

平成 22 年 5 月 14 日現在

研究種目：基盤研究（C）

研究期間：2007～2009

課題番号：19560339

研究課題名（和文） 実時間モデルベース動領域抽出 VLSI プロセッサの研究

研究課題名（英文） Study on real-time model-based video segmentation VLSI processor

研究代表者

深山 正幸 (MIYAMA MASAYUKI)

金沢大学・電子情報学系・講師

研究者番号：30324106

研究成果の概要（和文）：本研究はアフィン動きモデル推定による VGA 30 fps の動領域抽出用 VLSI アーキテクチャを提案する。VLSI 向けアルゴリズムは画像分割法、境界限定 ICM ラベリング、新領域の動き推定省略により特徴付けられ、精度の劣化無く演算量を 82%、メモリ容量を 95%、データ転送量を 99%削減できる。VLSI アーキテクチャは分割画像のパイプライン処理、複数領域の同時動きモデル推定、境界変更と新領域検出のラベリング処理共通回路により特徴付けられ、VGA 30 fps の動領域抽出を 167 MHz で実現でき、0.18 $\mu$ m プロセスのコア面積は 30 mm<sup>2</sup>である。

研究成果の概要（英文）：This study proposes a VLSI architecture for VGA 30 fps video segmentation with affine motion model estimation. The algorithm optimization for VLSI implementation is characterized by image division method, ICM labeling limited to region boundary, and omission of motion models estimation for new regions. The optimization reduces the computational costs by 82 %, the amount of memory by 95 %, and the amount of data traffic by 99 % without accuracy degradation. The VLSI architecture is characterized by pipeline processing of the divided images, concurrent motion models estimation for multiple regions, and a common processing element of update and detection labeling. The architecture enables VGA 30 fps video segmentation with 167 MHz frequency. The estimated core area using 0.18 $\mu$ m technology is 30 mm<sup>2</sup>.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007 年度	1,100,000	330,000	1,430,000
2008 年度	400,000	120,000	520,000
2009 年度	2,000,000	600,000	2,600,000
年度			
年度			
総計	3,500,000	1,050,000	4,550,000

研究分野：VLSI 画像処理

科研費の分科・細目：電気電子工学・電子デバイス・電子機器

キーワード：アフィン動き推定、動領域抽出、画像認識、実時間処理、VLSI、FPGA

## 1. 研究開始当初の背景

近年、ロボット、車載、監視といった今後の産業的な発展が予想される重要な応用分野において動画認識の高度化が望まれている。動画認識の例として以下の処理があげられる。

ロボット：視線認識、個人認識、顔表情認識、動作認識、立体物体認識、距離認識

車載：歩行者検出、車両検出

監視：流通解析、混雑解析、侵入者検出

これらの処理において動領域抽出は重要な前処理の一つである。動領域抽出とは輝度や色などの情報だけでなく動画における動き情報を使用して画像を領域に分割する処理である。

動領域抽出の一つの方法として、MRF(Markov Random Field)モデルを用いる方法がある。この方法では領域内各点の動きの正確性と、空間的な連続性(隣接する点は同じ領域に含まれる確率が高い)と、時間的な連続性(前フレームで推定された動きに従って移動した位置に同一点が存在する確率が高い)を表すエネルギー関数を定義し、エネルギーを最小化するように領域分割が行われる。通常、領域の動き推定と境界推定は互いに強く影響し合うため、収束するまでこれらの処理が何回も繰り返し行われる。収束を早めるためロバストな多解像度のアフィンパラメータ推定を導入し、動き推定と境界推定を分離し、1ステップで領域分割を行う方法が知られている。

この方法は計算量が比較的少ないという特徴があるが、VGA(640x480画素) 30fps(frame per sec.)のような解像度の比較的高い動画の実時間処理は莫大な演算量により現在のPCでは不可能である。一方、現時点で実現している動領域抽出の実時間処理は演算量を抑えるため用途を限定したアドホックなアルゴリズムに基づいている。

## 2. 研究の目的

本研究の具体的な目標は、平成21年度末までに、VGA 30fps解像度の動画に対して実時間処理可能な動領域抽出VLSIプロセッサを0.18 $\mu$ mプロセスの5mm角チップに実装することである。本研究は以下のことを明らかにする。

- ・アフィンパラメータによる動き情報を実時間計算するプロセッサ(アフィン動き推定プロセッサ)のVLSIアーキテクチャ(画像メモリ構成、メモリデータマッピング、データパス構成、制御方法等)と、VLSI設計結果(チップ面積、動作周波数、消費電力等)
- ・MRFモデルを用いて動領域抽出を実時

間処理するプロセッサ(モデルベース動領域抽出プロセッサ)のVLSIアーキテクチャと、VLSI設計結果

## 3. 研究の方法

最初にアフィン動きモデル推定プロセッサのVLSIアーキテクチャを研究する。具体的な研究内容は以下のとおりである。

(1)シミュレータの開発とシミュレーション、演算量プロファイリング

(2)ハードウェア化に向けたアルゴリズムの最適化

(3)VLSIアーキテクチャの研究、開発

(4)FPGAボード実装

(1)ではアフィン動き推定を行うソフトウェア・シミュレータをC言語により開発する。このシミュレータにテスト用動画を入力し、適切なアフィン動き推定が得られることを確認する。演算量のプロファイリングを行い、関数ごとに命令種類とその回数を調査する。(2)ではハードウェア化に向けたアルゴリズムの最適化を行う。具体的には浮動小数点から固定小数点への変換、ビット精度の最適化、メモリ容量を削減するための画像データキャッシュ法の検討等を行う。(3)ではVLSIアーキテクチャの研究、開発を行う。(4)では(3)で考案した回路をFPGAボードに実装し、アフィン動き推定プロセッサの機能、動作を検証する。

次にモデルベース動領域抽出プロセッサのVLSIアーキテクチャについて研究する。具体的な研究内容は以下のとおりである。

(1)シミュレータの開発とシミュレーション、演算量プロファイリング

(2)ハードウェア化に向けたアルゴリズムの最適化

(3)VLSIアーキテクチャの研究、開発

(4)FPGAボード実装

(1)では動領域抽出を行うソフトウェア・シミュレータをC言語により開発する。このシミュレータにテスト用動画を入力し、動領域抽出結果を確認する。(2)では境界推定と領域抽出の各ラベリング処理における収束までの繰り返しを局所的に行う方法について調査、検討する。(3)では境界推定と領域抽出を行うVLSIアーキテクチャについて研究する。境界推定と領域抽出はどちらもエネルギー関数の最小化を行うため回路を共通化する。(4)では前年度に導入したFPGAボードに(3)で検討した回路を実装し、動画をを入力して機能及び性能を検証する。

次に検討したアーキテクチャに基づく動領域抽出VLSIプロセッサを0.18 $\mu$ mプロセスで試作する。目標チップ面積は5mm角である。LSIテストを用いて試作チップの機能・性能評価と消費電力を実測する。

#### 4. 研究成果

##### (1)アフィン動きモデル推定プロセッサ

従来のアルゴリズムは推定精度を向上させるため、同じ階層において動きモデル推定を複数繰り返す増分推定( $\theta_{itr}$ )、複数解像度を用いて推定を行う多重解像度推定(level\_itr)、動きモデル推定時の外れ値を除外する重み $w$ による反復( $w_{itr}$ )を行っている。図1に従来および提案アルゴリズムのフローチャートを示す。従来では多重解像度推定の後に重み計算を1回行う。提案するアルゴリズムでは、各階層における増分推定毎に重み計算を行う。そして最新の重みを用いて次の増分推定を行う。

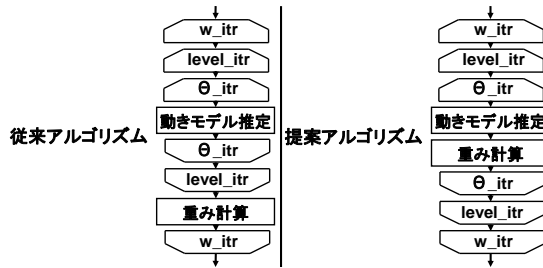


図1 アフィン動きモデル推定フローチャート

表1に正しい動きが既知の人工的なテスト動画像に対する従来法と提案法のシミュレーション結果を示す。動き推定精度の指標に Mean Angle Error (MAE: 平均角度誤差)を採用した。繰り返し回数を等しくしたとき提案法の精度は従指標の精度より若干良い。さらに提案法の繰り返し回数と精度の関係を図2に示す。提案法による動きモデル推定の収束に必要な $\theta_{itr}$ の回数は2回である。提案法は増分推定毎に重み推定を行うので $w_{itr}$ は不要である。従来法で提案法と同じ精度を得るには各ループの繰り返し4回が必要である。従って提案法は従来法の1/8の演算量で同等の精度が得られることが分かった。

表1 MAEによる提案手法の精度評価

MAE[°]	Translating Tree	Diverging Tree
従来法	0.143261	2.902186
提案法	0.134413	2.858279

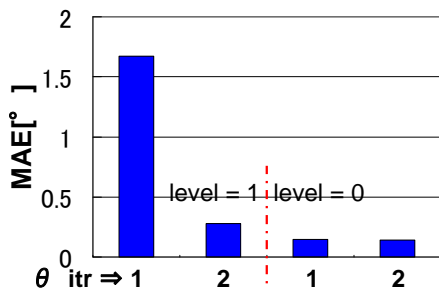


図2 Translating Treeの繰り返し回数と精度の関係

提案するアフィン動きモデル推定プロセッサのアーキテクチャを図3に示す。このプロセッサは多重解像度画像作成部、動きモデル

計算部、重み計算部、各種メモリから構成される。

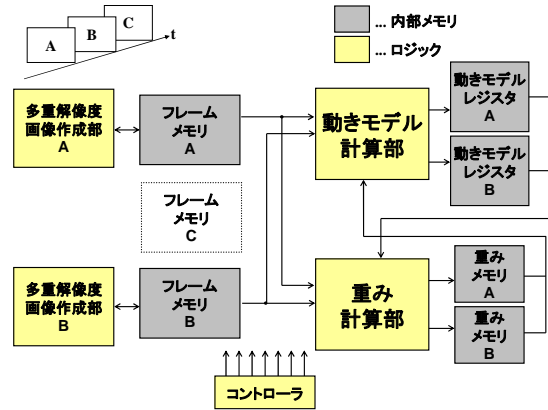


図3 アフィン動きモデル推定アーキテクチャ

提案アーキテクチャを前提とした従来および提案アルゴリズムのタイミング図を増分推定3回、解像度 level = 0, 1の2枚、重みの反復回数3回として図4に示す。図4の(b), (d)は(a), (c)の太枠部分の詳細である。従来法の場合、図2(a)のようにAB, BC間の動きモデルを同時計算させるフレームオーバーラップ手法により、AB間の動きモデル計算の間、BC間の重み計算を並列で行ないスループットを向上できる。しかし同時にABCの3フレームを処理対象とすることでチップ内部のメモリ量が増大する。提案法では $\theta_{itr}$ において直前の動きモデル推定結果を用いて同一の画像に対して重みを並列に計算する。これによりチップ内部に必要な画像を2フレームとし画像メモリを従来の2/3に削減した。前回の動きモデルがあれば現在の動きモデル計算中に対応画素の重みを計算できるため重みメモリも省略できる。

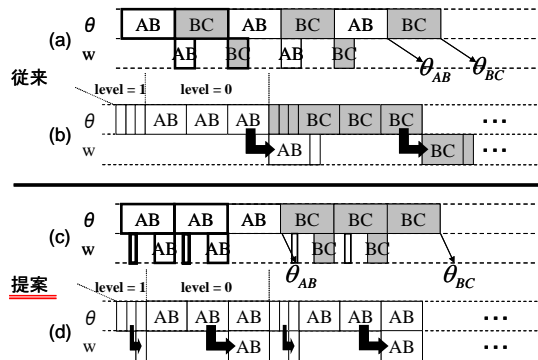


図4 アフィン動きモデル推定タイミング図

ローム 0.18  $\mu\text{m}$  5層メタルプロセスを用いて提案法に基づくVLSIプロセッサを試作した。レイアウト図を図5に示す。論理ゲート数は約60万ゲート、コア面積は4.0mm角となった。本回路の基本動作をLSIテストで確認した。本回路は動作周波数120MHzでVGA 240fpsのアフィン動きモデル推定が可能である。VGA 30fpsの推定に必要な周波数は15MHzであり、低電力化を達成した。

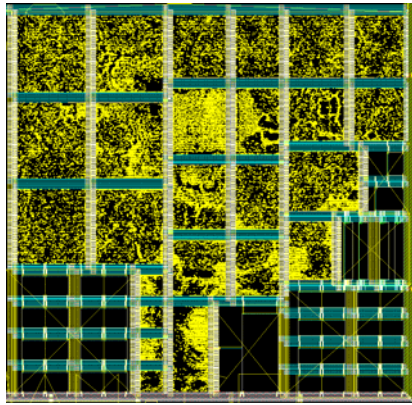


図5 レイアウト図

(2) 動領域抽出プロセス

従来の動領域抽出アルゴリズムを図6に示す。  $e(t)$  は時間  $t$  の画素と領域の対応を表すラベルマップ、  $\theta_t^{t+1}$  は時間  $t$  と  $t+1$  の間の各領域の動きモデルの集合を表す。

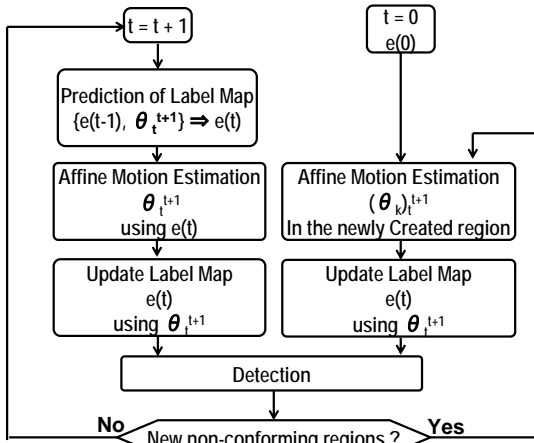


図6 動領域抽出フローチャート

アルゴリズムは次の4ステップからなる。

ステップ1: ラベルマップの予測

時間  $t-1$  で得られたラベルマップを動きモデルに従って移動させて時間  $t$  の分割ラベルマップ  $\tilde{e}(t)$  を得る。

ステップ2: アフィン動きモデル推定

分割ラベルマップ  $\tilde{e}(t)$  を用いて各領域の動きモデルの集合  $\theta_t^{t+1}$  を推定する。

ステップ3: ラベルマップの更新

分割ラベルマップ  $\tilde{e}(t)$  と動きモデルの集合  $\theta_t^{t+1}$  を用いて(1)式のエネルギー関数を最小化するラベルマップ  $\hat{e}(t)$  を得る。

$$U(e, o, \tilde{e}) = U_1(e, o) + U_2(e) + U_3(e, \tilde{e}) \quad (1)$$

観測値  $o$  は時間  $t$  と  $t+1$  の画像から成る。  $U_1$  項はラベルと観測値の適合性を表す。つまりラベルに対応する領域の動きモデルを適用して移動させた時間  $t$  の画素値が時間  $t+1$  の画素値と近ければ  $U_1$  項が小さくなる。  $U_2$  項はラベルの空間的な連続性を表す。  $U_3$  項はラベルの時間的な連続性を表す。画素毎に(1)式を最小化するラベルが選択される。

ステップ4: 新領域の検出

各領域において(1)式と同様のエネルギー関数を用いて動きがモデルと一致しない画素の領域を検出する。

提案する VLSI 実装向けアルゴリズムの特徴は以下のとおりである。

- ① 画像分割法
- ② 領域境界に限定した ICM ラベリング
- ③ 新領域の動きモデル推定の省略

①は1枚の画像を小さな複数の画像に分割して処理する方法である。これによりチップ内部のメモリ容量を削減し、チップ外部とのデータ転送量を削減できる。②は画素に領域ラベルを割り当てる処理を領域の境界に限定する方法である。これにより演算量が削減される上に、無意味な小領域の発生が抑止され、適切な領域分割結果が得られる。③はステップ4において新領域の検出の後の動きモデル推定を省略する方法である。これにより演算量が削減される。新領域の動きは0と仮定されて次の時間の処理が行われる。ロバストな動きモデル推定法を用いているため予測領域が多少不正確でも正しい動きモデルが得られ、適切な領域境界が得られる。新領域検出と動きモデル推定のループを除いた提案アルゴリズムのフローチャートを図7に示す。

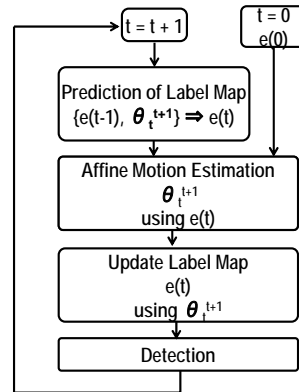


図6 提案フローチャート

従来および提案アルゴリズムによるシミュレーション結果を図7と図8に示す。どちらも適切な動領域抽出結果が得られている。提案法では分割画像の境界で領域境界が接続されている。VLSI 向けアルゴリズム最適化は精度を落とさずに演算量を82%、メモリ容量を95%、データ転送量を99%削減する。

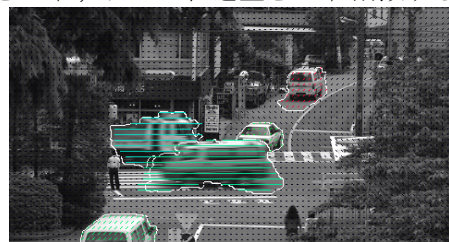


図7 従来法のシミュレーション結果



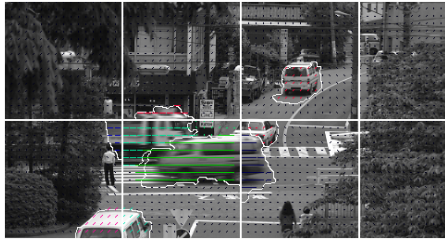


図8 提案法のシミュレーション結果

提案する動領域抽出プロセッサのアーキテクチャを図9に示す. このプロセッサは予測ラベルマップ作成部 (Prediction), アフィン動きモデル推定部 (PSM), ラベルマップ更新部 (Update)、新領域検出部 (Detection)、各種メモリから構成される. タイミング図を図10に示す. 分割画像を処理単位とするパイプラインはアフィン動きモデル推定のステージと領域分割のステージからなり、167MHzの動作周波数でVGA 30fpsの動領域抽出を行う. アフィン動きモデル推定部は複数の領域を並列処理できる. 0.18 $\mu$ mプロセスによるVLSI実装見積りを表2に示す. コア面積は約30mm<sup>2</sup>となり、低コストで実現できる見通しが得られた.

今後の課題として動領域抽出プロセッサの画像認識への応用があげられる. 本研究成果を発展させて車載、ロボット、監視などの分野に適用し、安全・安心な社会の実現に貢献していきたい.

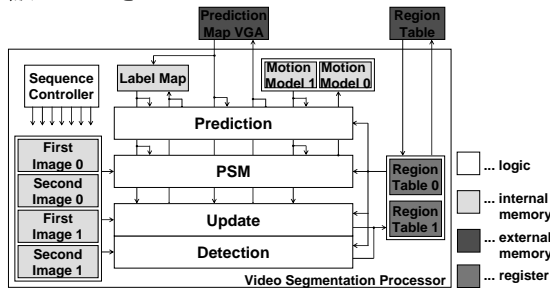


図9 動領域抽出アーキテクチャ

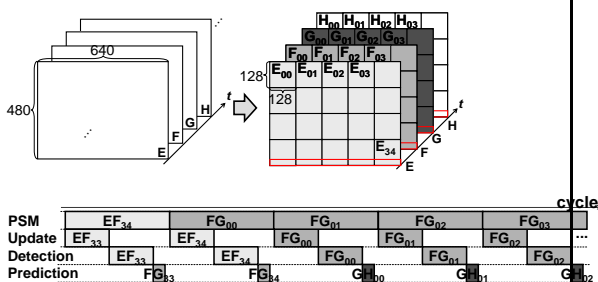


図10 動領域抽出タイミング図

表2 VLSI実装見積り

Performance	VGA 30fps
Operating Frequency	167MHz
Logic Gates(2 input NAND)	763,106 gates
Internal Memory	1port 85,380kByte
	2port 28,672kByte
Area(0.18 $\mu$ process)	29.536mm <sup>2</sup>

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表] (計4件)

1. 東郷康二、小松弘季、深山正幸、松田吉雄、“実時間動領域抽出向けアフィン動きモデル推定VLSIプロセッサ”、電子情報通信学会2010年総合大会、2010年3月10日、宮城県仙台市東北大学.
2. Masayuki Miyama, Yoshiki Yunbe, Kouji Togo, Yoshio Matsuda, “A VLSI Architecture for VGA 30 fps Video Segmentation with Affine Motion Model Estimation”, 12th International Symposium on Integrated Circuits (ISIC-2009), 2009. 12. 16, Suntec City, Singapore.
3. Y. Yunbe, M. Miyama, and Y. Matsuda, “A VGA 30fps Affine Motion Estimation Processor for Real-Time Video Segmentation”, IASTED Circuits & Systems, August 19, 2008, Kailua-Kona, Hawaii, USA.
4. 弓部 良樹、深山 正幸、松田 吉雄、“アフィン動きモデル推定プロセッサ用逆行行列計算回路”、電子情報通信学会2008年総合大会、2008年3月18日、福岡県北九州市北九州市立大学.

## 6. 研究組織

### (1) 研究代表者

深山 正幸 (MIYAMA MASAYUKI)  
金沢大学・電子情報学系・講師  
研究者番号：30324106

### (2) 連携研究者

松田 吉雄 (MATSUDA YOSHIO)  
金沢大学・電子情報学系・教授  
研究者番号：20401896