心とする大きさ $S \times S$ の探索領域を設定する. 候補画像上に点 P_2 を中心とする大きさ $L \times L$ のウィンドウ(候補ウィンドウ)を設定して, P_2 が探索領域に含まれる全てのウィンドウに対して次式に示す差の絶対値差分の和(SAD, Sum of Absolute Differences)を計算することにより求められる.

$$SAD(P_1, P_2) = \sum_{i=-\frac{L-1}{2}}^{\frac{L-1}{2}} \sum_{j=-\frac{S-1}{2}}^{\frac{S-1}{2}} |P(t+dt, x_1+i, y_1+j) - P(t, x_1+i, Y_1+j)|$$

ここで $F(t, a, b)$ は時刻 t の画像の座標 a, b の画素値を表す関数である. SAD が小さいほど2つのウィンドウが類似していると考えられる. SAD に基づく対応点探索の高信頼化を達成するために, ステレオビジョンと同様に, ウィンドウサイズ可変アルゴリズムを用いる.

さらなる高信頼化を達成するために, 画像対の時間間隔縮小に基づくオプティカルフローの高信頼化を提案する. Fig. 5に示すように, 時刻 t から $t+dt$ の対応点を求めるのに2枚(1対)の画像を用いるのに比べて, N 対の画像を用いると, 対となる画像間の時間間隔は $1/N$ となる. フレームの時間間隔が短くなるほど, 画像の変化は少なくなるため, 対応点探索の信頼性は向上する. このようして求めた部分的オプティカルフローを連鎖させることにより, 高信頼なオプティカルフローを求める.

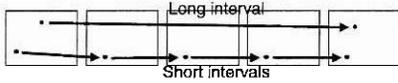


Fig. 5 Optical flow extraction using densely sampled images.

2.3 ステレオ動画生成

Fig.6に示すように、ステレオビジョンと最適化フローを利用して、2つの時刻で2台のカメラからの入力画像、計四枚の画像間での対応が求められ以下のようにステレオ動画を生成できる。

- 1) 時刻 t における2つのカメラの画像を用いたステレオビジョンに時刻 t 各点の三次元座標を求める。
- 2) 時刻 $t + dt$ におけるカメラの画像を用いたステレオビジョンに時刻 t 各点の三次元座標を求める。
- 3) カメラ1における時刻 t 及び $t + dt$ の画像間の対応を求めることにより、異なる時刻の3次元座標間の対応(動きベクトル)を求めることができる。

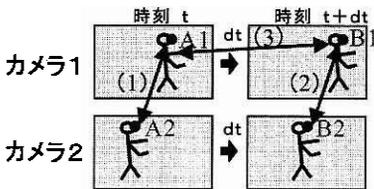
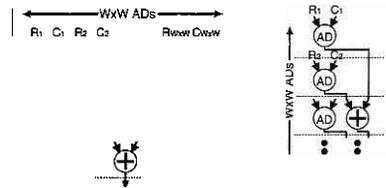


Fig. 6 Obtaining motion stereo image.

3. プロセッサアーキテクチャ

3.1 ステレオビジョンVLSIプロセッサ

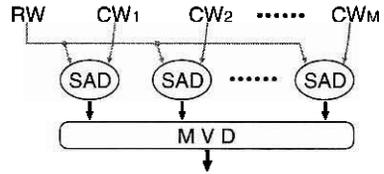
対応点探索においては、Fig. 7(a) に示すように1個のウィンドウ内の複数のピクセルに対して



Ri: Pixel value of a reference window
Ci: Pixel values of a candidate window
W: Window size

(a) Scheduled DFG of pixel-parallel SAD computation.
(b) Scheduled DFG of pixel-serial SAD computation.

Fig. 7 Data-flow graph of SAD computation.



RW: Reference Window CW: Candidate Window
MVD: Minimum Value Detection

Fig. 8 Data-flow graph of correspondence matching.

AD(Absolute Difference)を並列に計算できるといふ並列性と、Fig. 8 に示すように、複数の候補ウィンドウに対して並列にSADを計算できるという並列性がある。ピクセル間の並列性だけを活用して演算部を構成すると、並列度はウィンドウサイズ W に依存して変化するため、演算器の稼働率は低くなるという問題が生じる。一方、ウィンドウ間の並列性だけを活用する演算部では、並列度が入力画像の1行当たりのピクセル数で決まるために固定であるので、あらかじめ稼働率が100%になるような数の演算器を用意することができる。そこで、複数の候補ウィンドウに対する並列性を活用し、ウィンドウ内のピクセルに対しては、Fig. 7(b) に示すように直列的にADを計算することにより、演算器の稼働率を100%に保ったまま並列処理を可能にする。

Fig.9に、ステレオビジョンVLSIプロセッサの構成を示す。入力画像に対して、LOGフィルタを適

用することにより、ノイズ除去及び輪郭強調を行い、対応点探索の信頼性を向上させる。その出力画像は、画像メモリに格納される。画像メモリは大容量であるためアクセス時間が大きく、処理のボトルネックとなる。この問題を解決するために、頻繁に使用する画素データを高速なバッファに記憶する階層的メモリシステムを用いる。対応点探索はステレオビジョンにおいて最も計算量が多い処理であるため、対応点探索部を専用VLSI化する。対応点探索結果は、ホストプロセッサへ転送され、3次元座標が計算される。

Fig.10に示すように、対応点探索部においては、 n 個の処理要素(PE)を用いて n 個の候補ウィンドウに対して並列処理を行う。また、各PEにおいてはADの計算と累積加算のパイプライン処理を行う。MLは左画像を記憶するためのメモリモジュールである。右画像は並列アクセスのために n 個の領域に分割されメモリモジュール $MR_i (i = 1, 2, \dots, n)$ に記憶されるとする。各PEでは対応する画像の列に対してだけ演算を行うため、メモリ・PE間の相互結合網の複雑さは最小化される。隣接PE間で中間結果を転送するようなアロケーション及びブスケジューリングにより、相互結合網の複雑にすることなく、SAD演算を行うことが可能となっている。

対応点探索部に試作にはGate Arrayの一種である、Module-Based Array(Light Speed 社,50万ゲート)(MBA)を用いた。1個のMBAチップには、107個のPE及びバッファを集積でき、動作スピードは90MHzとなった。また、LOGフィルタ及び、最小値検出部、制御部はFPGA(Xilinx社Vertex1000)を用いた。Table 1にMBAチップを用いたステレオビジョンプロセッサの諸元を示す。従来の汎用プロセッサと比較して、400倍以上の高速化が達成されることが確認された。

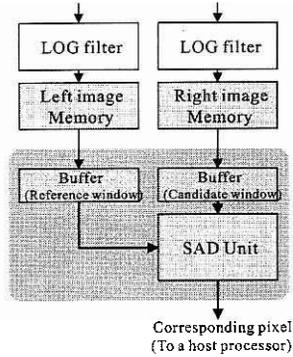


Fig. 9 Block diagram of the stereo vision processor.

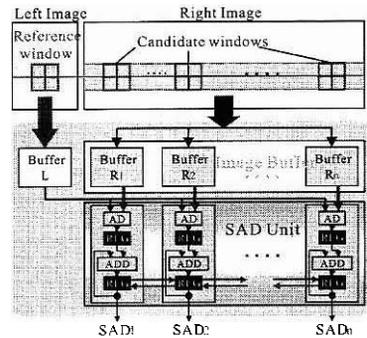


Fig. 10 Block diagram of the SAD unit.

3.2 オプティカルフローVLSIプロセッサ

ステレオビジョンVLSIプロセッサと同様に、同一ウィンドウ内のピクセルに対しては直列的にSAD演算を実行し、ウィンドウレベルの並列性を活用することにより、演算器の稼働率を100%に保ったまま並列処理を可能とする。

Fig.12は提案するウィンドウ間の並列性のみを最大限に利用したアーキテクチャ(ピクセル直列ウ

Image size	640 × 480 [pixel]
Processing time	1.2sec/ depth map
Window sizes	Min:3 × 3, Max: 25 × 25
Number of modules	6MBAs, 2FPGAs
Number of PEs	640
Clock frequency	MBA: 90MHz, FPGA: 45MHz

Table 1 Feature of the stereo vision processor.

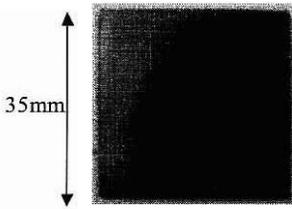


Fig. 11 Stereo vision VLSI processor

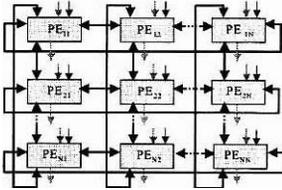


Fig. 12 Block diagram of the optical flow extraction VLSI processor.

インドウ並列アーキテクチャ)のブロック図である。探索範囲内の $S \times S$ 個のウインドウに対して並列にSADを計算するために、SAD演算部は $S \times S$ 個のPEから構成する。PEの基本的な構成は、ステレオビジョンVLSIプロセッサと同様であり、メモリモジュールと絶対値差分演算器と加算器をそれぞれ1個内蔵した構成となっている。

設計ルール	0.5 μ mCMOS二層メタル
トランジスタ数	5640 \times 400個
動作周波数	100MHz
画像サイズ	160 \times 160
探索範囲	20 \times 20
ウインドウサイズ	3 \sim 25

Table 2 Feature of the optical flow extraction VLSI processor.

また、ウインドウ間の並列性を最大限に利用するは、画像の $S \times S$ 個の画素を並列にアクセスする必要がある。この問題を解決するためにFig.13に示す長方形アロケーション^{2),3)}に従って、候補画像を各メモリモジュールに記憶する。さらに、隣接した

1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6
7	8	9	7	8	9	7	8	9
1	2	3	1	2	3	1	2	3
4	5	6	4	5	6	4	5	6
7	8	9	7	8	9	7	8	9

Fig. 13 Rectangular memory allocation.

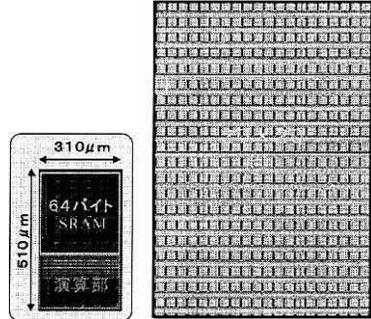


Fig. 14 Chip layout of the optical flow extraction VLSI processor.

ウインドウを隣接したPEで処理し、各PEでは1個のメモリモジュールの画素だけを処理するような演算器アロケーションによりメモリとPE間の通信を局所化できる。

Fig. 14は上に述べたオプティカルフロー抽出VLSIプロセッサのレイアウトである。また、Table2にその諸元を示す。2枚の画像間のオプティカルフローを30msで抽出することができ、汎用プロセッサ(Pentium III 500MHz)での処理時間と比較して10000倍の高速化を達成できる。

参考文献

- 1) M.Hariyama, T. Takeuchi, M. Kameyama: Reliable Stereo Matching for Highly-Safe Intelligent Vehicles and Its VLSI Implementation, IEEE Intelligent Vehicles Symposium, pp. 128-133(2000).

- 2) Lee Seung, Masanori Hariyama, Michitaka Kameyama, "A Three-Dimensional Instrumentation VLSI Processor Based on a Concurrent Memory-Access Scheme," IEICE Trans. Electron, vol.E80-C, No.11, pp.1491-1498,1997.
- 3) 張山昌論, 李昇桓, 龜山充隆, "転送ボトルネックのないセンサ・メモリアーキテクチャに基づくモーシヨンステレオVLSIプロセッサの構成," 電気学会論文誌, vol.120-E, no.5, pp.237-244,2000.