

ビジョンチップシステムを用いた高速視覚処理

High Speed Visual Processing Using a Vision Chip System

小室 孝† 鏡 慎吾† 渡辺 義浩† 竹内 大介† 神明前 方嗣† 石川 正俊†
Takashi Komuro†, Shingo Kagami†, Yoshihiro Watanaeb†, Daisuke Takeuchi†,
Masatsugu Shinmeimae†, Masatoshi Ishikawa†

† 東京大学情報理工学系研究科

† School of Information Science and Technology, the University of Tokyo

E-mail: Takashi_Komuro@ipc.i.u-tokyo.ac.jp

Abstract

This paper describes a vision chip which has a processing element in each pixel and that performs high frame rate visual sensing. A real-time visual processing system using the vision chip, and various visual processing which is realized by the system are also shown.

1 はじめに

コンピュータやロボット、家電などの知能化に伴い、リアルタイムの視覚センシングが重要な技術になってきている。イメージセンサと逐次処理プロセッサからなる従来型のシステムでは、リアルタイム処理に必要な演算を行うために高クロック動作を必要とし、機器の大型化、高消費電力化、高価格化を招いていた。また、画像の転送のために、フレームレートが固定されているかまたは上限が定められていた。

それに対し、図1に示すようなイメージセンサの画素毎に演算処理回路(PE)を備えたビジョンチップでは、画像の転送が不要な上、画素並列に画像処理を行えることから、高フレームレート(100~1000Hz)の視覚処理を、小型、低消費電力、低価格な装置で実行することができる。

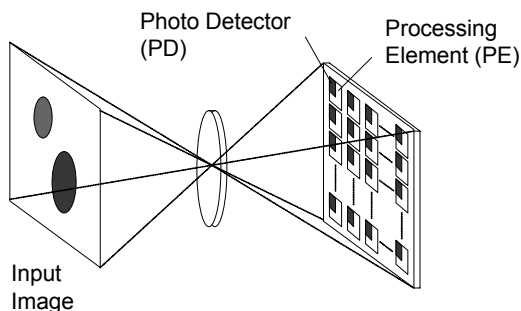


図1 ビジョンチップの概念図

本稿では、我々が開発したビジョンチップシステムとそれを用いた高速視覚処理の例を示す。

2 ビジョンチップシステム

2.1 プログラマブルビジョンチップ

従来研究されてきたビジョンチップの多くは、特定の演算機能しか持たないものであったが、我々は広い用途に利用できるように、汎用の PE を採用したプログラマブルなビジョンチップを開発した[1]。図2に開発したビジョンチップのアーキテクチャを示す。

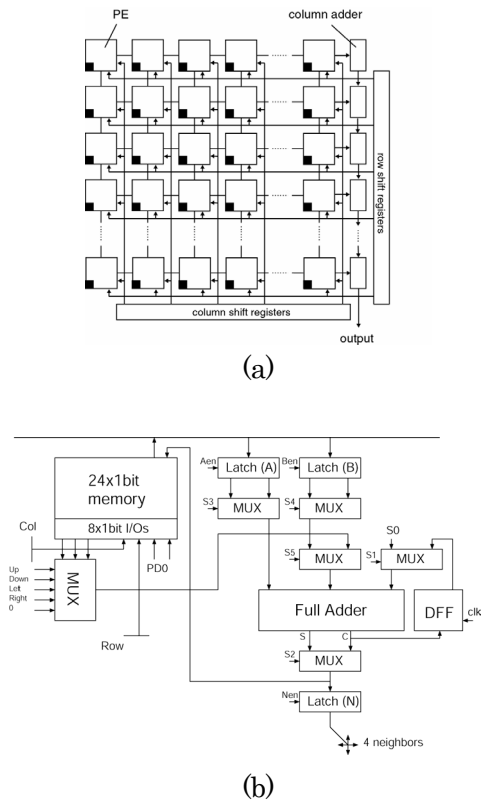


図2 開発したビジョンチップのアーキテクチャ:
(a) 全体 (b) PE

各 PE は 1 ビット ALU と 24 ビットのメモリを持ち、演算処理をビットシリアルに実行する。SIMD 型の並列処理により、全面素に共通な処理を画素並列に実行する。また、複数の PE を連結して総和などのグローバル演算を効率的に実行することも可能となっている。

開発したビジョンチップのレイアウトを図 3 に示す。このチップは 0.35 μm CMOS プロセスを使用し、5.4mm 角に 64 \times 64 画素が搭載されている。

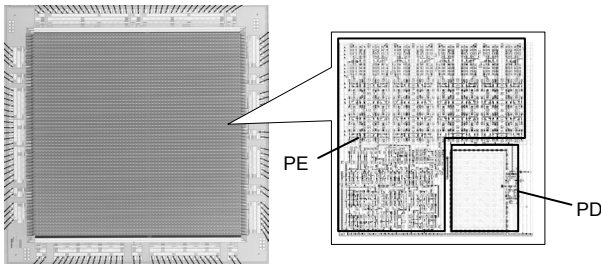
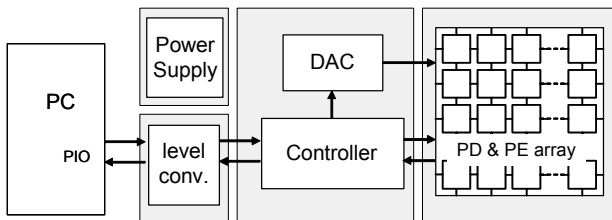


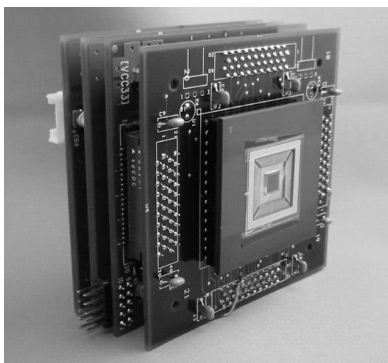
図 3 開発したビジョンチップのレイアウト

2.2 ビジョンチップシステム VCS-IV

上記ビジョンチップとコントローラ、通信インターフェースなどを搭載したシステムボード(VCS-IV)を開発した[2]。コントローラは、命令コードに基づいてビジョンチップの制御信号を生成するほか、プログラム制御や非並列の演算命令を行うことができる。同システムは、10MHz の命令レートで動作する。



(a)



(b)

図 4 ビジョンチップシステム VCS-IV:
(a) 構成 (b) 外観

また、VCS-IV のプログラミング環境として、専用のコンパイラを開発した[3]。同コンパイラは、ビットレベルの最適化により、メモリ消費の少ないコードを出力するようになっている。

3 高速視覚処理の例

3.1 初期視覚処理

ビジョンチップは画素並列に演算処理を行うので、すべての画素に対して共通の演算を行うような場合に高い効果を発揮する。特に、ある画素における演算結果がその画素自身とその近傍画素の値から決定されるようなアルゴリズムは極めて高速に実行できる。例として、エッジ検出・スムージングなどのフィルタ処理や、収縮・膨張などのモルフォロジー処理が挙げられる。

動画像処理については、ビジョンチップに向けた処理の一つにトラッキングがある。これは視野内の特定の対象を継続的に追跡する処理であり、高フレームレート下においてフレーム間の画像変動が少ないことを利用することで、簡単なモルフォロジー処理のみで実現される。

画像全体からグローバルな特徴量を計算する場合も、前述の PE を連結する機能を用いることで、画面全体の総和や総論理和などを高速に計算することができる。

図 5 に初期視覚処理の例と実行時間を示す。

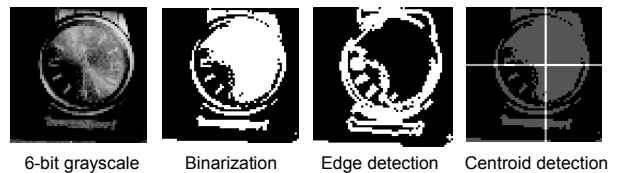


図 5 初期視覚処理の例

表 1 初期視覚処理の実行時間

Program	steps	time [μs]
Dilation, erosion (1-bit)	20	2.0
Edge detection (1-bit)	31	3.1
Edge detection (6-bit)	193	19
Smoothing (1-bit)	47	4.7
Smoothing (6-bit)	154	15
Target tracking (1-bit)	34	3.4
0 th moment (1-bit)	124	12
1 st moment (1-bit)	226	23
Centroid detect. (1-bit)	577	58

3.2 マルチトラッキング

画面内に対象が複数ある場合でも対象毎にウィンドウを持つことでそれぞれの対象をトラッキングすることが可能である。また、このマルチトラッキングの前処理として、二分探索によるラベリングアルゴリズムも提案している[4]。

これらの応用として、ライン上の対象物の計数処理

と球体の回転運動計測を行った[5]。

計数処理では、ビジョンチップは視野内の複数対象を追跡し、累積数をカウントする。500Hzのフレームレートで、1.0cmの対象を5.5m/s(20km/h)の速度まで計測できる。図6に実験の様子を示す。

回転計測では、ビジョンチップは球の表面に描かれた点模様を追跡し、軸の方向と回転速度を計算する。1kHzのフレームレートで、分速1200回転まで計測できる。これは、サッカーボールの最大回転速度(分速600回転)と野球のボールの最大回転速度(分速1800回転)の間である。図7に実験の様子を示す。

これらの視覚計測応用は、工場内の自動検査等に有効である。

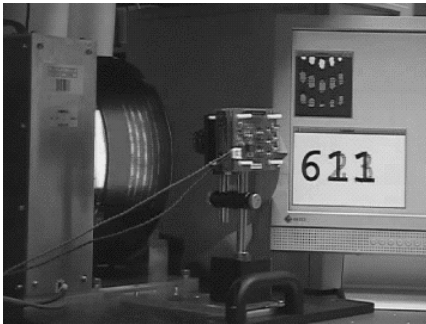


図6 計数処理実験

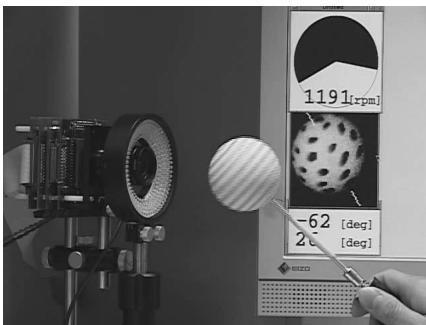


図7 回転計測実験

3.3 センサ特性のソフトウェア制御

ビジョンチップは画素内に処理回路を持っているため、撮像した画像に対して演算処理を行うだけでなく、撮像処理そのものをコントロールすることが可能となる。

開発したビジョンチップでは図8のように光検出器からの出力信号のAD変換の際にPEをカウンタとして用いて、キャパシタに蓄えられた電荷が光電流によって放電され、低下した電圧が外部から与えられた参照電圧を下回るまでの時間を計測している。

このとき図9のように比較器に与える参照電圧やカウントのタイミングを動的に変えることで、変換特性を変えることができる[6]。これによって、同一センサでリニアスケールや対数スケールなどさまざまな特性

を実現することができる。

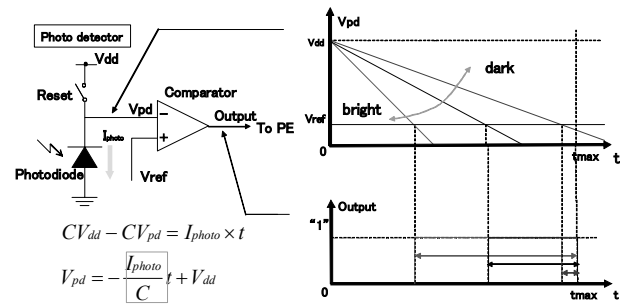


図8 AD変換の原理

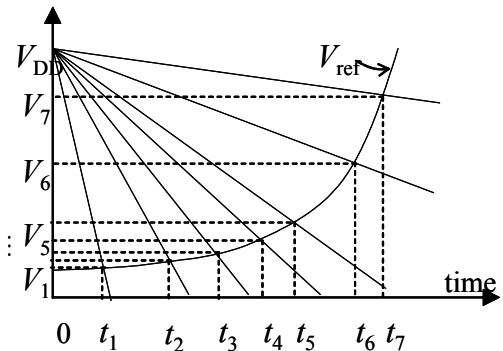


図9 AD変換特性のコントロール

また、比較器のオフセットばらつきに起因する固定パターンノイズ(FPN)の除去もソフトウェアで行うことができる。図10に手法を示す。参照電圧を二回同じ傾きで上昇させ、交差する二点間の時間を計測することで、光強度を測定する。

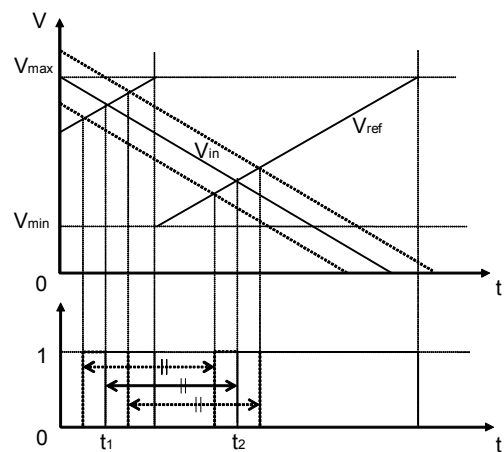


図10 ソフトウェアによる FPN 除去

ただし、この方法を用いる場合、比較器の入力が電源電圧に近いところでは正しく動作しないため、リセット電圧を可変にできるチップを新しく作成した。このチップを用いて、さまざまなAD変換特性で撮像した画像を図11に示す。

さらに、AD変換特性とRMS(Root Mean Square)

ノイズを計測した。RMS ノイズは、出力値を照度で正規化したものを用いて計算した。チップ表面をレンズを付けずに直接 LED 照明で均一に光を照射した。露光時間は 1.0ms とした。結果を図 12、図 13 に示す。FPN 除去を行うことで、除去前に比べ 3.3 倍の感度向上が実現された。

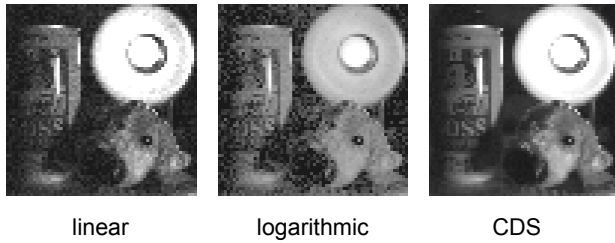


図 11 各種 AD 変換特性による撮像例

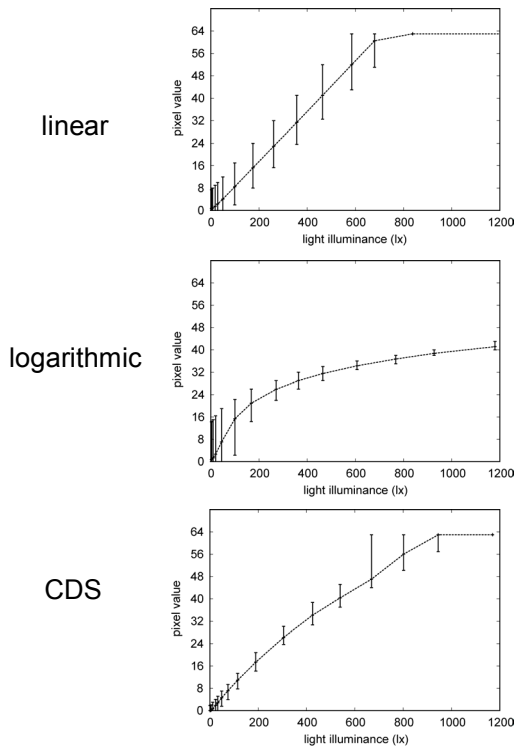


図 12 AD 変換特性の測定結果

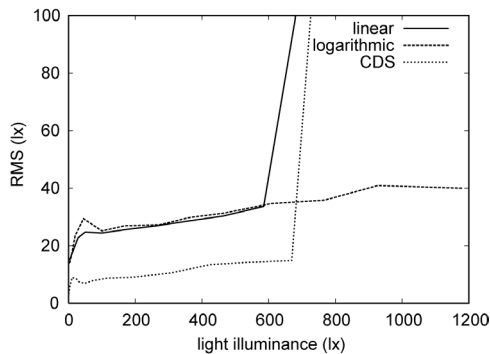


図 13 RMS ノイズの測定結果

3.4 モデルベース形状認識

ビジョンチップが持つ画素並列処理機能と高速なグローバル演算機能を利用することで、画像空間でのマッチングを高速に実行することが可能である。たとえば、入力画像を I 、候補として生成するモデル画像を M とすると、画像の差分絶対値総和(SAD)は以下の式で表される。

$$S(I, M) = \sum_x \sum_y |I(x, y) - M(x, y)|$$

このうち、差分絶対値は画素並列に、総和は PE 連結によるグローバル演算機能を用いて高速に実行できる。モデル画像の生成も座標値に基づいて画素並列に行うことができる。

取りうるすべての候補についてマッチングを行おうとすると計算量が爆発的になるので、モーメント特徴量を元に初期値を定め、収束解法を用いて探索を行う [7]。モーメント特徴量からモデル画像のパラメータを解析的に計算するのは逆問題となり困難なので、あらかじめ用意したテーブルを利用する。

さらに、動画画像の場合は前フレームの結果を利用することで、探索範囲を狭められる。ただし、いったん誤認識しても回復が可能なように、前フレームで推定したパラメータと前述のモーメント特徴量を用いて取得したパラメータの良いほうを初期値とする。図 14 にアルゴリズムを示す。

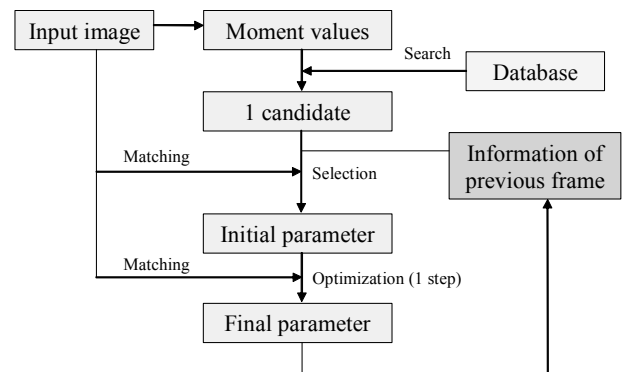


図 14 形状認識アルゴリズム

このアルゴリズムをビジョンチップに実装し、実験を行った。モデルには、三角形、平行四辺形、楕円の三種類を用意し、入力画像から図形の種類とパラメータを推定させる。収束演算ではパラメータのみを変化させ、図形の種類はモーメント特徴量からの初期値決定時のみ決められる。

図 15 に楕円を入力として与えた場合の実験結果を示す。最初は図形の種類を誤って識別しているが、いったん正しく識別すると、その後は安定して識別し続けていることがわかる。

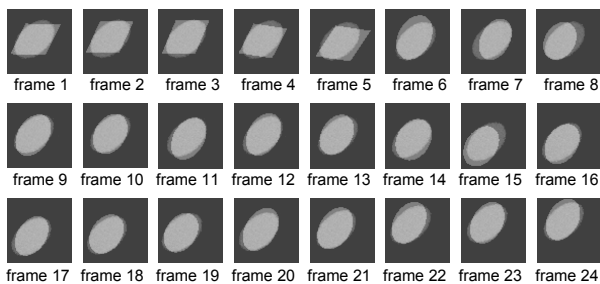


図 15 実験結果

フレーム毎の計算時間は約 25ms かかっており、ビデオレートのリアルタイム性は確保しているものの、ビジョンチップの高フレームレートを生かし切れてはいない結果となった。しかし、ハードウェアの改良などによる高速化の余地は十分残されている。

3.5 時間符号化光の検出・追跡

時間符号化光とは、LED 光源やレーザポインタのスポット像などを時間変調して情報を重畳したものであり、これを検出・追跡することで位置特定と対象識別が同時に可能となる。

時間符号化光をイメージセンサで検出する研究は既に多くの例があるが、高速に移動する光信号に対応したものは少ない。ビジョンチップを用いることで、光信号を追跡しながら十分な長さの信号を受信することが可能となり、より広いアプリケーションに利用できるようになる[8]。

図 16 に追跡処理の流れを示す。複数の同期されていない光信号を同時に認識できるように、各画素にはステートマシンが配置されている。信号が検出され、ステートが進行中の画素は、自己のステートを隣接する画素にコピーする。次のフレームで光信号の像が動いたとしても、コピーした範囲の画素ではステートマシンが矛盾なく動作し続ける。コピーした先の画素に光信号源の像が移動していなければ、ステートは進行せずリセットされる。これにより、追跡しながらの受信が可能となる。

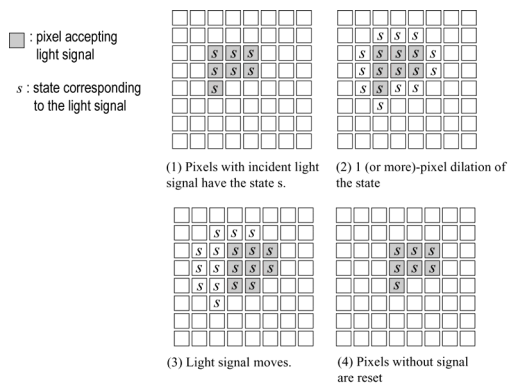


図 16 追跡処理の流れ

高速に移動する光信号への追従性能を評価するため、追跡処理を行った場合と行わなかった場合についてそれぞれ実験を行った。観測された光信号の画像上での座標をプロットしたものを 図 17 に示す。直線部などの高速に運動している部分では、追跡処理がない場合には検出できなかったが、追跡処理を行った場合は常に検出が可能であった。

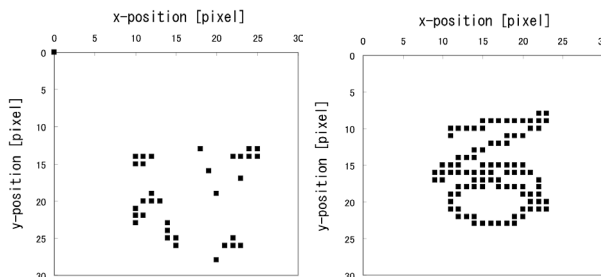


図 17 追跡実験結果

(左: 追跡処理なし, 右: 追跡処理あり)

さらに、複数の光信号の追跡や、衝突時の分離処理についても実験を行い、期待通りの結果を得た。

本成果は、光学マーカによるジェスチャ認識やカメラキャリブレーションなどに応用できる可能性を持っている。

4 おわりに

汎用の PE を採用したプログラマブルなビジョンチップとそれを用いた実時間処理システムを紹介し、そのシステムで動作する高速視覚処理の例を示した。これらの例でわかるように、プログラマブルな構成を取ることにより、さまざまな視覚処理を同一チップで実現することができる。従来のカメラとプロセッサを組み合わせたシステムに比べ、視覚センシングに特化しながらも汎用性を備えたビジョンチップは、近未来の新産業における標準デバイスとなる可能性を秘めている。

参考文献

[1] 小室 孝, 鏡 慎吾, 石川 正俊: ビジョンチップのための動的再構成可能な SIMD プロセッサ, 電子情報通信学会論文誌 D-II, Vol. J86-D-II, No. 11, pp.1575-1585 (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, pp.134-142 (2005)

[3] Takashi Komuro, Shingo Kagami, Masatoshi Ishikawa, Yoshio Katayama: Development of a Bit-level Compiler for Massively Parallel Vision Chips, Proc. IEEE International Workshop on

Computer Architecture for Machine Perception (CAMP'05) (2005, to appear)

[4] 渡辺 義浩, 小室 孝, 鏡 慎吾, 石川 正俊: ビジョンチップのためのマルチターゲットトラッキングとその応用, 電子情報通信学会論文誌 D-II, Vol.J86-D-II, No.10, pp.1411-1419 (2003)

[5] Yoshihiro Watanabe, Takashi Komuro, Shingo Kagami, Masatoshi Ishikawa: Multi-target Tracking using a Vision Chip and its Applications to Real-time Visual Measurements, Vol. 17, No. 2 (2005, to appear)

[6] 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)

[7] Takashi Komuro, Yoshiki Senjo, Kiyohiro Sogen, Shingo Kagami, Masatoshi Ishikawa: Real-time Shape Recognition Using a Pixel-parallel Processor, Journal of Robotics and Mechatronics, Vol. 17, No. 4 (2005, to appear)

[8] Shingo Kagami, Masatsugu Shinmeimae, Takashi Komuro, Yoshihiro Watanabe, Masatoshi Ishikawa: A Pixel-Parallel Algorithm to Detect and Track Fastly Moving Modulated Light Signals, Journal of Robotics and Mechatronics, Vol. 17, No. 4 (2005, to appear)