

高速ビジョンチップを用いたビジュアルサーボ基礎実験

石原聡之 辻田哲平 出口淳 杉村武昭 近野敦 栗野浩之 内山勝 小柳光正 (東北大学)

Preliminary Experiments of Visual Servoing Using High Speed Vision Chip

*T. Ishihara, T. Tsujita, J. Deguchi, T. Sugimura, A. Konno, H. Kurino,
M. Uchiyama and M. Koyanagi (Tohoku University)

Abstract— In order for mobile robots to walk around in a dynamic environment, the robots must detect changes in the environment in real time. Vision will be the most significant information to detect the changes. When the mobile robots must react with a rapidly changing environment, a fast image processing system like a human brain is needed. Against such a background, Koyanagi et al. developed a newromorphic vision chip imitating the human retina. In this work, visual servoing experiments were carried out to evaluate the performance of the vision chip. A parallel link camera head is used in the experiments to achieve the quick eye motions. This paper describes the detail of the vision chip, the camera head and the experimentation result.

Key Words: Vision chip, Parallel mechanism

1 はじめに

移動ロボットには周囲の環境を認識し、状況に応じて作業を行うことが求められている。そのためには、高速な視覚処理が欠かせない。例えば人間においては外界情報のほとんどが視覚を通して得られていると言われている。移動ロボットにも人間同様に視覚により外界の情報を得て、状況を認識し行動することが求められる [1, 2]。例えば災害現場などの複雑環境では、周りの状況をすばやく認識し、行動することが求められる。そのため高速に視覚情報を処理することで状況認識能力を向上させる必要がある。現在多くの研究において高速な画像処理を用いたロボットによる状況認識が試みられている。しかし未だ人間のように多種多様な環境において柔軟に対応することは達成されていない。本研究では、高速に画像処理を行うことだけでなく、人間の眼球を模擬することを目的としている。そこで我々は人間の網膜機能による情報処理を模擬することと、人間の眼球運動を模擬することの二つの面から研究を進めている。

より高速な画像処理システムの構築のため、光センサの情報を同じチップ上で処理した上で、メインコンピュータへデータを送る手法が提案されている。そのことによりピクセルごとに並列演算処理が行われるため、高速に画像処理が行える。しかしこのようなビジョンチップでは高機能な画像処理を実現しようとすると光センサに比べ、処理回路が相対的に大きくなる。その結果光センサの開口径の低下、ひいては集光量、解像度の低下を招くことになり、高機能化と高開口率化はトレードオフの関係にある。

小柳らは三次元集積技術を用い、複数層から成る三次元ビジョンチップの開発を行った [3]。小柳らが開発した三次元ビジョンチップは人間の網膜機能を模擬したもので、最上層が細胞を模擬した光センサ、二層目が水平細胞と双極細胞を模擬した画像処理回路、三層目が神経節細胞を模擬したデータ出力回路になっている。このように積層化することで、光センサの開口径を大きく保ったまま、大規模、高性能な処理回路が実現できる。

さらにこのビジョンチップを高速に動かすために、本研究ではカメラヘッドをパラレルリンク機構を用いて構成している。現在のカメラヘッドはシリアルリンク機構のものが多く、しかしパラレルリンク機構を用いたことにより、シリアルリンク機構と比較して高速かつ軽量のカメラヘッドが実現が可能となった。本研究では、パラレルリンク機構を用いることで高速応答性を実現し、これにより高速な人間の眼球運動をも模擬することを目指している。

これら高速ビジョンチップと高速カメラヘッドを用いる研究は、単に移動ロボットの視覚システムだけではなく、人間の眼球の研究さらには視覚における人間の脳への研究へとつながるものと考えられる。

2 視覚システム

現在開発を行っている視覚システムは、移動ロボット用視覚システムといった目的に加え、人間の視覚システムの模擬といった目的を持つ。そこで本研究では人間の網膜を模擬したビジョンチップの開発および、人間の眼球の動きを模擬したカメラヘッドの開発という二つの方向から開発を進めている。

2.1 三次元集積化ビジョンチップ

小柳らが研究を行っている三次元集積化ビジョンチップの試作チップの仕様を Table 1 に、カメラに使用したレンズの仕様を Table 2 に示す [4]。

またビジョンチップカメラの外観を Fig. 1 に、外形寸法を Fig. 2 に示す。試作チップは、 64×64 [pixel] のフォトセンサアレイ、ラインセクタ、制御回路、A/D 変換器、出力エンコーダの五つの部分からなり、フロントセンサから出力された画像を 1 [pixel] ごと A/D 変換器でデジタルデータに変換し、システムクロックと同期して出力する。1 フレーム 64×64 [pixel] で、1 [pixel] あたり 6 [bit] の画像データが 8 [ms] の周期で出力される。今回のビジョンチップはまだ試作品であり、二層目、三層目の画像処理回路は組み込まれていない。そのためビジョンチップからは座標 (0,0) から



Fig. 1: An overview of the prototype camera.

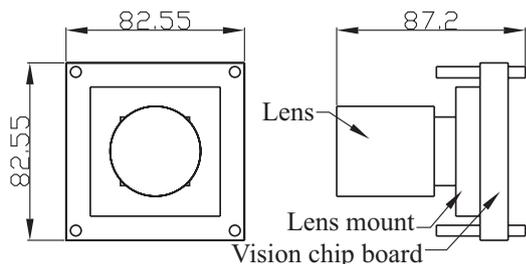


Fig. 2: Dimensions of the camera.

Table 1 Specification of prototype the vision chip

Technology	0.35 [μm] CMOS
Chip size	4.9 [mm] \times 4.9 [mm]
Photocell array	64 \times 64 [pixel]
A/D converter	6 [bit]
Operating frequency	500 [kHz]
Power supply voltage	3.3 [V]

Table 2 Specification of the camera lens

Focal length	25 [mm]
Aqueeze	F0.95 ~ 16
Horizontal angle of field	11 [$^\circ$]
Operating distance	0.5 ~ ∞ [mm]
Size (diameter \times length)	41.8 \times 45.2 [mm]
Weight	120 [g]

(63,63) までの画像データがシリアルに出力される。

2.2 ステレオカメラヘッド

Fig. 3 および, Fig. 4 に本研究で開発中のカメラヘッドを示す. 本研究のカメラヘッドは移動ロボットへ搭載し, 人間の眼球運動を模擬することを目的としている. 移動ロボットへ搭載することを考えるとカメラヘッドの重量を抑える必要がある. また人間の眼球運動を模擬する場合, 眼球運動において最高角速度であるサッカドと呼ばれる $600 [^\circ/\text{s}]$ の跳躍的眼球運動や, トレモアと呼ばれる $30\sim 100 [\text{Hz}]$ 程度の微小振動を模擬する必要がある. そこで本研究では, カメラヘッドに平行リンク機構を用いて構成している. そのためすべてのモータを台座に配置することができ, モータの出力を効率よくカメラの駆動へ用いることができる. このことにより, モータの大きさを最小限に抑えることができ, 軽量化にもつながる. 結果としてカメラヘッドの総重量は $2.7 [\text{kg}]$ に抑えることができた [5]. またステレオ視することでターゲットの三次元データが得られるため, より有効な情報が得ることができる.

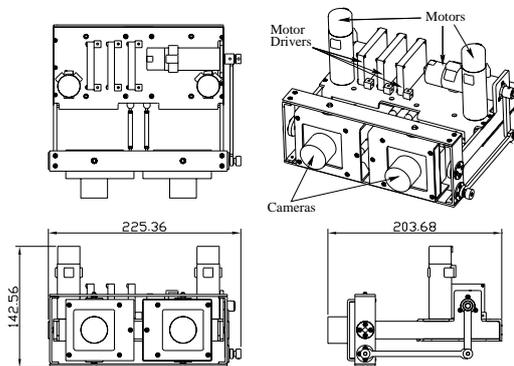


Fig. 3: Design of the binocular camera head.



Fig. 4: An overview of the binocular camera head.

2.3 システム構成

Fig. 5 にシステムの構成図を示す. 使用したコンピュータ PC/AT 互換機であり, CPU として PentiumIII $600 [\text{MHz}]$, メインメモリとして $512 [\text{MB}]$, OS として RTLinux-3.2-rc1 を使用した. カメラから得られる画像データはバスマスタ方式デジタル入出力ボードである, Interface 社製 PCI-2172C によって PC に入力される. 入力された画像データはユーザ空間で画像処理が行われ, 制御入力計算される. 計算された制御入力はリアルタイム処理が行われるカーネル空間へ渡される. カーネル空間では計算された制御入力をもとにモータの制御を行っている.

ビジョンチップからは $8 [\text{ms}]$ ごとに 1 フレームのデータが得られる. 現在は画像データの座標 (0,0) の情報であることを示す sync 信号を確実に捕えるため, 2 フレームの画像を読み込み, sync 信号から 1 フレーム分の画像を切り出している. したがって 1 フレームの画像情報を PC で得るには, $17.2 [\text{ms}]$ の時間を要する. そして $0.2 [\text{ms}]$ で対象物画像重心を計算し, $1 [\text{ms}]$ の周期のサーボループへ目標角度が送られる.

3 ビジュアルサーボ実験

3.1 制御則

ターゲットの重心が常に画面の中心へくるようにビジュアルサーボを行った. 重心は, ビジョンチップから得られる画像データに閾値を設け二値化を行い, x 座標, y 座標の平均値を計算し以下の式で求める.

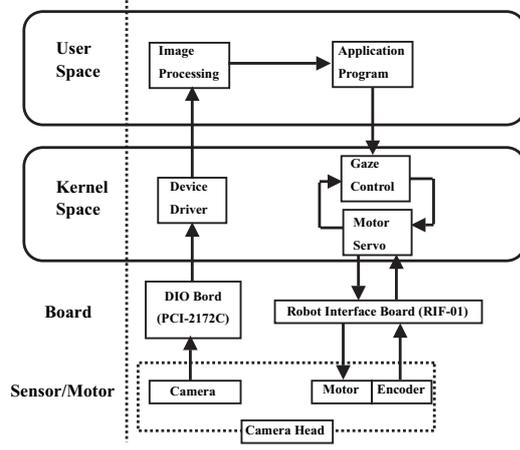


Fig. 5: System of the camera head.

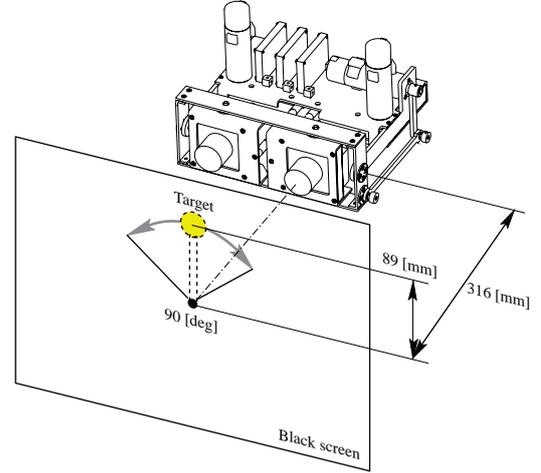


Fig. 6: Experimental setup.

$$C_{gx} = \frac{\sum_{i=1}^M X_i}{M} \quad (1)$$

$$C_{gy} = \frac{\sum_{i=1}^M Y_i}{M} \quad (2)$$

ここで C_{gx} [pixel], C_{gy} [pixel] はターゲットの重心位置の x, y 座標であり, X_i, Y_i は閾値以上の値をとった x, y 座標, M は閾値以上の値をとった画素数である. 閾値は実験により適当な値を設定した. 画像中心をターゲット重心に追従させる制御則は,

$$u = K_p(\theta_d - \theta) \quad (3)$$

で与えられる. この中で u [rad/s] は制御入力, θ_d [rad] は目標角度, θ [rad] は現在角度である. K_p [1/s] は比例ゲインであり, 本実験では 30 とした. カメラの現在角度はエンコーダの値を読み取り, 運動学を解いてカメラの姿勢角を計算し得られる. 目標角度は次式で計算される.

$$\theta_d = \theta + d\theta = \theta + \tan^{-1}(Ed_{pixel}/f) \quad (4)$$

ここで $d\theta$ [rad] は画像中心と重心点の角度差, E [mm/pixel] は画素差を角度差に変換する行列, d_{pixel} [pixel] は画面中心と重心点の画素差, f [mm] は焦点距離である.

3.2 実験方法

Fig. 6 に示す実験装置を用いてビジュアルサーボ基礎実験を行った. 本実験では両眼視覚でのビジュアルサーボを目標とした上で, 基礎実験として単眼での実験を行った. カメラからターゲットの回転中心軸までの距離は約 316 [mm], カメラの初期位置での光軸とターゲットの回転中心軸は一致させた. ターゲットのついた約 89 [mm] のアームを角度 ± 45 [°], 周期 0.5, 1.0 [Hz] で動かした. ターゲットは直径 12.6 [mm] の白い円盤状の物体にした. カメラヘッドの Tilt 角, Vergence 角は Fig. 7 で示すように定義する.

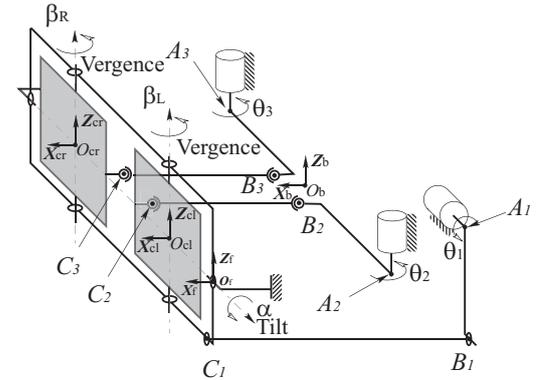


Fig. 7: Coordinate of camera head.

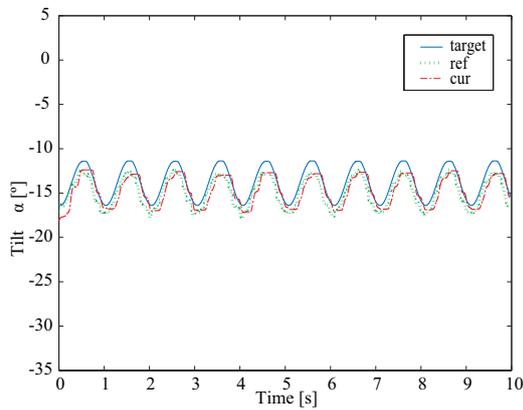
3.3 実験結果

アームの角度を周期 0.5 [Hz] で動かしたときの Tilt 角と Vergence 角の動きを Fig. 8 (a), (b) に, また 1.0 [Hz] の周期で動かしたときの Tilt 角と Vergence 角の動きを Fig. 9 (a), (b) に示す. 図中の target は実際のターゲットの角度, ref はビジョンチップから得られるターゲットの重心点の角度, cur はエンコーダから計算されるカメラの姿勢角を表す. また Fig. 10 (a), (b) に周期 1.0 [Hz] で動いているターゲットを 45 [°] と 0 [°] の位置でカメラがとらえた画像を示す.

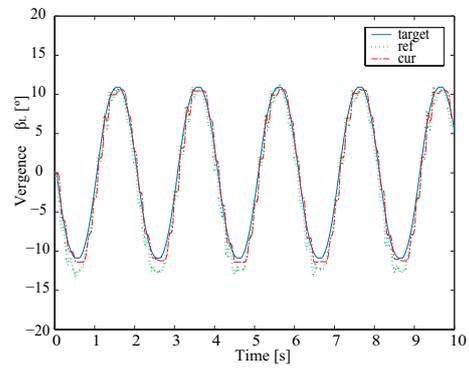
0.5 [Hz], 1.0 [Hz] とともにビジョンチップから得られる重心点は約 2 [°] の誤差内でとらえられていることが分かる. Fig. 19(a), (b) のカメラがとらえた画像からも, ターゲットが画面内にとらえられていることが確認できる. 2 [°] の誤差をピクセルに換算すると約 11 [pixel] に相当する. 誤差の原因として, Fig. 10(b) にみられるようにターゲットの残像が写っているためだと考えられる.

4 今後の課題

物体画像重心の計算方法において, 残像のような画像ノイズに影響されない画像処理を適応する必要がある. また今回は単眼でのビジュアルサーボを行ったので, 今後は両眼でビジュアルサーボを行う予定である. 今回の実験はビジュアルサーボ基礎実験であり, 比較的画像処理が容易な重心点を用いてビジュアルサーボ

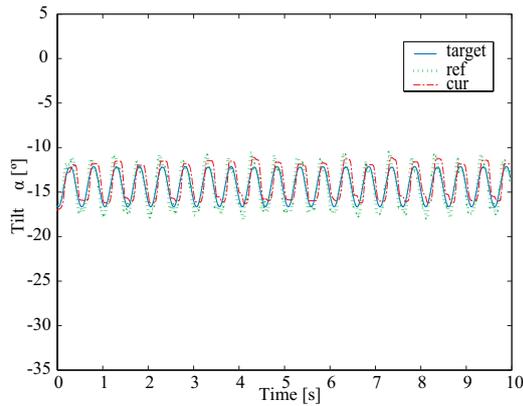


(a) Tilt

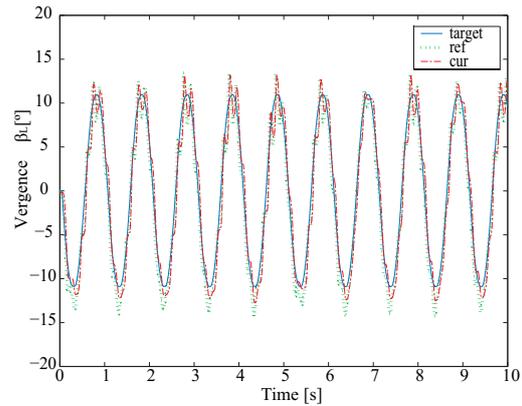


(b) Vergence

Fig. 8: Experimental results of visual servo 0.5 [Hz].

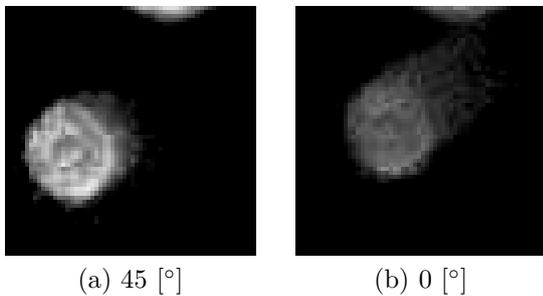


(a) Tilt



(b) Vergence

Fig. 9: Experimental results of visual servo 1.0 [Hz].



(a) 45 [°]

(b) 0 [°]

Fig. 10: Camera image while tracking an object. 1.0 [Hz]

を行い本研究のシステムの有効性を確認した．今後はロボットを制御する際にどのような情報が有効で、どのような画像処理を行うことが適当かを考えていく必要がある．そこで考えられた画像処理を今回の実験と同様に検証を行っていく予定である．またその検証によって得られる画像処理をビジョンチップで行うことにより、更なる高速化を目指す．

5 おわりに

本論文ではビジョンチップを用いた両眼視覚システムのための基礎実験として、単眼でのビジュアルサーボ実験を行った．その結果 Vergence 角, Tilt 角共に最大追従誤差が 2 [°] 以内であることを確認した．

参考文献

- [1] 三ツ谷祐輔, 四脚歩行ロボットの視覚に基づく行動制御に関する研究, 東北大学大学院工学研究科修士論文, 2001
- [2] 三ツ谷祐輔, 近野敦, 阿部幸勇, 内山勝, 四脚歩行ロボット搭載用両眼視覚システムの開発, 計測自動制御学会東北支部 第 188 回研究集会 188-14, 2000.
- [3] Mitsumasa Koyanagi, et al., Proc. of the 2001 IEEE International Solid State Circuits Conference, pp. 270-271, 2001
- [4] 栗野浩之, 中川源洋, 季康旭, 中村共則, 山田裕介, 諸岡哲, 稲村清, 朴起台, 小柳光正, 三次元集積技術を使ったビジョンチップ, 信学技報, pp. 29-35, 2001
- [5] 辻田哲平, 近野敦, 内山勝, パラレルリンク機構を用いた両眼視覚システムの開発, 計測自動制御学会東北支部 第 206 回研究集会 206-2, 2002.