

超並列・超高速ビジョンを用いた 1ms ターゲットトラッキングシステム

中 坊 嘉 宏* 石 井 抱* 石 川 正 俊*

1 ms Target Tracking System Using Massively Parallel Processing Vision

Yoshihiro Nakabo*, Idaku Ishii* and Masatoshi Ishikawa*

It is obvious that visual servo systems have various applications for real time robot control. But most conventional vision systems have serious restriction on their performance, since those systems always use CCD cameras to acquire the images which transmitted by serial video signals. Therefore the sampling rate is limited by the video frame rate, and this restriction on the speed is quite insufficient to control of the robot.

To solve this problem we have developed a massively parallel processing vision system called SPE-256 in which the photo-detectors and processing elements are directly connected. We have realized high speed visual feedback with 1 ms cycle time.

In this paper we describe our 1 ms visual feedback system and its performance in the application of high speed target tracking.

Key Words: Visual Servo, Active Vision, Visual Feedback, Vision Chip, Robot Vision

1. はじめに

未知環境、あるいは動的に変化する環境や対象物に対してロボットを制御するためには、外界の状況を把握する手段として、視覚情報を利用したビジョンシステムを用いることが有効である。特に最近では、視覚センサをフィードバックループのなかに組み入れてロボットを制御するビジュアルサーボの研究が盛んになされている。このようなシステムでは大量の画像情報を実時間で処理する必要があるが、近年の計算機能力の急速な向上によってこれが実現しつつあり、様々な応用が期待されるようになってきている。

しかし、現在研究されているほとんどのビジョンシステムで問題となるのは、画像の取り込み部分に CCD カメラなどを用い、ビデオ信号によるシリアル伝送を行っている点である。なぜなら、いくら高速な画像処理を実現したとしても、ビデオ信号を用いる限りそのフレームレートより高速なセンサフィードバックは不可能であり、結果としてサーボシステムのフィードバックレートはビデオレート (30 [Hz]) で制限されることになるからである。それに対して、一般的なロボットのサーボコントローラでは 1 [kHz] 程度の制御レートが必要だと考えられることから [1], ビデオ信号を用いた従来のシステムでは、ビジュアルフィードバックだけを用いて実時間でロボット制御を実現

するのは、実際には困難である。

このような問題を解決するためには、従来の方法でボトルネックとなっている画像のシリアル伝送の部分解消し、センサ情報を伝送せずに画像を取り込んだその場で処理を行えばよい。そのために、画像センサと画像処理部を一体化し、それぞれの画素について完全な並列処理を実現すればよいと考えられる。このような観点に従って、従来から、センサとアナログ演算回路による処理部を一体化したビジョンチップ [2] などいくつか提案されており、実際に試作されたものもある。ところが、これらのビジョンチップは特定の処理や用途に固定されたものがほとんどであり、汎用性の面で問題がある。特に、実際にビジョンチップを用いてロボットを制御しようとするとき、画像に含まれている様々な情報を、制御システムの必要に応じて抽出する能力が必要となり、アナログ固定回路によるビジョンチップのように用途が限定されるのは好ましくない。

汎用性のあるビジョンチップを実現するためには、Fig. 1 に示すように、並列処理回路としてデジタル回路による汎用プロセッシングエレメント (PE) を用いることが有効となる。ただし、センサと一体化した超並列演算処理機構を実現しようすると、既存の汎用プロセッサ程度のもものでは回路規模が大き過ぎて、多数の PE を同一チップ上に載せることができなくなる。そこで集積化を念頭において回路規模をなるべく小さく、しかもある程度の汎用性を保つという相矛盾する要求仕様を満たすものとして、石川らは、SPE (Sensory Processing Elements) を提案し [3], 実際に PE を 8 個集積化した SPE-8 を試作して

原稿受付 1996年2月2日

*東京大学大学院工学系研究科

*Faculty of Engineering, Univ. of Tokyo

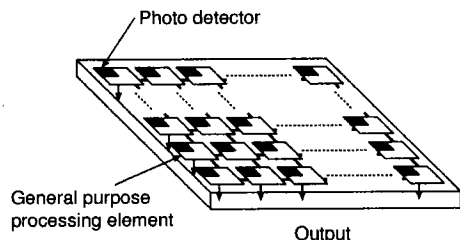


Fig. 1 Concept of general purpose vision chip

いる [4].

この SPE のアーキテクチャを用いることで、超並列・超高速ビジョンチップに様々な画像処理アルゴリズムを実装することが可能となり、ロボットの制御をはじめとして、様々なビジュアルフィードバックシステムへの応用が開かれることになる。特に、ロボットの制御の場合、このようなビジョンチップを用いた画像処理の速度は従来のシステムををはるかに超えるものとなり、エンコーダやポテンショメータと同等な速さでセンサ情報を得ることが可能となるので、作業領域における対象や自分自身の動きを画像から認識し、それを実時間のサーボループに直接組み入れることができるようになる。

このことを踏まえて、本論文では、将来の超並列・超高速ビジョンチップに対する応用システムとして、我々が開発した 1ms ターゲットトラッキングシステム [5] を位置づける。すなわち、画像処理部として SPE を用い、センサと処理部を 1 対 1 に直結することで、センサ情報に対する完全並列な処理を実現した。その結果、1[ms] という従来不可能であった高速なサイクルタイムでのビジュアルフィードバックを可能にし、実際にこれを用いたターゲットトラッキングを行ってその有効性を確認した。以下では、このビジュアルフィードバックシステムの構成と基本的な性能を示し、このシステムを用いたターゲットトラッキングの実験結果を示す。

2. 1ms ターゲットトラッキングシステムの構成

我々が開発した 1ms ターゲットトラッキングシステムは、Fig. 2 に示すように、画像を入力して並列演算処理を行う SPE-256 システムと、その処理結果に基づいて視線を制御するアクティブビジョンシステムからなり、全体として 1[ms] のサイクルタイムでのビジュアルサーボを構成している。以下の節ではこれらのシステムを順に説明していく。

2.1 SPE-256 システム

並列演算処理部となる SPE は SIMD 型の制御によりビットシリアルに演算が行われる。演算ユニット機構としては、AND, OR, EX-OR, 加減算, 乗算が可能で、1PE あたり 24[bit] のレジスタを持っている。また各 PE はメッシュ状に配置され、それぞれ 4 近傍との入出力、センサ入力と処理結果の並列出力が可能である。この SPE の内部構造を Fig. 3 に示す。

このアーキテクチャでは集積化のために 1PE あたりのゲート数を大幅に削減し、337 ゲートというコンパクトな回路が実現されており、現在 8 個の PE が集積化された LSI として SPE-8 が試作されている。将来はフォトディテクタなども一体化し

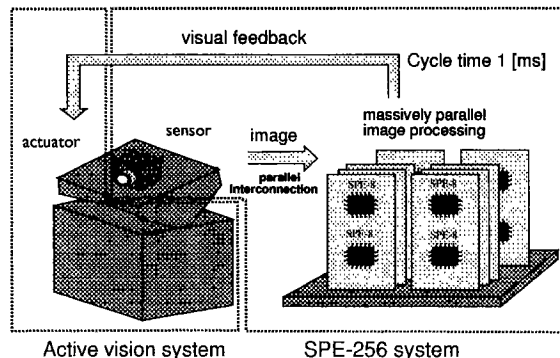


Fig. 2 Architecture of 1ms target tracking system

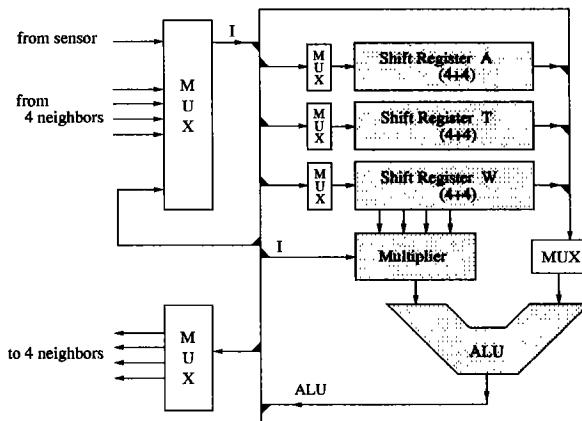


Fig. 3 Block diagram of SPE

た上で、さらに集積化を進めることが予定されているが [6]、本論文では将来のシステムのスケールアップモデルとして、この SPE-8 を 32 個用いて SPE-256 を試作した。

SPE-256 では $16 \times 16 = 256$ 個の PE をマトリクス状に配置し、これに対応した画像センサには浜松フォトニクス社製の 16×16 素子 PIN フォトダイオード (PD) アレイ S3805 (放射感度 $0.5 [A/W]$ at $800 [nm]$, 応答速度 $7 [ns]$) を用いた。この PD アレイの出力は全画素並列に取り出すことができるので、それぞれの PD 出力を PE と 1 対 1 に結線し、入力画像の各画素に対して完全並列な処理を実現している。また、各 PE に送られた PD 出力はまず 2 値化され、それに対して種々の画像処理を行っている。

SPE ヘインストラクションを送る方法としては、専用のシーケンサなどを用いて高速化を図るべきであるが、今後、ロボット制御システムへと拡張することを考慮して Texas Instruments 社の並列処理用 DSP, TMS320C40 (256 MOPS, $320 [Mbyte/sec]$, $20 [Mbyte/sec]$ の非同期データ転送) をホストとして用い、パラレル IO ポート (PIO) を介して直接インストラクションを送り出している。このため、現在のシステムでは PIO の動作速度でサイクルタイムが制限されており、本来 SPE-8 のインストラクションサイクルが、最高 $100 [ns]$ であるのに対して、サイクルタイム $12.5 [\mu s]$ での動作となっている。

SPE では、ALU やレジスタを利用したマイクロプログラミングを行うことにより、得られた画像に対して様々な画像処理を行うことができる。Table 1 に、一般的な画像処理を行った

Table 1 Execution time for algorithms on SPE-256

Algorithms	Steps	Execution time		
		present	max	
Edge	2-neighbor	23	287.5 [μ s]	2.3 [μ s]
	4-neighbor	33	412.5 [μ s]	3.3 [μ s]
Skeleton	4-neighbor	149	18.6 [ms]	149 [μ s]
	8-neighbor	360	45 [ms]	360 [μ s]
Motion detection	7		87.5 [μ s]	0.7 [μ s]
Trace	4		50 [μ s]	0.4 [μ s]
Poisson eq.	125		312.5 [ms]	2.5 [ms]
Smoothing	51		637.5 [μ s]	5.1 [μ s]
Matched filter	248		3.1 [ms]	24.8 [μ s]
Laplacian filter	84		1.05 [ms]	8.4 [μ s]
Optical flow	672		1.68 [s]	13.44 [ms]

ときの実行時間を示す。ただし、細線化、ポアソン方程式、オプティカルフローでは取束演算を行っている。

PE から出力される処理結果はパターンとして得られるので、ここでは、そのパターンの重心や総和を求めたものを画像特徴量とし、ビジュアルフィードバックに利用する。これまでに、このようなパターン情報の重心や総和を並列かつ高速に求めるものとして、抵抗ネットワーク回路などが提案されているが[7][8]、今回はホストコンピュータとして用いた DSP, TMS320C40 を利用し、逐次処理によって重心を計算することにした。具体的には PE アレイの一边と PIO を接続し、そこから SPE で並列処理した処理結果を DSP へ送り出している。しかし、実際にはこの処理結果の排出のために 1 サイクルごとに 500 [μ s] 程度の時間を消費しており、システム全体のサイクルタイムである 1 [ms] の半分を占めている。このことから、集積化ビジョンチップの処理の並列性を生かすためには、重心や総和などの出力計算を並列に行う機構をあらかじめ同じチップに組み込んでおくことが必要になる [9]。

2.2 アクティブビジョンシステム

アクティブビジョンシステムの可動部には PD アレイとレンズが載せられており、直交するチルト、パンの 2 軸についての回転自由度をもっている。それぞれの軸のアクチュエータにはハーモニックドライブシステムズ社製 DC サーボアクチュエータ RH-8 (チルト), RH-11 (パン) を用いており、これらのサーボとしてサーボランド社製アナログサーボコントローラ SMCM2-AI (ステップ応答 0.2 [ms] 以下) を用いて速度制御を行っている。したがってこのサーボに適切な速度指令値を与えてやれば、ビジョンシステムの視線方向を自由に変わることができる。

前述したように、PD で取り込まれた画像は各画素並列に SPE-256 へ送られ、2 値化された上で与えられたプログラムに従って画像処理が行われる。その処理結果は DSP へ送られ、重心を計算して画像特徴量としている。さらに DSP では、その画像特徴量の目標値からの偏差から、PID 制御によりアクチュエータへの指令を計算している。アクチュエータのサーボはこの DSP で計算された指令値に従って速度制御を行っている。

これら一連の処理によって、画像情報をフィードバックして視線を制御するビジュアルサーボを構成しており、そのサイク

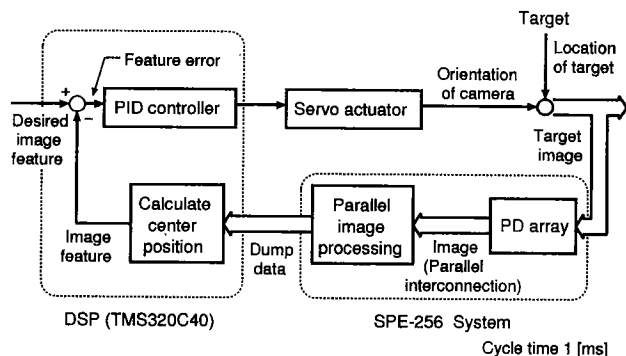


Fig. 4 Block diagram of 1 ms target tracking system

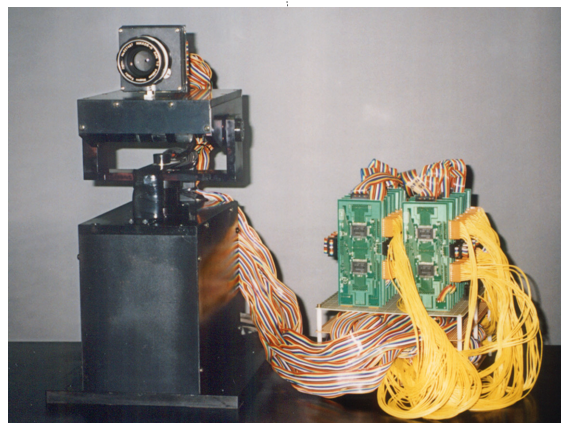


Fig. 5 Picture of 1 ms target tracking system

ルタイムは 1 [ms] となっている。SPE-256 を含めた全体システムでの処理の流れを Fig. 4 に示す。またシステム全体の写真を Fig. 5 に示す。

3. 実験結果

3.1 アクティブビジョンの周波数応答特性

まず、アクティブビジョンシステムの基本的な動特性を評価するために、サーボループからビジュアルフィードバックの部分を除いた開ループでの周波数応答を測定した。このなかにはアクチュエータやサーボコントローラ、視線の方向を変えるための機構部分の動特性が含まれている。これを最初に測定しておくことで、ビジュアルフィードバックをサーボループに組み入れたときに、そのシステムの性能がどの部分で制限されているかを評価することができる。

測定結果を Fig. 6 に示す。アクチュエータに与えた速度指令に対して、実際に得られた速度のゲインと位相遅れを示している。破線がパン方向の回転軸、実線がチルト方向の回転軸の応答を表している。ゲインのグラフ中に点線で引いた -3 [dB] の線とグラフの交点から、カットオフ周波数は約 20 [Hz] (時定数: $T_m = 50$ [ms]) であることが分かる。以下では、この測定で得られたパラメータをもとに、デジタルサーボを組む際にどの程度のサンプリングレートでビジュアルフィードバックを行えばよいかを考察する。

3.2 ビジュアルサーボにおけるサンプリング間隔の影響

一般に、デジタルサーボでは、サンプリング間隔を短くすれ

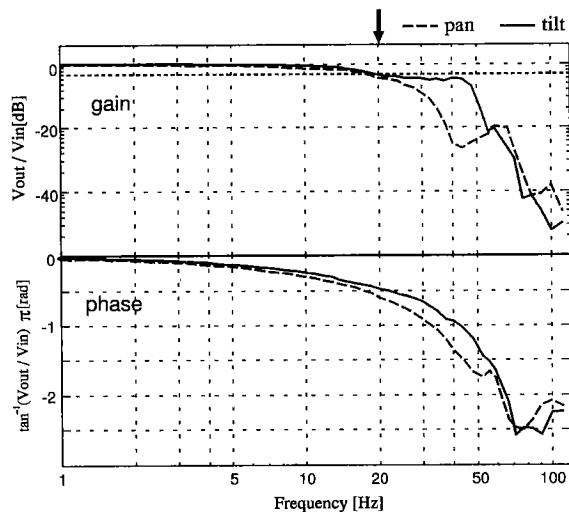


Fig. 6 Frequency response of active vision

ばするほどサーボの安定性と速応性が向上するということが広く言われている。例えば従来のビデオ信号を用いたビジュアルサーボのサンプリングレートとしては、ビデオ信号をそのまま用いて 30 [Hz] のビデオフレームレートでサーボを構成する場合と、フィールドレートを利用して 60 [Hz] のサンプリングレートで制御を行う場合の 2 通りが考えられる。

ところが、ロボットの制御レートについて考えると、ロボットの機械系が有する固有振動数 $1/T_m$ に対して、最も下位の関節サーボのサイクルタイム T_s が満たすべき条件として $T_s < T_m/10$ 程度が必要であると言われている [10]。これをこの実験で用いたアクティブビジョンシステムに適用すると、 $T_m/10 = 5$ [ms] となり、ビデオ信号を用いたシステムでこの条件を満たすためには、補間や予測を用いなければならなくなる。これに対して、この論文の SPE-256 システムではビデオ信号の制限がなく、1 [ms] のサイクルタイムで画像を処理できるので、特別な方法を用いることなくこの条件を満たすことができる。

さらに、一般的なロボットの固有振動数が 20 ~ 50 [Hz] 程度であることから、ロボットのサーボのサイクルタイムとしては 1 ~ 5 [ms] が必要となる [1]。例えば、産業用ロボットの Unimation 製 PUMA560 では、関節ごとのサーボは 1 サイクル 0.875 [ms] で計算を行っており、PID 制御に従ってアクチュエータを制御している [11]。このことから、本論文のシステムを用いて構築したビジュアルサーボは、ロボットのサーボとして十分な速度を有することが分かる。

このようなデジタルサーボにおけるサンプリング間隔の影響を、より一般的に、定量的に調べたものとして、水谷らは、システムの開ループ伝達関数が 1 次遅れ系とみなせるときに、サンプリング間隔がサーボの性能に与える影響を数値計算を用いて考察している [12]。

これに従って本論文のアクティブビジョンシステムを一次系とみなし、系の時定数 $T_m = 50$ [ms] を基準として評価すると、本システムのような 1 [ms] のサンプリング間隔では、ほとんど連続値系とみなして設計できるのに対して、従来のシステムの

ようにビデオレート (33 [ms]) の制限のためにサンプリング間隔が $0.3T_m \sim 0.5T_m = 15 \sim 25$ [ms] 以上になると、もはや連続値系とみなすことはできなくなり、安定性を確保するためにはアクチュエータが本来持っている動作性能を犠牲にし、ゲインを低く設定してサーボ系を設計せざるを得なくなる。

これらのことから分かるように、本システム程度の動作周波数を有するアクティブビジョンシステムに対してビジュアルフィードバックによる制御を行う場合、従来用いられているビデオ信号のサンプリングレートでは不十分であり、速応性を犠牲にして低速なシステムとして設計するか、何らかの予測を用いてフィードバックの遅れを補償するしかない。逆に、本論文のシステムのように、1 [ms] のサンプリング間隔でのビジュアルフィードバックが実現された場合、ロボットなどの制御のためのビジュアルサーボとしてそのまま利用できるだけの十分な高速性を有しているということが言える。

3.3 ターゲットトラッキング

次に、実際にこのビジュアルフィードバックシステムを用いたターゲットトラッキングの実験の結果を示す。この実験では、PD から入力したターゲットの画像を 2 値化し、得られたパターンをそのまま用いてその重心を画像特徴量とし、ビジュアルフィードバックを行った。また、通常のターゲットトラッキングでは、移動するターゲットに対してターゲットの画像の中心が画面の中央に見えるように視線を変化させるが、この実験の場合、追跡対象となるターゲットを高速ビジョンに見合うほど速く、かつ正確に動かすのは難しい。そこで、等価的に同じ状況を作るために、画面のなかでターゲットが見えるべき位置を指定してこれを目標位置とし、ターゲットがその位置にくるように視線を制御することにした。

実験結果を Fig. 7 に示す。目標位置として 5 [Hz] の三角波を与えてこれに追従させた。実線が 1 [ms] のサイクルタイムによるターゲットトラッキング動作を示す。破線は、本システムにおいて動作サイクルをわざと落とし、ビデオレートと同じ 33 [ms] のサイクルタイムでトラッキングを行ったものである。1 [ms] では目標に対して十分追従しているのに対して、33 [ms] では追従性能が著しく悪くなっているのが分かる。

特に 1 [ms] でのトラッキングでは目標に対する遅れは約 20 [ms] であり、センサフィードバックの遅れの影響はほとんどなく、アクチュエータなどの機構部分の動作周波数の限界だけから遅れが生じていると言える。これに対して 33 [ms] のサイクルタイムでは約 70 [ms] の遅れが生じており、オーバーシュートも大きくなっている。これは、アクチュエータの動作周波数の限界だけでなく、センサフィードバックが遅すぎるために、目標の動きの高速な変化に対応できないのだと考えられる。また、一般にセンサフィードバックの遅れを補償する手段として、対象の動きを予測するという方法がとられるが、この実験のように、対象の動き自体が予測不可能な動きをする場合には、その方法は有効ではなくなると考えられる。これに対して 1 [ms] という高速なセンサフィードバックが実現された場合には、このような対象の動きに対しても、簡単な制御で十分な追従性能が得られることが分かった。

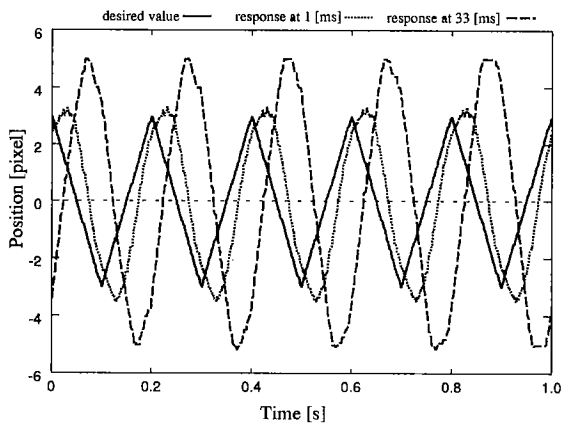


Fig. 7 Tracking response to a 5 [Hz] triangular wave (pan)

4. おわりに

汎用の PE を用いた完全並列処理によるビジョンシステムを実現し、ビジュアルフィードバックによるターゲットトラッキングを行った。その結果、センサ情報を並列に処理することで 1 [ms] という高速なビジュアルフィードバックが可能となり、これを用いて高速なターゲットトラッキングが実現できることを示した。

今後の課題として、実際にセンサと処理部を一体化し、それらをワンチップに収めた集積化ビジョンチップの実現が挙げられる。これによって、解像度や、処理部のクロックサイクルを大幅に引き上げることが可能となり、振動制御などの超高速ビジュアルフィードバックに応用できるようになる。また、システムの小型化によって、このようなワンチップビジョンシステムを多数個用いた分散型のセンシングなども容易に行えるようになるほか、機構部分の動特性が改善されることによって、より高速なアクチュエータと組み合わせて、今まで不可能とされてきたような超高速ターゲットトラッキングが実現できる。

また SPE の能力を利用すれば、本システムに対してより複雑なアルゴリズムを適用し、対象物を認識しながらのターゲットトラッキングを 1 [ms] のサイクルタイムで実現することも可能となる。その場合、対象物は 1 [ms] の間でほとんど動かないため、

従来の画像処理アルゴリズムを単純化することが可能であり [13]、制御の単純化と合わせて高速化の利点になると考えられる。さらに、これによって、現在ロボットの姿勢角センサとして用いられているポテンシオメータやエンコーダと同等なサンプリングレートで、画像情報に基づいた実時間のサーボループを構成することができるようになり、これをロボットの制御に用いて動的な環境に実時間で適応できるシステムを構築したり、実際のタスクを行う領域での制御精度を向上させることができるようになる。

このように、高速な視覚情報処理が求められるような様々なシステムに対して、超並列・超高速ビジョンシステムが幅広く応用できるものと期待される。

参考文献

- [1] ロボット学会編：ロボット工学ハンドブック、V.2.2 ソフトウェアサブの実現。pp.532-534, コロナ社, 1990.
- [2] C. Mead and M. Mahowald: "A Silicon Model of Early Visual Processing," *Neural Networks*, vol.1, pp.91-97, 1988.
- [3] 石川正俊：“超高速・超並列ワンチップビジョンとその応用”，日本ロボット学会誌，vol.13, no.3, pp.335-338, 1995.
- [4] M. Ishikawa, A. Morita and N. Takayanagi: "High speed vision system using massively parallel processing," In Proc. IEEE Int. Conf. on Intelligent Robotics and Systems, pp.373-377, 1992.
- [5] 中坊嘉宏, 石井抱, 石川正俊：“並列ビジョンシステムを用いた高速ターゲットトラッキング”，SICE'95 予稿集，pp.21-22, 1995.
- [6] 小室孝, 鈴木伸介, 石川正俊：“超並列ビジョンチップアーキテクチャ”，電子情報通信学会技術報告，CPSY95-19, pp.63-69, 1995.
- [7] A. Utsugi and M. Ishikawa: "Construction of inner space representation of latticed network circuits by learning," *Neural Networks*, vol.4, no.1, pp.81-87, 1991.
- [8] 向井利春, 石川正俊：“並列ビジョンのための 2 次元座標変換回路”，電子情報通信学会技術報告，PRU92-72, pp.111-116, 1992.
- [9] 坂口隆明, 小室孝, 石井抱, 石川正俊：“ビジョンチップのためのモータメント出力回路”，SICE'96 予稿集，pp.829-830, 1996.
- [10] R.P. Paul: *Robot Manipulators*. MIT Press, 1987.
- [11] J.J. Craig: *Introduction to ROBOTICS—mechanics and control*, Addison-Wesley, 1989.
- [12] 水谷隆, 岡本清和, 大庭信男, 張玉武：“デジタルサーボ系におけるサンプリング間隔の影響に関する基礎的考察”，精密工学会誌，vol.56, no.6, pp.142-146, 1990.
- [13] I. Ishii, Y. Nakabo and M. Ishikawa: "Target tracking algorithm for 1 ms visual feedback system using massively parallel processing vision," In Proc. IEEE Int. Conf. on Robotics and Automation, pp.2309-2314, 1996.



中坊嘉宏 (Yoshihiro Nakabo)

1972年2月22日生。1995年、東京大学工学部卒業。現在、同大学院修士課程在学中。超並列・超高速ビジョン、ロボット制御に興味を持つ。計測自動制御学会会員。（日本ロボット学会学生会員）



石井抱 (Idaku Ishii)

1970年2月14日生。1994年、東京大学大学院修士課程修了。1996年、東京大学大学院博士課程中退。同年、同大学院工学系研究科計数工学専攻助手。超並列超高速ビジョン、人工現実感の研究に従事。日本バーチャルリアリティ学会会員。計測自動制御学会会員。（日本ロボット学会正会員）



石川正俊 (Masatoshi Ishikawa)

1954年8月22日生。1979年、東京大学大学院修士課程修了。同年工技院製品科学研究所、1989年、東京大学大学院工学系研究科計数工学専攻助教授。生体計測、触覚センサ、超並列・超高速ビジョン、光コンピューティング、センサフュージョンの研究に従事。工学博士。1984年計測自動制御学会論文賞、1988年度工業技術院長賞、1989年応用物理学会光学論文賞受賞。（日本ロボット学会正会員）