

高速ビジョンシステム

石川正俊 並木明夫 小室孝 鏡慎吾

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 16 年度の成果として、前年度までに開発した超並列ビジョンシステムを、ネットワークに接続するためのシステム開発の概要を報告する。

2 高速ビジョンネットワークの設計方針

人間とロボットが共存する環境において、人物やロボット、あるいはその周囲の物体の動きを視覚により認識することを考えると、観測対象の領域を十分に広く取れ、オクルージョンによる死角などが可能な限り存在しないことが望ましい。また、複数の観測対象を同時に追跡でき、かつ個々の対象を詳細に観測できることが重要となる。このような要求を満たすためには、複数のビジョンシステムの協調が不可欠となる。

高速ビジョンシステムのネットワーク化を行う際の通信メディアを選定する際、まず重要なのは、1000 fps の高速な視覚情報を、その価値を失わずに伝達できる相応の帯域と遅延性能を持つ点である。

さらに、環境中に偏在するさまざまな情報機器から、高速ビジョンシステムが得た視覚情報を自在に利用でき、連携アプリケーションを柔軟に構成できることが求められる。その際、ユーザとな

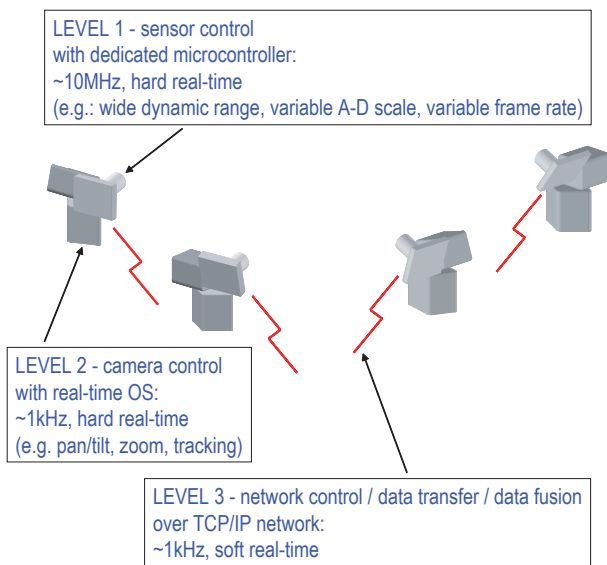


図 1: 高速ビジョンネットワークのコンセプト

る情報機器の側に特殊な入出力デバイスや特定の OS を仮定しないことが望ましい。特に、本プログラムでは研究ユニット間の連携が重要なポイントであり、この点に関する要求は強いものとなる。

以上を満たすものとして、本研究では 100BASE-TX イーサネット上の TCP/IP によるネットワークを採用した。TCP/IP ネットワークの高い汎用性と相互接続性により、さまざまなユーザシステムとの連携アプリケーションの実現が期待できる。

なお、より柔軟なシステム構築のためには無線による接続が望まれるが、現時点では特に遅延に関して十分な性能を得ることが困難と考えられるため、対応を見送った。将来的には対応を検討したい。

図 1 に、高速ビジョンネットワークのコンセプトを示す。リアルタイム性の粒度・性質の観点から見ると、3 つの異なるレベルからなる階層的な構造となる。

LEVEL 1 最も細粒度のレベルでは、ビジョンチップ専用コントローラによる、命令サイクルの粒度でのリアルタイム性保証がなされる。これにより、広ダイナミックレンジ撮像、可変 A-D 変換特性、可変フレームレートなど

のセンサ制御をソフトウェアで確実に行うことができる [1]。

LEVEL 2 第 2 のレベルでは、リアルタイム OS によるミリ秒の粒度でのリアルタイム性が保証される。フレームごとの処理タスクや、アクティブビジョンを採用する際のパン・チルト制御、ズーム制御などはこのレベルで実行できる。

以下で述べる第 3 レベルではリアルタイム性が保証されないため、例えば視覚情報のロスにより対象を見失う場合が生じ得るが、この第 2 レベルでハードリアルタイムを保証することにより、個々のカメラの単位では対象は継続して追跡できており、その情報を第 3 レベルに反映させることで見失った対象の追跡を回復するといったことが可能となると考えられる。

LEVEL 3 第 3 のレベルは、厳密なリアルタイム性の保証がない TCP/IP 上での通信を介する。時間粒度としては、第 2 のレベルと同等のミリ秒の粒度を指向するが、ハードリアルタイムである第 2 レベルとは異なり、ソフトリアルタイムな階層となる。このレベルにて、視覚情報の送受や複数センサ情報の統合がなされる。

このレベルではリアルタイム性を保証する機構はないため、視覚情報の一部が突発的に遅延したり、失われたりすることがあり得る。よって、より高次の処理レベルでの予測や補間、また、複数のセンサあるいは異種のセンサからのデータの統合による情報の確実化などが必要とされる。

そもそも画像の理解という問題は困難であり、観測対象の検出が常に確実に行われるとは必ずしも期待できないことを考えると、このような高次の処理レベルにおける情報の確実化は、ネットワーク構成の如何に関わらず本質的に必要なものであると考える。

3 設計と実装

前年度までに開発を行って来たビジョンシステム VCS-IV [2] をベースとして開発を行った。VCS-IV では、センサボード、コントローラボード、インタフェースボードなど複数の構成要素に機能を分散し、必要に応じてそれらをスタックして使用する。この構成ボードの 1 枚として、前述のレベル 2 に対応するネットワークプロセッサボードを開発するのが最終像であるが、本年度はまず試験的な実装として、市販の組み込み用プロセッサボードを接続したプロトタイプを開発した。

VCS-IV のビジョンチップコントローラは FPGA に実装されている。コントローラボードと外部システムとの入出力は、すべてこの FPGA の I/O を通じて行うように構成されている。市販のプロセッサボードを接続する場合は、その仕様に合わせてインタフェースロジックやその他のグルーロジックを用意する必要があるが、この構成により、FPGA 内のモジュールの入れ換えだけで済ませることが可能となった。

プロセッサボードは、アルファプロジェクト製 MS104-SH4 を採用した。CPU はルネサス SH-4/240MHz (SH7750RF240)、イーサネットコントローラは SMSC LAN91C111 を搭載している。リアルタイム OS として、 μ ITRON 4.0 仕様準拠の市販 OS であるミスポ製 NORTi 4 がインストールされている。TCP/IP プロトコルスタックも利用可能である。

システム構成のブロック図を図 2 に示す。MS104-SH4 の外部バスである PC/104 バスのスレーブインタフェースをビジョンチップコントローラボードの FPGA 内に実装し、接続した。ソフトウェアからは、CPU のメインメモリ空間にビジョンチップコントローラのプログラムメモリ・データメモリ空間がマップされたように見え、共有メモリ型で通信が行われる。実装されたシステムの写真を図 3 に示す。

ユーザシステムからのリクエストを受け取り、それに基づいてシステムの制御を行うリアルタイムタスクを実装した。従来の構成である PC との平行 I/O 接続上での通信と同等の機能を

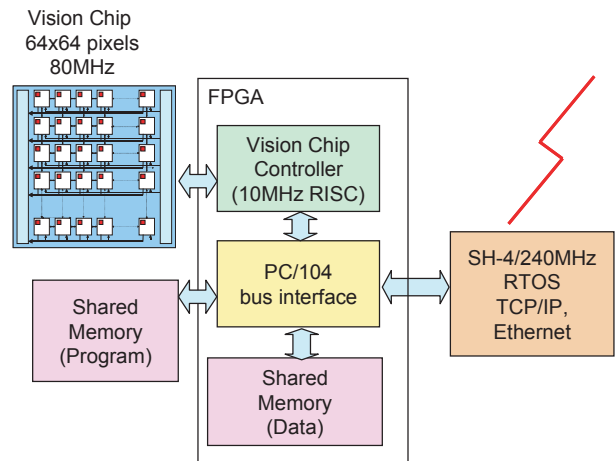


図 2: システムの構成



図 3: 実装されたビジョンシステムの外観

TCP 上で実現したほか、UDP または TCP を用いて、カメラからネットワーク上の他のノード (他のカメラや PC 等を含む) にデータを一定の周期 (最小 1 ms) で送信し続けるモードを実装した。100BASE-TX イーサネットスイッチで接続された LAN 環境で通信実験を行ったところ、重心等の画像特徴量のサイズのデータであれば、遅延やジッタは概ね数百 μ s に収まることが分かった。

一方で、データを送信するノードが増えた場合、通信の衝突によりパフォーマンスが低下することが予想される。ネットワークシミュレータ ns-2 を用いて評価したところ、20 ~ 30 台を超えたあたりから遅延・ジッタが数 ms を超え始め、急激に悪化することが分かった。その場合も、カメラを

グループ化してクラスタ化すれば 100 台程度の通信は 1 ms 以下の遅延・ジッタに概ね収まり，マルチホップさせることによる遅延を考慮しても，その方が望ましいという見通しを得ている．

4 おわりに

これまで開発を継続して来た超高速ビジョンを他のシステムとの相互接続性の高い TCP/IP ネットワークで複数台接続することのできるネットワークシステムを開発し，1000 fps の視覚情報をそのネットワーク上で送受できることを確認した．今後ソフトウェア環境を整理し，ユーザシステムから柔軟に利用可能な高速ビジョンシステムネットワークとしての整備を継続して行く．

参考文献

- [1] 鏡慎吾, 小室孝, 藤村英範, 石川正俊. デジタルビジョンチップのためのソフトウェア A-D 変換手法. 映像情報メディア学会誌, Vol. 57, No. 3, pp. 385–390, 2003.
- [2] 鏡慎吾, 小室孝, 渡辺義浩, 石川正俊. ビジョンチップを用いた実時間視覚処理システム VCS-IV. 電子情報通信学会論文誌 (D-I), Vol. J88-D-I, No. 2, 2005. (to appear).