

# HDTV 解像度実時間動画画像認識応用 SIFT 特徴量抽出プロセッサ

発表者：神戸大学大学院 システム情報学研究科 水野 孝祐

Email: mi-no@cs.kobe-u.ac.jp

## Abstract

実時間動画画像認識応用 SIFT 特徴量抽出プロセッサを提案する。オリジナルの SIFT アルゴリズムの処理を VLSI 向きに改良し、さらに高並列アーキテクチャと 3 ステージ ROI パイプラインの導入をおこなった。その結果、HDTV 解像度の動画画像を 30fps で処理できる性能を低消費電力で実現した。

## Introduction

近年の画像認識技術の応用例としてはデジタルカメラに搭載されている顔検出や車載カメラ画像からの道路交通標識の認識などが挙げられる。画像認識はどのような環境で撮影しても、ロバストに高速・高精度処理を行うことが要求される。この要求に対して、モデルを局所記述子によって得られる特徴ベクトルで表現し、撮影画像の特徴ベクトルと比較・識別するものが提案されている。代表的なものとして SIFT[1]が挙げられる。SIFT を用いると、輝度値の変化・回転の影響を受けずに、高精度に画像認識を行うことができる。しかしながら SIFT を計算するためには莫大な演算量が必要となる。Intel Core2Duo E8400(3.0GHz)で SIFT 特徴量抽出を実行した場合の処理時間を図 1 に示す。VGA 解像度において 2.77 秒かかっており、実時間動作が難しい。さらに将来の高知能ロボット応用を想定した場合、認識に用いる動画画像の高解像度化が考えられる。そのため、高解像度対応の実時間 SIFT 特徴量抽出プロセッサの実現が必要不可欠であると言える。

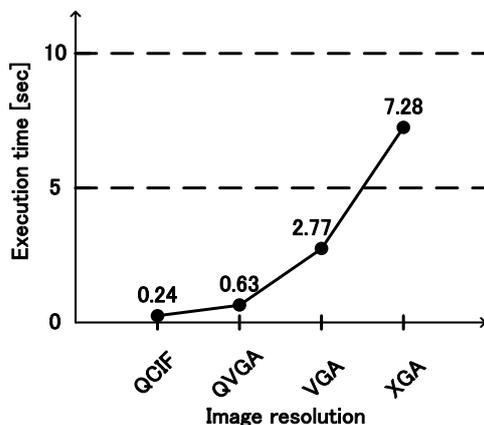


図 1 SIFT 特徴量抽出実行時間

高解像度対応の SIFT 特徴量抽出プロセッサを実

現するために、回路を高並列化する必要がある。HDTV 解像度における SIFT 特徴量抽出を行うための必要演算量の解析をおこなった。図 2 に SIFT 特徴量抽出の 3 大処理である、ガウシアン平滑化、特徴点抽出、特徴量記述の演算量の解析結果を示す。縦軸が演算量、横軸が特徴点数を表している。ガウシアン平滑化の演算量が支配的となっており、高速化阻害の一番の要因となっている。そのため、本設計ではガウシアン平滑化回路の並列度を 120 とし、HDTV 解像度動画画像に対してリアルタイム処理可能な性能を達成した。

特徴点数は写っている物体により大きく変わり、画像の複雑度によっては数万点となることも想定される。数万点の特徴点を扱う場合、特徴量記述が占める演算量も無視できない値となる。そのため特徴量記述の専用演算回路化も必要となる。

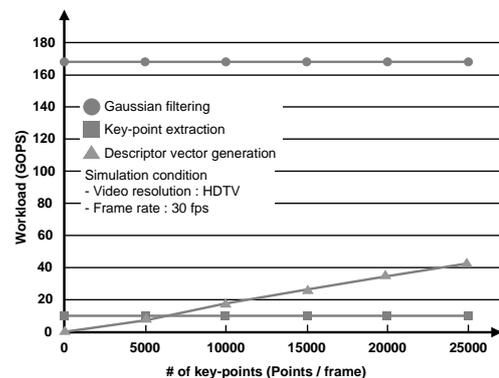


図 2 演算量解析

SIFT 特徴量抽出を実行するためのメモリ帯域の見積もりをおこなった(図 3)。横軸を解像度、縦軸をメモリ帯域としている。最も大きな帯域を占めるのはガウシアン平滑化であり、HDTV 解像度においては 5374Gbps である。特徴点抽出は 1254Gbps となり、ガウシアン平滑化に次いで 2 番目に大きなメモリ帯域を必要とする。演算量のみ考慮した場合、特徴点抽出は汎用 RISC プロセッサで十分であると考えられる。しかし上記のような大きなメモリ帯域が必要となるため、専用ハードウェアを用いた解決が必要となる。本設計ではガウシアン平滑化部、特徴点抽出部にシストリックアレイアーキテクチャを導入することで画素再利用、処理結果再利用を促進しメモリ帯域の削減をおこなった。

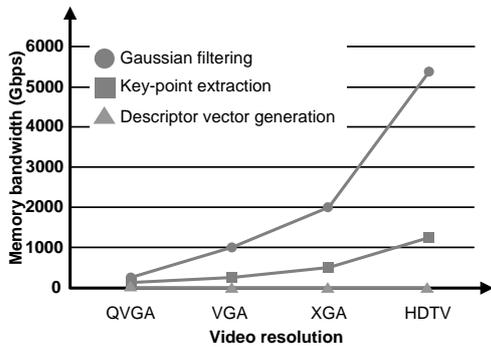


図3 メモリ帯域解析

さらに ROI 単位で処理を行う 3 ステージ ROI パイプラインアーキテクチャを導入することで、低消費電力化と回路の拡張性の確保を実現する。

### Sift Algorithm for VLSI Implementation

図4にVLSI向きに処理を改変した SIFT 特徴量抽出アルゴリズムを示す。入力画像を ROI に分割し、ROI 毎に処理していくことで内部メモリ容量を大幅に削減する。さらに全てのガウシアン平滑化を入力画像から直接生成することにより、処理遅延の問題を解決する。

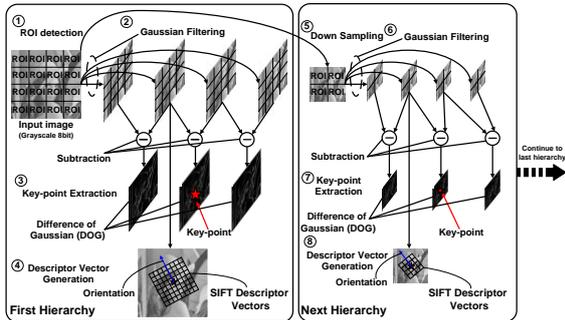


図4 VLSI 向き SIFT 特徴量抽出アルゴリズム

図5に提案アルゴリズムの導入による認識精度への影響を示す。図に示す通り、オリジナルに比べて 2.1%の精度劣化に抑えることができおり、シビアな安全性が求められるアプリケーションに対しては十分適用可能な値になっていると言える。

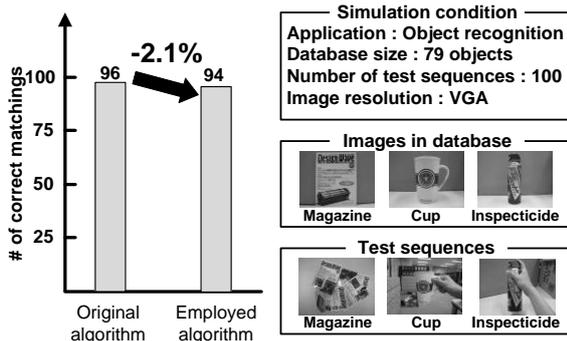


図5 認識精度評価

### Three-stage Pipelined Architecture

図6にROIを効率的に処理することができる3ステージROIパイプラインアーキテクチャのブロック図を示す。ガウシアン平滑化画像用SRAMを3枚実装しており、3つのモジュール(ガウシアン平滑化、特徴点抽出、特徴量記述)は別々のSRAMにアクセスする。

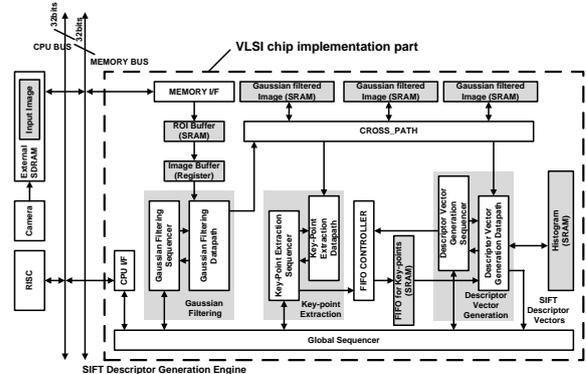


図6 3ステージROIパイプラインアーキテクチャ

図7に3ステージROIパイプラインの動作を示す。縦の点線で区切られた範囲が1つのROIを処理するために必要なサイクルタイムを表す。ガウシアンピラミッド生成部、特徴点抽出部、特徴量記述部はそれぞれ独立に動作し、それぞれ別々のROIに対する処理を行う。

1つのROIの処理にかかる時間は、3つのモジュールの内、最も処理に時間のかかるモジュールに依存することになる。1つのモジュールの処理の完了を待つ間、他の2つのモジュールにはアイドルサイクルが発生する。そのためクロックゲーティングなどを用いてアイドルサイクル中に無駄な電力が発生しないようにしなければならない。本パイプラインを用いることで、各モジュールが必要とするサイクル数が事前にある程度把握できるため、回路の低消費電力化技術であるクロックゲーティングや周波数制御などを適用することが容易となる。

提案アーキテクチャを用いることでROIを効率的に処理することができ、大幅な消費電力削減効果が得られる。

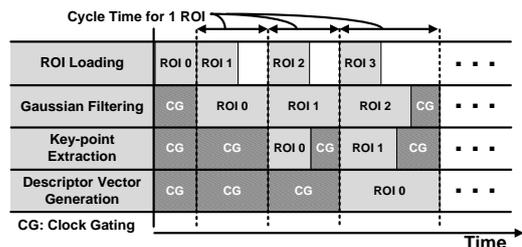


図7 3ステージROIパイプライン

## Gaussian filtering Architecture

図 8 に本設計で実装した 120 並列ガウシアン並列化アーキテクチャを示す。本アーキテクチャは 120 個の 1 次元リング状シストリックアレイ(SA)から構成される。1つの SA は 2つの積和演算器, 14 段のリング状レジスタ, セレクタ回路から構成されている。本アーキテクチャによるガウシアン平滑化ではウィンドウサイズとして 15 を想定している。この時, 1 次元ガウシアン平滑化処理を 1 画素ずれた点に対して連続的に行うことを考えると, 隣り合うガウシアン平滑化処理同士では 14 画素が重複することになる。そのため, 隣り合うガウシアン平滑化画素の生成の際に再利用できる 14 画素分のレジスタを用意している。2つある MAC はそれぞれが 1 次元ガウシアン平滑化に対応している。

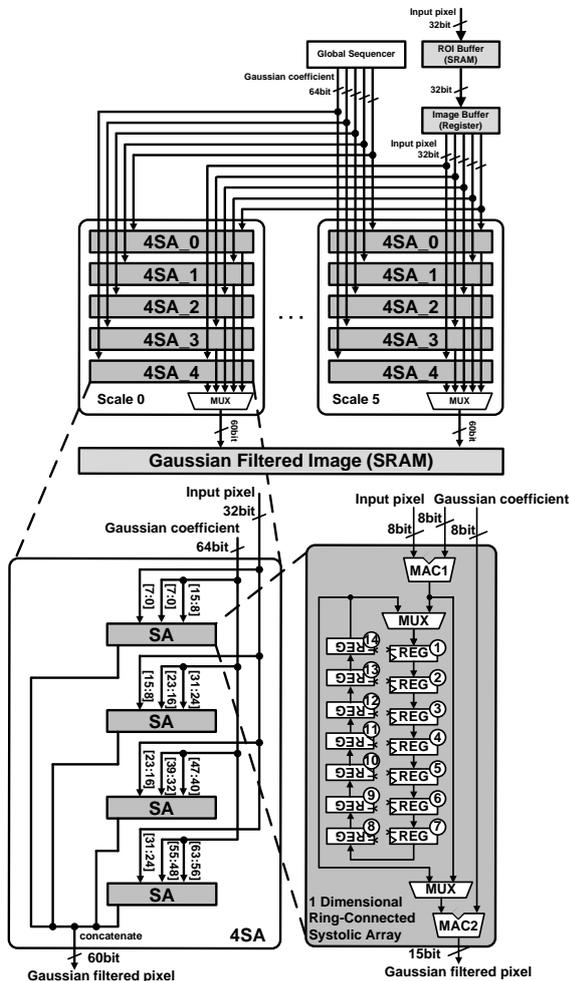


図 8 ガウシアン平滑化アーキテクチャ

## Key-point Extraction Architecture

図 9 に 27 並列特徴点抽出シストリックアレイアーキテクチャを示す。全体としては 27 個の Processing Element(PE), 9 個の差分演算器(Subtractor), AND 回路から構成される Systolic Array 型のアーキテクチャである。中段の真ん中の PE からの出力は全 PE に接続されており, 比較演算のために用いられる。上

段の PE は中段の PE へ, 中段の PE は下段の PE へ接続されており, 1 サイクル毎に画素データがシフトしていく。この構造により画素の再利用が実現される。

次に PE への画素ロードの流れについて説明する。1 サイクルごとにガウシアン平滑化画像 SRAM から 18 画素を読み出し, 差分演算器により DoG 画像 9 画素を生成する。この 9 画素を毎サイクル PE に入力する。シストリックアレイの動作は, PE に初期画素をロードすることから始まる。初期化するためには 27 個の PE 全てに画素をロードする必要があるため, 3 サイクル必要となる。初期化のための 3 サイクルが経過すると, 27 個の PE に画素データが蓄積し, 極値検出の結果が出力される。その後は 1 サイクル毎にデータが一段ずつ下にシフトしていき, 1 サイクル毎に 1 点に対する極値検出結果が出力される。

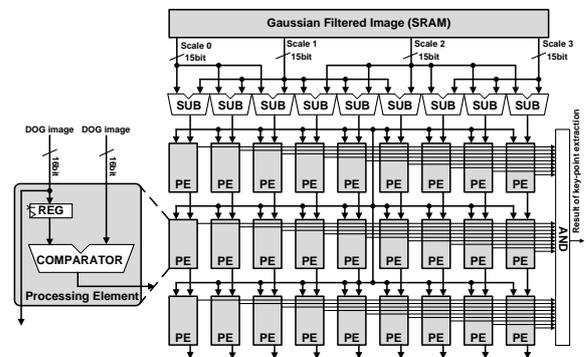


図 9 特徴点抽出アーキテクチャ

## Descriptor Vector Generation Architecture

図 10 に特徴量記述アーキテクチャのブロック図を示す。特徴量記述アーキテクチャは制御回路(Sequencer), 2 並列の向き計算部(Orientation), 2 並列の特徴量記述部(Description)から構成される。本アーキテクチャの最も特徴的な部分である向き計算部は, 勾配計算部, COOrdinae Rotation Digital Computer (CORDIC)演算部, ヒストグラム生成部によって構成されている。CORDIC はハードウェアで三角関数を計算するための手法であり, 広く用いられている。

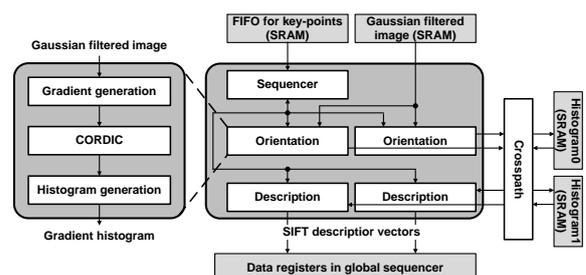


図 10 特徴量記述アーキテクチャ

## Performance Evaluation

図 11, 12 に提案アーキテクチャの導入による gaussian 平滑化, 特徴点抽出, 特徴量記述それぞれのサイクル数とメモリアクセスの削減効果を示す. SIMD アーキテクチャと比較して, Gaussian 平滑化においてはサイクル数を 82%削減し, メモリアクセスを 99.8%削減する. 特徴点抽出においてメモリアクセスを 66%削減する. 提案アーキテクチャは SIFT 処理のボトルネックとなっている (1)Gaussian 平滑化の演算量, (2)Gaussian 平滑化のメモリ帯域, (3)特徴点抽出のメモリ帯域の 3つの問題を解決する.

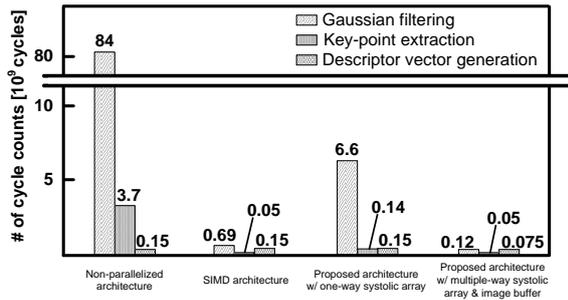


図 11 サイクル数削減効果

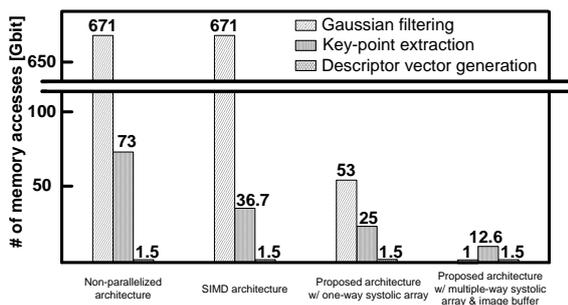
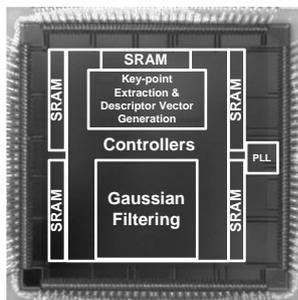


図 12 メモリアクセス削減効果

## Hardware Results

設計したプロセッサは論理ゲート 1.1M ゲートで SRAM 容量は 1.38M ビットになった. 図 13 に設計したプロセッサのチップ写真と仕様を示す.



Technology	65nm CMOS
Supply Voltage	Core 1.2V, I/O 3.3V
Chip Size	4.2mm x 4.2mm
Core Area	3.2mm x 3.4 mm
Logic Gate Count	1.1M gate
On-chip SRAM	1.38Mbit
Operating Frequency & Power Consumption	38.2mW@78MHz for 1920x1080(HDTV)/30fps 14.5mW@27MHz for 640x480(VGA)/60fps
Target Application	SIFT descriptor generation

図 13 チップ写真

図 14 に提案プロセッサの周波数・消費電力特性の実測結果を示す. 青線がコア電圧 1.2V 固定時のグラフ, 赤線が最低動作電圧時のグラフである.

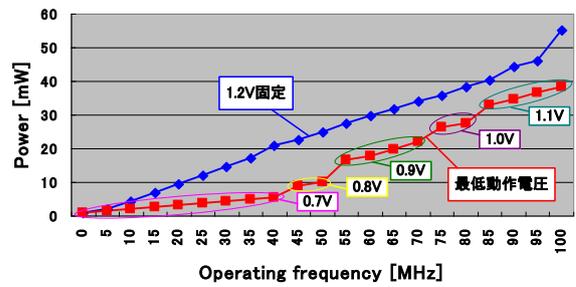


図 14 P-f 特性

図 15 に過去に提案されている SIFT 特徴量抽出プロセッサ[2][3]との性能比較を示す. [2]と比較してフレームレートについては 53 倍となっており, 高速性が要求されるアプリケーションに適用可能である. さらに消費電力については 38.2mW となり, 39%削減することができる. そのため限られたバッテリー環境で動作する, モバイルアプリケーションへの適用が期待できる.

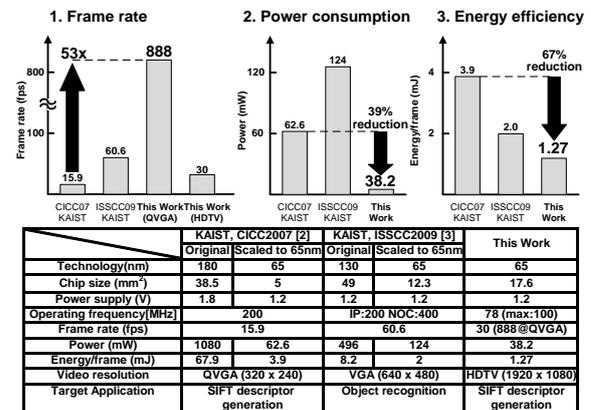


図 15 性能比較

## References

- [1] D. Lowe, "Distinctive image features from scale-invariant keypoints," in International Journal of Computer Vision, vol. 20, 2003, pp. 91–110.
- [2] D. Kim, et al., "An 81.6 GOPS Object Recognition Processor Based on NoC and Visual Image Processing Memory", CICC, pp.443-446, Sep. 2007.
- [3] J. Y. Kim, et al., "A 201.4GOPS 496mW real-time multi-object recognition processor with bio-inspired neural perception engine", ISSCC Dig., pp.150-151, Feb. 2009.