

# 実世界情報システムプロジェクト ～視聴覚研究グループ石川・橋本研究室～

石川正俊 橋本浩一 小室 孝  
情報理工学系研究科システム情報学専攻

## 概要

1 秒間に 1000 フレームの画像処理能力を持つ高速ビジョンシステムを複数台用いて人間の動きや意図を人間が行動を終了する前にあらかじめ察知するようなセンシングシステムを開発する。そして、人を超える性能を提示するデモンストレーションを実現することで、システムの有効性を目に見える形で示すことを目標とする。

## 1 はじめに

実環境で人とロボットがインタラクトする場合、システムが人間と直接的に触れ合うのでビジョンなどのセンサによって人間の動きを素早く観測し、正確かつ安全にロボットを制御する必要がある。

一方、我々はこれまでに、視覚情報を実時間に取得・処理する視覚情報処理デバイスとして、イメージセンサの画素毎に処理回路を取り付けたビジョンチップおよびその技術を応用した高速ビジョンシステムの開発を行っており、それを用いて従来の視覚センサでは不可能な 1ms の高速視覚フィードバックを実現している。視覚情報は実世界情報の中でも特に情報量が多く、人間の活動の基本となるものであり、これを実時間に取得・処理することが、情報システムの高度化において重要であるといえる。

本研究では、この 1 秒間に 1000 フレームの画像処理能力を持つ高速ビジョンシステムを複数台用いて人間の動きや意図を人間が行動を終了する前にあらかじめ察知するようなセンシングシステムを開発する。

具体的には、視覚によって微少な人間の行動を高速に検出するアルゴリズムおよび視覚特徴量から人間の意図を知るための学習アルゴリズムを開発する。さらに、人と対戦するゲームなどにおいて人を超える性能を提示するデモンストレーションを実現することで、システムの有効性を

目に見える形で示すことを目標とする。

本報告では、平成 14 年度の成果として、超並列アーキテクチャに基づく、1 秒間に 1000 フレームの撮像・処理が可能な高速ビジョンシステムの概要について報告する。

## 2 ビジョンチップ

従来の画像処理システムにおいては、ビデオカメラから処理装置にデータを送信する際の走査によるボトルネックが存在し、フレームレートに制限があった。それに対し、センサと並列の処理要素 (PE) を一体化し、ワンチップに収めたビジョンチップと呼ばれるデバイスを用いることで、ビデオレートをはるかに超える視覚処理が可能となる。このような高いフレームレートを用いることにより、安定したビジュアルフィードバックが実現され、たとえば視覚情報を用いたロボットのリアルタイム制御が可能となる。

これまで研究されてきたビジョンチップは主に処理要素がアナログ回路で構成されていたが、近年では PE にデジタル回路を採用することにより高度な視覚処理を実現するビジョンチップも少数ではあるが登場している。しかしながら、デジタル回路はアナログ回路に比べ、回路規模が大きくなってしまいうという問題があり、高機能化・高画素化のためには高集積化設計が必須となる。そこで、従来のスループットを重視するコンピュータ設計から、回路面積を重視する設計に方針を切り替えることで、十分な数のデジタル PE をワンチップに搭載できるようになった。そのようにして設計したビジョンチップのアーキテクチャを図 1 に示す。

PE は、主に ALU とローカルメモリから構成されている。ALU は全加算器、レジスタといくつかのマルチプレクサからなる簡単な構成で、論理演算と算術演算を共通の回路で実行する。

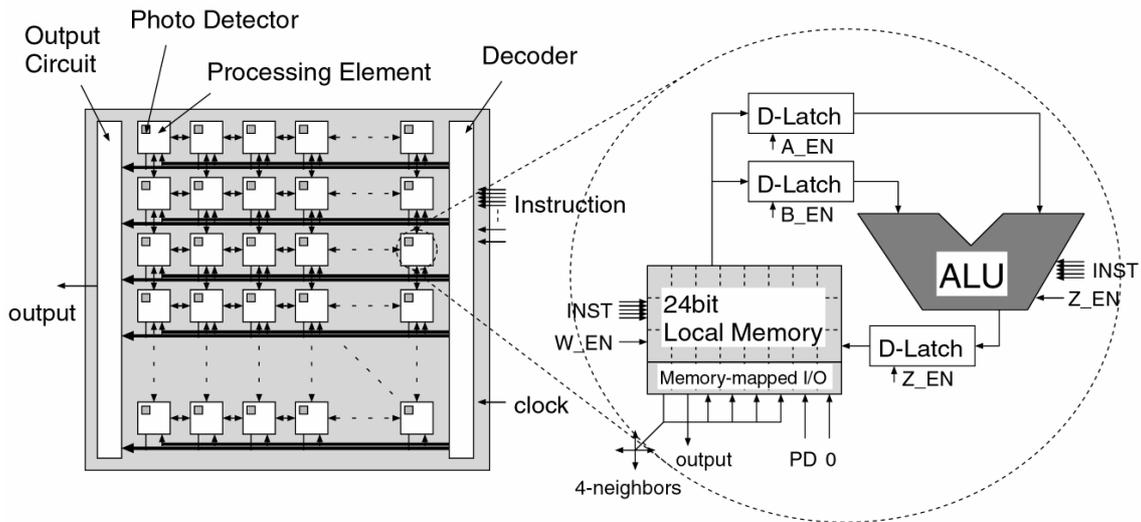


図1 ビジョンチップのアーキテクチャ

一度に実行できるのは1ビット分の演算のみであり、多ビット演算はこれを繰り返すことで実現されるというビットシリアル演算を採用している。ローカルメモリは24ビットのランダムアクセスメモリ (RAM) に加え、8ビットの I/O ポートが同一のアドレス空間に割り当てられている。これらはビット単位でランダムにアクセスすることが可能となっている。I/O ポートは上下左右の4近傍との通信、センサからの入力、および零信号と接続されている。

フォトディテクタからの入力は、電圧値がインバータのスレシールドを切るまでの時間を計測するという仕組みになっている。その際、PE をカウンタとして用いることにより、回路を追加することなく A/D 変換を実現している。

本アーキテクチャに対し、エッジ検出やノイズ除去といった初期視覚処理アルゴリズムを実装したところ、どれも  $\mu\text{s}$  のオーダーで実現されており、ビデオレートをはるかに超える視覚処理が可能であることが示された。

さらにセルフウィンドウ法と呼ばれるアルゴリズムを用いることで、動物体の追跡処理が可能である。このアルゴリズムは高速ビジョンにおいてはフレーム間で対象がほとんど動かないという特性を活かし、極めて単純な演算のみで実現されている。

また、視覚処理の結果を画像で出力するわけには行かないので、なんらかの特徴量をビジョンチップ上で高速に計算する必要がある。本アーキテクチャには特徴抽出のための回路は備えられて

おらず、繰り返し演算によって求めることになるが、専用回路の提案もなされており、それが導入されると高速化される見通しである。

上記アーキテクチャに基づき試作されたフルカスタムチップは、 $0.35\ \mu\text{m}$  プロセスを用い、PE を回路面積  $105 \times 105\ \mu\text{m}^2$  で実装することができ、 $64 \times 64$  画素を  $8.7\text{mm} \times 8.7\text{mm}$  のチップ上に集積されている。写真を図2に示す。

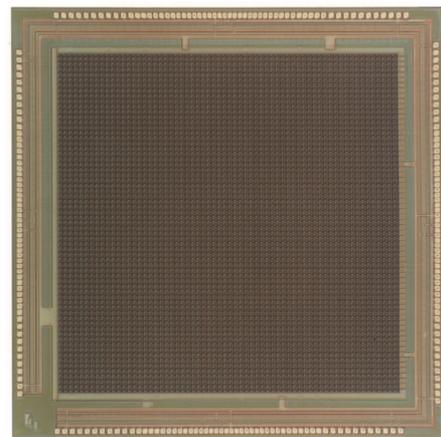


図2 試作されたビジョンチップ

### 3 ビジョンチップを用いた高速ビジョンシステム

ビジョンチップを用いてリアルタイムのセンサ情報処理を行おうとした場合、通常の視覚処理を大幅に上回るフレームレートで動作するビジョンチップの制御には、厳しい実時間性が要求さ

れる。その要求を満たすため設計したビジョンチップのコントローラのアーキテクチャを図3に示す。

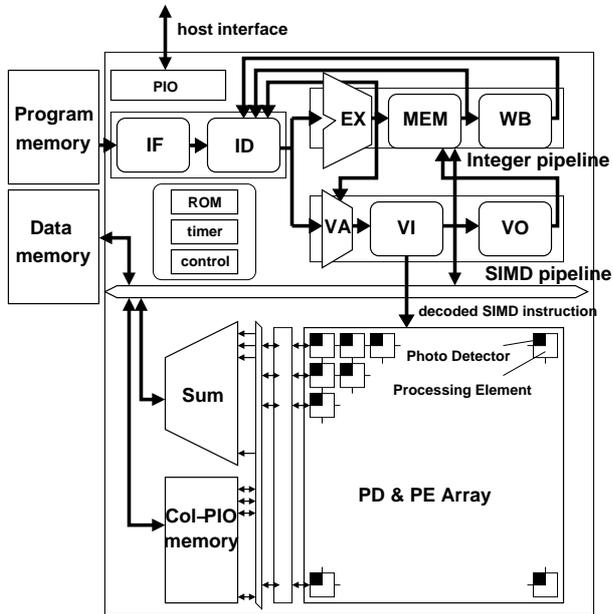


図3 ビジョンチップコントローラのアーキテクチャ

RISC 型 32 ビットマイクロコントローラのパイプラインにビジョンチップの制御フローを統合することでインストラクションレートの向上を図り、かつ、パイプライン内/パイプライン間フォワードイング、遅延分岐/ロード、固定レイテンシのメモリアクセスなどによりパイプラインのストールを回避し、インストラクションサイクルの粒度での実時間性を保証している。

本アーキテクチャに基づき開発されたコントローラとビジョンチップを実装したワンボード型のシステムを図4に示す。

ビジョンチップコントローラは XILINX 社の FPGA XC4062XLA に実装されており、FPGA を書き換えることで複数の種類のビジョンチップによるインタフェースの差異を吸収することができる。ビジョンチップコントローラは FPGA 上に、約 27000 ゲート換算の規模で実装されている。これは 64×64 画素のビジョンチップと比較するとおよそ 1 割弱の規模であり、将来的にビジョンチップとともにワンチップに集積することも十分現実的であると考えられる。この FPGA 上の実装で、1 パイプラインステージ 100ns での動作が確

認できており、その分解能でインストラクション供給時刻を制御できる。

プログラムメモリおよびデータメモリは、容量は各々 32K ワードでデュアルポート構成となっており、ホストコンピュータ上で生成した実行オブジェクトのダウンロードや、メモリ空間の内容の確認などを任意に行うことができる。モニタリング用のルーチンを実行オブジェクトに組み込むことで、処理結果の画像をホストコンピュータ上でリアルタイムに表示しながらの動作検証も可能である。

ホストコンピュータとのインタフェースはシンプルな平行 I/O 接続を採用しており、汎用の平行 I/O インタフェースを利用することで PC や DSP 等のさまざまなホスト環境に対応できる。

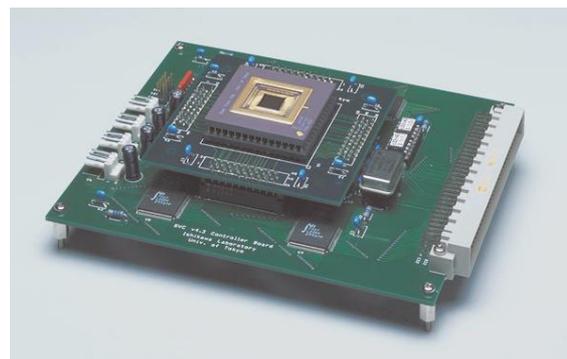


図4 ビジョンチップシステム

ビジョンチップシステム上で動作するソフトウェアの効率的な開発を可能とするためには、ソフトウェア開発環境の整備が重要となる。

ビジョンチップシステムに対応したプログラミング言語として、SPE-C が提案されている。これは ANSI C ライクな文法を基本として、ビジョンチップの各画素上のデータに対応した型 parallel を導入することにより、SIMD 型の並列計算を中心として記述されるビジョンチップシステム上の処理を、自然に表現できるよう拡張された言語である。

先に述べたビジョンチップシステムに対応して、この SPE-C 言語による開発を可能とする開発環境を用意している。コンパイラ・アセンブラ・リンカといった処理系に加えて、PC 上でのシミュレーション環境が用意されている。シミュレータの実行画面を図5に示す。

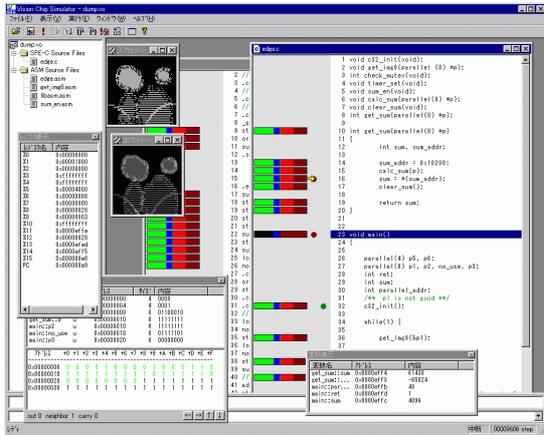


図5 ソフトウェア開発環境

#### 4 列並列高速ビジョンシステム

センサとプロセッサのワンチップ化は、速度やサイズの点では有利であるが、センサの感度や精度、多画素化・高機能化の可能性、開発の困難さなどを考慮すると、フォトディテクタとPEアレイを分離して作成するシステムが有利になる場合もある。この場合、PEやコントローラの構造はビジョンチップのものを流用できるが、LSIのピン数に制限があるため、完全並列タイプのアーキテクチャは採用できず、フォトディテクタとPEアレイの接続に列並列転送を採用することになる。このようなタイプを列並列高速ビジョンシステムと呼び、浜松ホトニクス(株)と実用化に向けて研究を進めている。列並列高速ビジョンシステム概念と開発したシステムの外観を図6,7に示す。128×128画素に対して、完全並列処理により、1msのフレームレートでの画像処理を実現している。

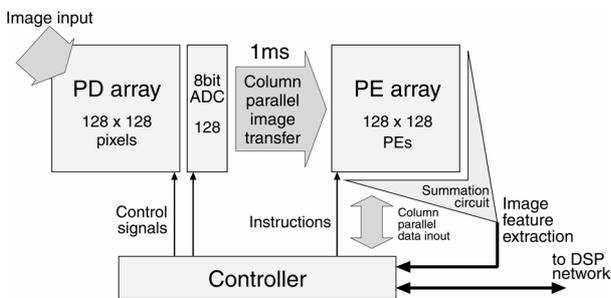


図6 列並列ビジョンシステム概念



図7 開発された列並列ビジョンシステム

#### 5 おわりに

実環境において人とロボットがインタラクトする際に必要となる、視覚情報を実時間に取得・処理するための、1秒間に1000フレームの画像処理能力を持つ高速ビジョンシステムについて報告した。

本システムを複数台ネットワークでつなぎ、そこからの情報の統合を行うアルゴリズムを開発することで、人の運動を計測・認識することが次の課題である。

#### 参考文献

[1] Shingo Kagami, Takashi Komuro, Idaku Ishii, Masatoshi Ishikawa : A Real-Time Visual Processing System using a General-Purpose Vision Chip, Proc. 2002 IEEE International Conference on Robotics and Automation, pp. 1229-1234, 2002

[2] 石川正俊, 小室孝 : デジタルビジョンチップとその応用, 電子情報通信学会論文誌 C, Vol.J84-C, No.6, pp.451-461, 2001