学術・技術論文

アクティブビジョンの高速化を担う光学的視線制御システム

奥村光平*奥 寛雅*石川正俊*

Optical Gaze Control System to Realize More High-speed Active Vision

Kohei Okumura*, Hiromasa Oku* and Masatoshi Ishikawa*

The purpose of our study is to realize an active vision capable of casting its gaze in any desired direction in msorder. It is, however, difficult even if a vision sensor itself is controlled directly by high-speed actuators. Therefore, we propose a high-speed gaze control subsystem, called "Saccade Mirror." Using this system and a high-speed vision, image sequences that each image has different direction of gaze in msorder has been successfully obtained. Measuring the frequency response of the gaze control, this result has exceeded the preceding study. That is to say, the utility of our proposal has been ascertained.

Key Words: Active Vision, Gaze Control, Visual Servo, Visual Feedback

1. はじめに

人間が外界から受け取る様々な感覚情報のうち,視覚器官に よる情報は80%以上を占めると言われている[1].これを基に, ロボットを制御する際に外界の状況を把握する手段としても,視 覚情報を利用したビジョンシステムが有用であると以前から考 えられてきた.そしてよりよい視覚情報処理に向け,核となる ビジョンセンサはこれまで様々な革新を遂げてきた.筆者らは その革新の中でも特に「アクティブ化」と「高速化」の2点に 着目している.

まずビジョンセンサの「アクティブ化」について説明する.人間は、自らが行動することで能動的に視覚情報を取得し、それ を基に次の適切な行動をとることができる.これを既存のビジョ ンシステムに応用することが議論され始め、能動的な視覚情報 の取得を担うシステム、いわゆる「アクティブビジョン」とい う概念が誕生し、1980年代半ば以降広く研究されるようになっ た[2].本論文では、アクティブビジョンとはビジョンセンサの 視線をなんらかの手法によってアクティブに制御できる機能を 持つビジョンシステムの総称と定義する.

一方,ビジョンセンサの「高速化」も注目すべき革新と考えら れる. 視覚情報処理では通常 CCD カメラなどフレームレート が 30 [fps] のものが用いられることが多いが,画像のサンプリ ング速度がビデオ信号の伝送レート (30 [Hz])で制限され,高 速な視覚制御を実現するには不十分である.そこで現在は、こ のような場面では通常のビデオフレームレートのビジョンセン サではなく、1 [ms] サイクルといった非常にわずかな時間内で 撮像・画像処理が可能な「高速ビジョン」と呼ばれるビジョン センサが用いられている.

筆者らはこの二者の革新を同時に満たすシステム「高速アク ティブビジョン」の有用性を期待している.すなわち,高速ビ ジョンの視線を撮像・画像処理レートに見合う速度で動的に制御 すれば,従来では速度面で不可能であった様々なダイナミクス をも認識することのできる新たなアクティブビジョンが実現さ れると考えられる.例えば超高速飛翔物体の認識や,衝突・破裂 といった瞬間的な物理現象の視覚的評価などが挙げられる.よ り広く言うならば,応用先は視覚情報処理を含む多くのフィー ルドであり,それはロボットビジョン,放送・映像メディアコ ンテンツ,計測システムなど様々に期待される.

ところが現実には、ビジョンセンサの視線を制御するのに要 する時間が撮像・画像処理に要する時間に比べはるかに長く、こ れがシステム全体の高速化のボトルネックとなっている. なぜ なら、質量の大きいビジョンセンサ自体を撮像・画像処理レー トの1[kHz] に見合うほど高速に動かすのは容易ではないから である. 高速アクティブビジョンを実現させるには、ビジョン センサを動かすのではなく、なんらかの別の手法で視線を制御 する必要があると言える.

そこで筆者らは、ビジョンセンサ自体は固定したまま、それに 別途視線の制御だけを担う外付けのサブシステムを使用するこ とで、本ボトルネックを解消できるのではないかと考えた.この サブシステムの概要図を **Fig.1** に示す、本システムは、軽量ゆ え高速に駆動できる 2 軸の小型ミラーと、実用的な画角を実現

原稿受付 2009 年 12 月 20 日

^{*}東京大学大学院情報理工学系研究科

 $^{^{*}\,\}mathrm{Graduate}$ School of Information Science and Technology, The Univ. of Tokyo

[■]本論文は有用性で評価されました.



Fig. 1 A concept of our proposal: A high-speed gaze control is realized by mirrors and an angle of view is extended by Pupil Shift Lenses

させるための瞳転送系と呼ばれる光学系から構成される.また人間の無意識の高速眼球運動に準え、サッカードミラー(Saccade Mirror)と呼ぶものとした[3].

本論文では、サッカードミラーの構成と仕組み、プロトタイ プシステムの構築、およびその応答性能の評価実験について順 次解説し、高速アクティブビジョンの実現に向けた本システム の有効性を実証する.

2. 視線制御の手法比較

筆者らの提案する駆動ミラーによる視線制御の手法および, その他の視線制御の手法とを本章にて比較・考察していく.

2.1 ビジョンセンサ自体の向きを機械的に制御する手法

本手法のメリットは、アクチュエータの仕様次第で広範囲な視 線制御を実現できる点、視線制御機構自体による画角の制約がな い点などが考えられる.監視用途やロボットビジョンなど広く用 いられている一般的な手法と言える[4]~[7].一方で高速ビジョ ンの撮像レートに見合うほど視線制御を高速化できないという デメリットが存在する.先行研究として、中坊らは1,000 [fps] の 高速ビジョン CPV-I の視線を Pan、Tilt2 軸のアクチュエータ を用いてビジョン自体の向きを制御し、1-ms ターゲットトラッ キングシステムを実現している[6][7].しかしこのシステムにお ける視線制御の遮断周波数は、オープンループ系で約 20 [Hz], 視覚フィードバック系で約 10 [Hz] と撮像・画像処理レートの 1 [kHz] に比べはるかに低い値を示している.

2.2 レンズ系を駆動させ視線を制御する手法

本手法はビジョンセンサ全体を動かすのに比べ質量の小さい レンズを駆動させることで高速な応答が期待できる.しかしな がらレンズを駆動させることでの視線の制御は,あくまでも光 軸を微調整するという程度でしか行えず,デジタルカメラの手 ぶれ補正 [8] などには適しているものの,本論文で目標とする 高速アクティブビジョンの実現に寄与するとは考えられない.

2.3 広角な画像を取得し、そこから所望の画像を切りだす手法

本手法は物理的な駆動箇所を持たず, すべてソフトウェア処 理で行えるというメリットがある.実際の高速ビジョンでも, し ばしばこのような区域を限定して計算機へ転送するという処理 は用いられている[9].ただしこれはフレームレートを高めるた めに1枚の画像の情報量を落とすという手法である.よって, その区域は撮像中では一定であり, リアルタイムにミリ秒オー ダで切り替えられるような仕様には現状なっていない.

仮にそれが行えたとすれば、どの程度のイメージセンサが必

要かを考えてみる.例えば、10 [deg] の画角である動的対象を 撮像することを仮定する.また使用するカメラ画像の解像度を, 現行の高速ビジョンの解像度を踏まえ 512×512 とする. Pan, Tilt 方向に 60 [deg] まで視線を制御できる視線制御システムを 用いた場合,この視線制御システムがカバーする撮像可能領域 の解像度を擬似的に表すと,

$$1 + \frac{1}{2\tan 5^{\circ}} \times 512 \approx 2,000 \tag{1}$$

となる. 2,000 × 2,000 の解像度とは, フルハイビジョン 1,920 × 1,080 の約2倍に当たる解像度であり, このような高解 像度のイメージセンサを高速ビジョンに直接搭載するのは考え にくい.

2.4 ミラーを駆動させ視線を制御する手法

本手法のメリットは質量の小さいミラーを駆動させることで 高速な応答が期待でき,かつある程度の広範囲な視線制御を実 現できる点にある.このような観点から筆者らは高速アクティ ブビジョンを実現させる上で先述の手法よりも優れた手法だと 考え,本手法を採用するに至った.

ここでいくつかミラーを用いた関連研究を紹介する. 韓国. Chonbuk Natinal Univ. の H. Kim ら[10] や, カナダ, National Research Council Canada の J. Taylor ら [11] は、それ ぞれ別の研究としてビジョンセンサ、レーザ、駆動ミラーを用 いて奥行き方向の計測を行うことのできるシステム(レーザレ ンジファインダの一種)を提案している.これらのシステムは 二次元範囲に高速にレーザをスキャニングするために、2軸の ミラーを用いている、レーザのような広がりを持たない光線の 高速制御には、本手法は適していると考えられている.しかし 我々が扱う光線はレーザではなく映像であるため、画角の制約 を強く受けてしまうことが予想される。一方映像を扱う研究で あれば、東京工芸大の木戸らはミラーを駆動させることで任意 の位置を拡大することのできるマルチビジョンシステム Optz を開発している[12]. このシステムは画角の制約の低い1枚の ミラーを2軸のアクチュエータによって制御することで二次元 範囲にビジョンセンサの視線を制御している.しかしこの手法 では筆者らが追求する高速性が損なわれてしまう.

すなわち筆者らが本手法を採用するにおいて考慮すべき点は, 2 軸ものミラーを用いながらいかに実用的な画角を実現させる かであると言える.

3.1 画角と高速性のトレードオフ

Fig.2のようなピンホールカメラモデル[13]において,カメ ラの撮像できる光景の範囲をイメージセンサの各端点に届く光 線束が成す角として表した値を画角(Angle of View)と呼ぶ. ビジョンシステムにおける重要なパラメータの一つである.ま た,画角の中心は外界の光線束が集光する点で,これは光学用 語で瞳(Pupil)と呼ばれる領域の中心である[14].

次に, 瞳から任意の距離 d だけ離れたある面における光線束 の通過面積 $S(\theta)$ と画角 θ の関係を考える. この $S(\theta)$ は, 視 線の制御に用いるミラーの最低必要領域と見なすことができる. また, 図より θ が大きいほど $S(\theta)$ も大きくなることが分かる.



Fig. 2 An angle of view in a pinhole camera model



Fig. 3 A pupil shifting using a positive lens

つまり、システムの画角を広くとるならば、それだけ大きなミ ラーが必要となることが言える。ところがミラーが大きいとそ れだけ質量も大きくなってしまうため、本研究の趣旨でもある 高速性が損なわれてしまう。以上より、ミラーを用いたこの手法 において、画角と高速性はトレードオフの関係になっているこ とが分かる。特に本提案手法では、Pan、Tilt と 2 軸のミラー を用いるため、画角への制約は格段に大きくなることが分かる。

3.2 瞳転送系による画角の拡大

このような画角と高速性がトレードオフになる問題を解決す るため、サッカードミラーには瞳転送系と呼ばれる光学系を装 備させている。そもそも凸レンズにはレンズを介した対面側に 瞳を転送させる性質がある。このような目的で用いられる凸レ ンズを瞳転送レンズ (Pupil Shift Lens) と呼ぶ [15]. **Fig.3** に これを示す.

図中(A)はピンホールカメラモデルの場合で、凸レンズを通 過させることで本来のカメラ瞳Aと別にもう一つの転送された 瞳Bを実現できる.この転送された新しい瞳B付近にミラーを 配置することで、ミラーを大きくせずとも画角だけを向上する ことが見込めるわけである.なお、現実のビジョンセンサにお いて1枚の凸レンズを用いるだけでは図中(B)のようにフォー カスがずれてしまう.これを解決するために、瞳転送系は1枚 のレンズではなく複数のレンズを用いるのが一般的である.次 章で紹介するプロトタイプシステムでは3枚のレンズから構成 される.



Fig. 4 The optical setup of the prototype system



Fig. 5 A photograph of the Saccade Mirror

4. 実装および基本的な光学性能の検証

市販の光学素子を用いてサッカードミラーのプロトタイプシ ステムを構築した.この光学系をFig.4に、また実際の写真を Fig.5に示す.視線の制御を担うミラーについては、レーザの スキャニングに用いられるガルバノミラーと呼ばれる光学素子 を用いるとした.そのなかでも有効なビーム径が大きめのもの を理想的と判断し、GSI GroupのM3を使用した.本製品の有 効なビーム径は最大 30 [mm] とガルバノミラーのなかでも最大 級であり、光学走査角は±30 [deg] と広範囲の視線制御が期待 される.また入力指令電圧に対し走査角がリニアに応答するた め非常に扱いやすい製品である.一方瞳転送系は、対物レンズ、 フィールドレンズ、コリメータと呼ばれる3枚の凸レンズから 構成されている.これらのレンズの仕様を以下にまとめる.な お詳細な機能などは付録に別途記載した.

- •対物レンズ:45105-K, Edmund Optics.
 - -焦点距離:80 [mm].
 - -レンズ口径:40 [mm].
- •フィールドレンズ:45353-K, Edmund Optics.
 - -焦点距離:100 [mm].
 - レンズ口径:50 [mm].
- ・コリメータ: 45218-K, Edmund Optics.
 - 焦点距離: 60 [mm].
 - レンズ口径:40 [mm].

また,このシステムで実現できる最大の対角線画角を測定した. Fig.6 に測定手法と結果を示す.プロトタイプシステムの 瞳位置からおよそ 1,000 [mm] 離れた地点にホワイトボードを 設置し,そこでの撮像範囲をメジャーで測定した.このとき視 線とホワイトボードの面は直交している.測定結果は,目標設 計値 40 [deg] に対して,半対角線の長さが約 350 [mm],すなわ



Fig. 6 A measurement of the angle of view

ち対角画角換算で 38.6 [deg] であった.

5. オープンループ系での性能評価

5.1 視線制御の周波数応答

サッカードミラーの基本的な動特性として、オープンループ 系での視線制御の周波数応答を測定した.結果を**Fig.7**に示 す.図中(A)はミラーへの入力指令に対する実際に得られた走 査角のゲインを、図中(B)はその位相遅れを表している.この グラフより、ゲインが約3[dB]低下する周波数、すなわち遮断 周波数が Pan、Tilt とも 200 [Hz]以上であることが分かった. 中坊らのシステムでは、同条件の実験で約20 [Hz][6][7]とあ り、本システムにより10倍の高速化が可能と考えられる.ま た、遮断周波数付近での位相遅れはおよそ90 [deg]であること が分かった.

5.2 所望の視線に変位するまでに有する時間

ある視線を向いている状態から別の所望の視線に変位するま でに有する時間を測定するため, Fig.8のようなオープンルー プの撮像システムを構築した.実際に用いた機器の詳細は以下 のとおりである.

●高速カメラ:Phantom V4.2, Vision Research.

- ●任意波形発生器: 33220A, Agilent.
- ●データロガー:NR-600, キーエンス.

ここでは最も時間のかかる場合を想定し, ミラーの角度を最 大に変位させる場合, すなわち最大角視線制御に要する時間を 測定するものとする.

5.2.1 実験概要

任意波形発生器からの入力指令 *V_{in}*[V] は,時刻を *t*[ms] として,次式で与えられる台形波とする.

$$V_{in} = \begin{cases} \frac{3}{2}(t-2) & (0 \le t \le 4) \\ 3 & (4 \le t \le 10) \\ -\frac{3}{2}(t-12) & (10 \le t \le 14) \\ -3 & (14 \le t) \end{cases}$$
(2)

なお、本来このような応答を評価する場合は矩形波を用いるの が一般的だが、本システムでは用いたミラーの仕様上、安定動 作が保証される有限の傾きを持つ台形波を用いている.ただし 傾きの絶対値については限界まで大きいものを用いている.

さらに実際の視線の変位を視覚的に評価するため、入力電圧 の変化開始と同時に高速カメラが撮影を開始するように、任意 波形発生器よりトリガ信号が出力される.撮影に用いる高速カ



Fig. 7 The frequency response in a case of a open loop system; (A) gain [dB], and (B) phase delay [deg]



Fig. 8 A system setup for examining the settling time



Fig. 9 The experimental result by a data logger; the temporal change of the instruction voltage, the mirror angle and the trigger

メラのフレームレートは 1,000 [fps] である.また入力指令,ミ ラーの応答およびトリガ信号はデータロガーによってサンプリ ングレート 200 [kHz] にて記録される.

5.2.2 実験結果

データロガーによる定量的な実験結果を Fig.9 に示す.入力 指令の立ち上がりで4[ms] かかったあと,さらに2[ms] ほど遅 れてミラーの応答が安定していることが分かる.一方高速カメ ラによる視覚的な実験結果を Fig.10 に示す.本結果で6[ms] の画像には若干だがブラーが残っている.これについては以下 のように考察できる.本実験ではトリガを用いて映像とミラー の同期を取っているが、トリガの入力と、ある画像に対する露光 の開始時間が完全に同期がとれるとは限らず、撮像時間の1[ms]



Fig. 10 The experimental results by a high-speed camera; (A) a whole subject and (B) the image sequences

未満の誤差まで発生しうるのである.この誤差も考慮した上で, 本システムでは7[ms] 以内に視線の変位が完了し所望の向きへ の視線が実現されることが判明した.

6. 視覚フィードバックによるターゲットトラッキング

動的な対象の認識という点で最も基本的かつ有効なタスクで あるターゲットトラッキングは、視線制御の性能指標と判断で きる. そこでサッカードミラーに視覚フィードバックを組み込 み、ターゲットトラッキングシステムを構築した.

6.1 実験システム

Fig. 11 に構築した実験システムを示す.また,実際に使用した機器の詳細は以下のとおりである.

- •計算機:Xeon2.8GHz, 4GB RAM, Windows XP.
- 高速ビジョン: CPV-IV, 浜松ホトニクス [9].
- ●データロガー:NR-600, キーエンス.

計算機は、高速ビジョンで得られた画像の画像処理とミラー の制御とを両方行う.画像の入力インターフェースとしてカメ ラリンクを用い、ミラーへの指令値出力は DA ボードにより出 力される電圧値で行った.なお、本ターゲットトラッキングシ ステムの制御では 1 [ms] オーダの制御周期が求められるため、 Windows XP が備える高い時間分解能のタイマー割り込みであ る、Multimedia Timer を用いた.このタイマーは最短で 1 [ms] ごとのタイマー割り込みを指定できるが、Windows XP は非 リアルタイム OS であるので高い精度の割り込み周期の維持は 保証されていない.しかし、不要なプロセスを止めて計算機の 負荷を低く保つことで比較的安定して割り込み周期を保てるこ とが分かっており、カメラリンクインターフェースのデバイス ドライバーがそのまま使えるという利点もあるため、今回実装 したプロトタイプシステムの評価であれば支障はないとみなし て導入した.

6.2 トラッキングアルゴリズム

トラッキングのアルゴリズムを **Fig. 12** に示す. 以下図中各 項目について解説する.

(1) Capture images.

高速ビジョンによって対象が1[ms]ごとの連続画像として撮影される。



Fig. 11 A system setup for target tracking



Fig. 12 Tracking algorithm

(2) Remove the fixed pattern noise.

連続画像が計算機へ転送され、固定パターンノイズ(Fixed pattern noise)と呼ばれるハードウェア起因のノイズを取り除く、これについて詳しく説明する。

実は何も見えない真っ暗な状態であっても、イメージセンサ を構成するフォトダイオードは微小ながら電流を出力している. これを暗電流(Dark current)と呼ぶ.そして暗電流は各ピク セルのフォトダイオードに依存するので、画像中には常に固定 のパターンを持つノイズが確認される.よって取得画像から固 定パターンノイズだけの画像を減算することでキャリブレーショ ンをしている.なおこの画像は0から1の値を持つ画素の集合 関数 *I*_o(*x*, *y*) とみなすことができる.

(3) Binarize to recognize the only object.

本アルゴリズムでは簡単のため、この対象は画像中で最も明 るいと仮定する.このとき適当な閾値 i ($0 \le i \le 1$)を取ること で、対象とそれ以外とが 2 値として区別されたバイナリーデー タが生成される.これは次式のような 2 値関数 $I_b(x,y)$ として 与えられる.

$$I_b(x,y) = \begin{cases} 0 & (i > I_o(x,y)) \\ 1 & (i \le I_o(x,y)) \end{cases}$$
(3)

(4) Calculate the center of mass.

生成されたバイナリーデータから画像の重心位置を算出する. 以下計算法を示す.式(3)を用いて,(p+q)次の画像モーメ ントと呼ばれる値が

$$m_{pq} = \sum_{x} \sum_{y} x^p y^q I_b(x, y) \tag{4}$$

と定義される. 例えば、0 次モーメント m_{00} は画像中における対象の面積和である. x 座標、y 座標それぞれの重心座標 $G(x_{cm}, y_{cm})$ を得るためには、1 次モーメントを 0 次モーメントで割った値、 $m_{10}/m_{00}, m_{01}/m_{00}$ を求めればよい. これを次式に示す.

$$\begin{pmatrix} x_{cm} \\ y_{cm} \end{pmatrix} = \begin{pmatrix} \frac{\sum_{x} \sum_{y} xb(x,y)}{\sum_{x} \sum_{y} I_{b}(x,y)} \\ \frac{\sum_{x} \sum_{y} yb(x,y)}{\sum_{x} \sum_{y} I_{b}(x,y)} \end{pmatrix}$$
(5)

(5) Scan mirrors.

 $G(x_{cm}, y_{cm})$ と画像の中心との距離をx成分,y成分それぞれについて算出し,それらが0になるようにミラーの角度を逐次制御する.

(6) Observe the target.

(1)~(5)のプロセスを繰り返すことでトラッキングを実現し、ターゲットを継続的に観察することができる [16].

6.3 高速回転体による視覚フィードバック系での周波数応 答測定

視覚情報から目標位置を推定するのはいわゆる逆問題を解く ことになる.そこで本実験では回転運動という規則的な運動を 取り扱い,目標位置を比較的高い精度で定量化し,視覚フィー ドバック系での周波数応答を測定する.ターゲットに用いたの は,回転数を可変できる市販のDC 軸流ファン(F410T-12LC, 日本電産コパル電子)の1点に付けた白色マーカである.これ を Fig. 13(A) に示す.

6.3.1 目標位置の定量化

まず目標位置を定量的に表す. ここでは回転面と視線の光軸 は直交しているとし,交点を原点とする **Fig. 13** のような座標 系を考える.回転半径を r [mm],回転角を θ [rad], プロトタ イプシステムの瞳位置から原点までの距離を d [mm],視線の Pan 走査角を ϕ_{pan} [rad] とする. これらのパラメータについて Fig. 13 (B) より幾何的に次式が導出される.

$$\phi_{pan} = \tan^{-1} \left(\frac{r \cos \theta}{d} \right) \tag{6}$$

ここで 回転周波数を f [Hz], 経過時間を t [s] として, $\theta = 2\pi f t$ と置くと ϕ_{pan} を時間の関数として次式のように表せる.

$$\phi_{pan}(t) = \tan^{-1}\left\{\frac{r\cos(2\pi ft)}{d}\right\}$$
(7)

Tilt 走査角についても同様にして,

$$\phi_{tilt}(t) = \tan^{-1}\left\{\frac{r\sin(2\pi ft)}{d}\right\} \tag{8}$$

となる. なお,本実験では,r = 20およびd = 200を採用した.



Fig. 13 A relationship between θ and ϕ_{pan}



Fig. 14 The frequency response in a case of a visual feedback system; (A) gain [dB], (B) phase delay [deg], and (C) the method of measuring the phase delay

6.3.2 撮像・画像処理レートと視線制御の動作周波数の整合 条件

Paul によれば、ロボット等をフィードバックを用いて安定に 制御する場合、システム内の最高位の制御レート f_s [Hz] と機 械系固有振動数 f_m [Hz] は次式の関係を満たすことが条件であ るとされている [17].

$$f_m \le \frac{1}{10} f_s \tag{9}$$

本ターゲットトラッキングシステムにおいて f_s は撮像・画像処 理レートに当たり、 f_m は視線制御の動作周波数に当たる. し たがって $f_s = 1,000$ を代入すれば

$$f_m \le 100 \tag{10}$$

と導出される. すなわち視覚フィードバック系における視線制 御の遮断周波数が 100 [Hz] 以上であれば Paul による条件を完 全に満たすことのできるシステムであるといえる.

6.3.3 実験結果

視覚フィードバック系での視線制御の周波数応答を Fig. 14 に示す. 図中(A)は式(7)(8)より推定された目標位置の電圧 に対する応答電圧のゲインを表している. 遮断周波数は, Tilt については約 100 [Hz] であり, Pan についてはそれより高い値 と推測されるが, 図中(B)に示したように, 100 [Hz] 付近では 位相遅れが 180 [deg] に達し, これ以上の値を測定することがで



Fig. 15 Image sequences while the rotating object ((A) 11 [Hz], (B) 33 [Hz], and (C) 50 [Hz]) was being tracked; numbers under the images are the elapsed time [ms]

きなかった.ただしいずれにしても Paul による条件を限界まで満たせるほどに高速な視線制御を実現できることが分かった.

また、本来位相遅れを測定する場合、各システムどうしの厳 密な時間同期が必要となるが、本システムの計算機は非リアル タイム OS を用いているため、それが困難である。そこで位相 遅れについては、図中 (C) のように高速ビジョンより得られた 画像から、視線中心と回転中心を結ぶ直線およびターゲットの 重心と回転中心とを結ぶ直線の成す角 θ_{delay} を測定している.

なお、オープンループ系での遮断周波数における位相遅れに 比べ本実験でのそれはより大きな値が確認されている.これにつ いては以下のように考察することができる.ターゲットトラッキ ングシステムでは Windows XP を搭載した計算機を使用するた め、推定 2 [ms] 程度の遅延がどうしても避けられない.従来の アクティブビジョンでも Windows を用いた場合は計算機由来の 遅延は発生していたと考えられるが、例えば遮断周波数 10 [Hz] に対する 90 [deg] の位相遅れとは 25 [ms] なので、計算機の遅 延の影響はそこまで大きくない.一方、筆者らのシステムでの 遮断周波数 100 [Hz] に対する 90 [deg] の位相遅れは 2.5 [ms] で あるため、計算機の遅延が問題となってしまう.これは高速ア クティブビジョンに関して今後解決すべき課題の一つと考えら れる.

さらに,高速ビジョンが実際にとらえた連続画像を Fig. 15 に示す.各画像下の数値は経過時間をミリ秒単位で表示したも のであり,また各画像の2本の対角線の交点を視線中心とする. 図中(A)は回転周波数が11[Hz]の場合で,先行研究における 限界とされていたものだが,本システムを用いる場合では,位 相遅れはほとんど目立たず視線中心に常にターゲットを確認す ることができる.

図中(B)はその3倍にあたる33[Hz]の場合で,図中(C)は 位相遅れが90[deg]に近い値を持つ50[Hz]の場合である.周 波数を高めるごとに位相遅れの影響が大きくなるが,ゲインは いずれも取れているので視線の中心は回転するターゲットの軌 道上にほぼ向けられている.

6.3.4 まとめ

中坊らの先行研究に対しサッカードミラーを用いる場合では, 遮断周波数を評価の基準として,視覚フィードバックの有無双 方について約10倍程度の高速化が見込まれる結果となった.す なわち本システムを用いることで,高速アクティブビジョン実 現に向けたボトルネックの解消が実現できたと考えられる.

ただし別の問題として,計算機の遅延が位相遅れに直接影響 してしまうことも分かった.これについては今後計算機の仕様 を改善するなり, FPGA を用いた制御に変更するなりで対処し ていかねばならないと考えている.

6.4 アプリケーションの提案

6.4.1 スポーツの世界に貢献する評価アプリケーション

本ターゲットトラッキングシステムは大変高速な応答を実現 できるため、高速運動物体や瞬間的な物理現象の観察に大変有 効であると考えられる。そこで筆者らはスポーツの世界に大き く貢献するアプリケーションを提案する。

野球,サッカー,テニスなどボールが高速に運動するスポー ツは多く存在するが,通常のカメラで速度ベクトルの変化が激 しいボールの詳細な動きを継続的にとらえるのは現状困難であ る.しかしこのような映像を取得できたならば,それは放送コ ンテンツとしての映像の高質化,その映像による選手の自主練 習のサポート,あるいはスポーツ用品メーカーの製品研究開発 用途など多くの需要があるといえる.そして,このような需要 を満たすアプリケーションの開発には、ミリ秒オーダで高速に 視線を制御することの可能なサッカードミラーが大きく貢献す ると考えられる.

6.4.2 予備実験

本アプリケーション構築に向けた予備実験としてゴムボール の追従観察実験を行った.

このゴムボールは突然視野内に入り込み、卓上で跳ね返り、さ らにラケットによって打たれるという複雑な動作をするもので ある.従来のアクティブビジョンを用いた場合、特に急激に速 度ベクトルが変化する卓上での跳ね返りや、ラケットに打たれ



Fig. 16 Image sequences captured by a high-speed vision. A ball is (A) falling from a hand, (B) bouncing against a table, (C) going up, and (D) being hit by a puddle. Numbers in the images are elapsed time [ms]



Fig. 17 The gaze angle while the object was being tracked

た直後に視野から外れてしまっていた.サッカードミラーを用いた本ターゲットトラッキングシステムを用いることで安定な 観察ができると期待される.

観察結果を Fig. 16 に示す. 図中 (A) はボールが自由落下し て視野内に入り込みターゲットトラッキングが開始される様子で ある. 図中 (B) はボールが卓上で跳ね返る前後の様子を1[ms] ごとに表示したものである. 衝突の瞬間の変形などをはっきり と確認することができる. 図中 (C) は跳ね返ったあとのボール の様子である. 図中 (D) はそのボールがラケットで打たれる前 後の様子を1[ms] ごとに表示したものである. また, 画像中の 数値は (A)~(D) の各状態の時刻を示している. 例えば衝突の 瞬間など, 各状態で最も重要だと思われる瞬間を0と決定して いる.

一方このときの視線の時間変化を **Fig. 17** に示す. 本システ ムの特徴として, Pan, Tilt 両方について 60 [deg] という広範 囲な視線制御範囲角を有することが挙げられるが,本実験では ボールが視野内に入ってから卓上で跳ね返るまでに, Tilt 方向 で 50 [deg] も変位している. すなわち本システムの性能を大き く生かした実験結果であると考えられる. またラケットで打た れた直後の Pan 方向にボールが勢いよく飛んでいく様子も,本 結果から読み取ることができる.

6.4.3 まとめ

上述の複雑な動作を有するゴムボールを本システムを用いる ことで観察することに成功した.なお筆者らが何度か同様の実 験を行った結果であるが、人間がラケットでボールを打つ動作 にはムラがあり、良いと実感できるスマッシュが打てる場合と、 そうでない場合が存在する.それぞれボールの回転・並進の動 きなどを観察すると肉眼で分かるほどに違いがある.現状では 筆者らの感覚による評価だが、プロフェッショナルの意見など も参考になんらかの評価軸を設け将来的には上述の需要を満た すアプリケーションの実現に尽力していきたい.

7. おわりに

本論文における主張を以下にまとめる.

- (1) 高速アクティブビジョンを実現させれば、これまで不可能 であった様々なダイナミクスを認識・評価・計測すること ができると期待されるが、現実には視線制御の遮断周波数 が低くボトルネックとなっている。筆者らはこれを解決す る新たなサブシステム、サッカードミラーを提案した。
- (2) 市販の光学素子を組み合わせることでプロトタイプシステムを構築し、その動特性を評価した.その結果、オープンループでの使用、視覚フィードバックを用いる場合での使用、双方について先行研究と比較して10倍程度の高速化を実現させた.

今後の展望として,第一に計算機による遅延を改善すること である.高速回転体を用いたターゲットトラッキングでは,視 線の制御は逐次行われたものの,計算機による遅延の影響を大 きく受けてしまった. これは現在の OS が実時間処理を最重視 するリアルタイム OS を用いていないことが原因と考えている. これを変更することや,あるいは補正として位相進み制御を用 いることで,改善が見込まれると考えている.

第二に提案したスポーツ評価アプリケーションの実装である. 予備実験として既存のターゲットトラッキングシステムを用い て複雑に動作をするゴムボールの観察を行ったが、今後は出来 具合などの評価軸を設け本格的な評価アプリケーションを実現 していきたいと考えている.

第三に、アクティブな撮像用光学系の導入が考えられる.フォー カスの制御についても近年大きな進展があり、奥らによって、従 来数百ミリ秒から1 秒近くかかっていた焦点調節を、わずか ミリ秒オーダで実現させる光学デバイス、ダイナモルフレンズ (Dynamorph Lens; DML)が開発されている[18]. もし視線 の制御だけでなく撮像用光学系も同時に制御することができれ ば、奥行き方向の情報も取得可能となる.その結果、三次元空 間における任意の対象を高速に認識できるようなビジョンシス テムの実現が期待される.これは、ロボットビジョン、放送・映 像メディアコンテンツ、計測システムなど、視覚情報処理を担 う多くのフィールドで応用が可能であると考えている.

参考文献

- [1] 橋本:視覚情報とヒューマンエラー. 海文堂出版, 1996.
- [2] J. Aloimonos, I. Weiss and A. Bandyopadhyay: "Active Vision," International journal of computer vision, vol.1, no.4, pp.333–356, 1988.
- [3] 奥村,奥,石川: "駆動鏡面式超高速アクティブビジョン",第27回
 日本ロボット学会学術講演会予稿集 DVD-ROM, 3R1-02, 2009.
- [4] 大池, 呉, 加藤, 和田: "鮮明な画像撮影のための高速追従型アクティ ブカメラ", 情報処理学会研究報告. CVIM, [コンピュータビジョン とイメージメディア], vol.2004, no.40, pp.71–78, 2004.
- [5] D. Comaniciu, V. Ramesh and P. Meer: "Kernel-Based Object Tracking," IEEE TRANSACTIONS ON PATTERN ANALY-SIS AND MACHINE INTELLIGENCE, vol.25, no.5, pp.564– 575, 2003.
- [6] 中坊、石井、石川: "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",日本ロボット学会誌,vol.15, no.3, pp.417– 421, 1997.
- [7] Y. Nakabo, M. Ishikawa, H. Toyoda and S. Mizuno: "1 ms column parallel vision system and its application of high speed target tracking," Proc. of IEEE Int. Conf. Robotics and Automation, pp.650–655, 2000.
- [8] 芝崎: "4. ディジタルー眼レフカメラの手振れ防止技術",映像情報 メディア学会誌, vol.61, no.3, pp.279-283, 2007.
- [9] Y. Sugiyama, M. Takumi, H. Toyoda, N. Mukozaka, A. Ihori, T. Kurashina, Y. Nakamura, T. Tonbe and S. Mizuno: "A High-Speed CMOS Image Sensor with Profile Data Acquiring Function," IEEE J. of Solid-State Circuits, vol.40, no.12, pp.2816–2823, 2005.
- [10] H. Kim., C.-S. Lin, C.-B. Yoon, H.-J. Lee and H. Son: "A Depth Measurement System with the Active Vision of the Striped Lighting and Rotating Mirror," Progress in Pattern Recognition, Image Analysis and Applications, vol.3287, pp.108–115, 2004.
- [11] J. Taylor, J.-A. Beraldin, G. Godin, L. Cournoyer, R. Baribeau, F. Blais, M. Rioux and J. Domey: "NRC 3D imaging technology for museum and heritage applications," THE JOUR-NAL OF VISUALIZATION AND COMPUTER ANIMATION, J. Visual. Comput. Animat. 2003, vol.14, pp.121–138, 2003.
- [12] 木戸, 鈴木: "全方位マルチビジョンシステム OPTZ を用いたビジュ

アルサーボイング",第27回日本ロボット学会学術講演会予稿集 DVD-ROM, 3R1-04, 2009.

- [13] 松山, 久野, 井宮: コンピュータビジョン. 新技術コミュニケーショ ンズ, pp.80-82, 1998.
- [14] 辻内, 黒田 ほか:光学技術ハンドブック. pp.94-95, 朝倉書店, 2002.
- [15] 永田:レンズのわかる本. p.118, 日本実業出版, 2002.
- [16] B.K.P. Horn, ROBOT VISION. pp.46–49, The MIT Press, 1986.
- [17] R.P. Paul: Robot Manipulators. MIT Press, 1987.
- [18] H. Oku and M. Ishikawa: "High-speed liquid lens with 2 ms response and 80.3 nm root-mean-square wavefront error," Applied Physics Letters, vol.94, 221108, 2009.
- [19] 左貝:光学の基礎. pp.62-63, コロナ社, 1997.

付録 A. 瞳転送系の設計

瞳転送系を設計する上で撮像する対象は無限遠にあるものと 定義する.なぜなら,カメラ瞳は一般に 5~10 [mm] と大変小 さく,一般的なカメラの被写体,つまり 2 [m] 程度以上離れた 対象を撮像する場合は,相対的に無限遠にあると近似して差し 支えないからである.

以下レンズの設計手順概要を **Fig.18** に示す.まず図中(A) のように2枚の凸レンズを用意し、その距離が両焦点距離の和 になるように設置する.これで出力光も入力光同様平行光とな る.図中(A)の光線束(i)は2枚のレンズを通過した後も平行 光を維持している.すなわち無限遠の対象に関してフォーカス を合わせた状態を保っている.このとき光の入力側にあるレン ズを対物レンズ、出力側にあるレンズをコリメータと呼ぶ.

しかし、この2枚のレンズだけでは光線束の量的な問題が残っ ている.対物レンズに入射した光線束のすべてがコリメータを 通過するとは限らないのである.図中(A)の光線束(ii)がその ような場合で、このときビジョンセンサ側では、視野の外周が 光量不足から暗くなる口径食(Vignetting)という現象が起き ている.

これを防止するため、図中(B)のように、対物レンズとコリ メータの両焦点距離位置にフィールドレンズと呼ばれるもう1 枚の凸レンズを挿入する.このレンズは光線束の拡散や集光に は寄与せず、光線束そのものを屈折させるだけの効果を持つ.図



Fig. 18 The process of developing the Pupil Shift Lenses; (A) a field Lens: OFF, and (B) a field Lens: ON

中(B)の光線束(ii')は、この光線がフィールドレンズにより屈 折しコリメータへ入射する様子を示している.

付録 B. 画角とレンズパラメータの関係

Fig. 19 のような任意の光学系において、ある物体が光軸上に存在し、その交点から射出している光線束の角度を θ_e 、物体の大きさを y_e 、物体位置での屈折率を n_e 、同様に像と光軸との交点に集光する光束の角度を θ_i 、像の大きさを y_i 、像位置での屈折率を n_i とすると、次式の関係を満たすことが知られている.

$$n_e y_e \tan \theta_e = n_i y_i \tan \theta_i \tag{B.1}$$

この量を、ヘルムホルツ・ラグランジュの不変量(Helmholtz-Lagrange invariant)と呼ぶ [19].

ここで物体を転送されたシステムの瞳,像を本来のビジョン センサの瞳とみなすと本理論がそのまま適用できる.まず設計 する瞳転送系において両瞳位置での屈折率 n はともに空気中で あるから等しく,それぞれ適当な値,例えば1として考えてよ い.したがって **Fig. 20** のように、本来のビジョンセンサの画 角を α 、ビジョンセンサの入射瞳の口径を d_{α} 、瞳転送系の使 用によるシステムの画角を β 、転送された瞳の口径を d_{β} とす ると、

$$d_{\alpha} \tan\left(\frac{\alpha}{2}\right) = d_{\beta} \tan\left(\frac{\beta}{2}\right) \tag{B.2}$$

が成立する.一方瞳転送系において、対物レンズの焦点距離 f_o とコリメータの焦点距離 f_c は $\tan(\alpha/2)$ および $\tan(\beta/2)$ と 共役である.すなわち次式が成立する.

$$f_o: f_c = \tan\left(\frac{\alpha}{2}\right): \tan\left(\frac{\beta}{2}\right)$$
 (B.3)

以上,式(B.2)(B.3)より画角,瞳口径,対物レンズおよびコ リメータの焦点距離の関係式を以下のように導くことができる.

$$f_o: f_c = \tan\left(\frac{\alpha}{2}\right): \tan\left(\frac{\beta}{2}\right) = d_\beta: d_\alpha \quad (B.4)$$

また, これを Fig. 20 に示す.



Fig. 19 Helmholtz-Lagrange invariant



 $Fig.\ 20 \quad {\rm Focal \ lengths \ and \ an \ angle \ of \ view}$

付録 C. 全体の光学配置と各素子の仕様決定

サッカードミラーの実装に向け,各素子にどのようなものを 用い,どのような手順で配置するかを以下解説していく.なお 設計手順については必ずしも一意に定まるものではないため用 途や目的,その場の状況(現在所有している光学素子が限定さ れる等)に添ったそれぞれの手順が考えられる.本論文では,画 角や走査角をあらかじめ目標値として定め,それに沿った設計 を行う.

まず初めにシステムの重要パラメータである画角の目標値 β を、例えば本プロトタイプシステムを例にとれば 40 [deg] など と設定する.次に、視線の走査角がどの程度必要かを想定する. 走査角はミラーの仕様に依存するのでこの段階で素子として用 いる 2 軸のミラーを決定する.そしてこの画角を実現し、かつ 口径食のないミラーの配置をシミュレーションする.なお、ミ ラーを走査することで光軸に対するミラーの断面積は変化する ため、最も厳しい条件である両ミラーとも最大まで走査されて いる状態を考えなければならない.さらに、その後方に位置す る対物レンズの口径 D_o が決定する.本プロトタイプシステム では $D_o = 40$ としている.このとき、Pan ミラーと対物レン ズがぶつからない距離を保てる大きさを考える.ここまでで決 定したパラメータを **Fig. 21** に示す.

また,その後の光線束の軌道を考え,コリメータの口径 Dc お



Fig. 21 A simulation to determine positions of the mirrors: (A) Both mirrors are driven for clockwise. (B) Both mirrors are driven for counterclockwise



Fig. 22 A simulation to determine positions and sizes of the lenses $\label{eq:Fig.22}$

よびフィールドレンズの口径 D_f を決定する. 一般的にフィー ルドレンズを大きめにするとよい. 本プロトタイプシステムで は $D_c = 40, D_f = 50$ としている. 一方ビジョンセンサ固有の 画角 α について, コリメータの口径 D_c を絞りとみなせばコリ メータからビジョンセンサの入射瞳までの距離 l_{c-p} に依存す ることより計算が可能である. α, β が決定すれば, その値を式



奥村光平(Kohei Okumura)

2008年早稲田大学理工学部応用物理学科卒業.2010 年東京大学大学院情報理工学系研究科システム情報 学専攻修士課程修了.同年同専攻博士課程進学,現 在に至る.高速視線制御システムの研究に従事. (日本ロボット学会学生会員)



石川正俊(Masatoshi Ishikawa)

1977 年東京大学工学部計数工学科卒業.1979 年東 京大学大学院工学系研究科計数工学専攻修士課程修 了.同年通産省工業技術院製品科学研究所研究員. 1989 年東京大学工学部計数工学科助教授.現在東 京大学大学院情報理工学系研究科創造情報学専攻教 授.2002,2003 年度東京大学総長特任補佐.2004

年度東京大学副学長. 2005 年度理事・副学長. 超並列・超高速ビジョン, センサフュージョン, メタ・パーセプションなどの研究に従事. 1998, 2001, 2008 年度本学会論文賞受賞. 工学博士.

(日本ロボット学会正会員)

(B.4) に代入することで、各レンズに必要な焦点距離およびその配置を算出することができる。以上のプロセスでサッカード ミラーを構成する各光学素子のスペックおよび配置が決定され る.本プロトタイプシステムでは $\alpha = 30$ として式(B.4)に代 入することでレンズの仕様を $f_o = 80, f_c = 60, f_f = 120$ と決 定した。これらを**Fig. 22**に示す。



奥 寛雅 (Hiromasa Oku)

1998 年東京大学理学部物理学科卒業.2000 年東 京大学大学院工学系研究科計数工学専攻修士課程修 了.2003 年同専攻博士課程修了.2003~2005 年 科学技術振興機構研究員.2005 年東京大学大学院 情報理工学系研究科システム情報学専攻助手.2007 年同助教,現在に至る.ダイナミックイメージコン

トロール,ダイナモルフレンズなどの研究に従事. 2009 年度本学会 研究奨励賞受賞,2010 年度本学会論文賞受賞.博士(工学).

(日本ロボット学会正会員)