

指定カメラワーク下でのカメラ動作修正の抑制

新 康孝[†] 亀田 能成^{††} 美濃 導彦^{††}

[†] 京都大学情報学研究科 京都市左京区吉田本町

^{††} 京都大学学術情報メディアセンター 京都市左京区吉田二本松町

E-mail: [†]atarashi@mm.media.kyoto-u.ac.jp, ^{††}{kameda,minoh}@media.kyoto-u.ac.jp

あらまし 本稿では動物体を首振りカメラで撮影する際に、カメラ制御修正回数を抑制しつつ指定されたカメラワークを実現する方法について提案する。単一対象を撮影する際のカメラワークは、各フレームごとの対象の位置・大きさ・速度・大きさ変化、および背景の速度で定義できる。撮影を行う際、対象の動きによっては指定カメラワークと実現されたカメラワークとの間にずれが生じる。このずれを修正する際、指定カメラワークの正確な実現を目指して制御を行うと、カメラ動作の修正が頻発し、映像が対象の動きに伴うずれに対して敏感に反応して見にくい映像となる。従って、見易い映像の実現のためには、ずれの修正に際して、カメラ動作修正のタイミングと修正量について考慮を払うべきである。我々はこの点に注意し、実現カメラワークが指定カメラワークに対してある許容範囲内にあり続け、かつカメラ動作の修正を抑制するようなカメラ制御方法について提案する。

キーワード 自動撮影、カメラワーク、首振りカメラ

Reduction of Camera Motion Adjustments under a Planned Camera-work

Yasutaka ATARASHI[†], Yoshinari KAMEDA^{††}, and Michihiko MINOH^{††}

[†] Graduate School of Informatics, Kyoto University Yoshida-Honmachi, Sakyo-ku,

^{††} Academic Center for Computing and Media Studies, Kyoto University Yoshidanihonmatsucho, Sakyoku,
Kyoto

E-mail: [†]atarashi@mm.media.kyoto-u.ac.jp, ^{††}{kameda,minoh}@media.kyoto-u.ac.jp

Abstract We propose a method to control a camera in order to realize a planned camera-work with reduced number of camera motion changes. A camera-work to shoot a moving object can be defined by the position, velocity, size and magnification rate of the object, and the velocity of the background in the video image. When we obtain video images, unexpected motion of the object makes the difference between the resultant camera-work and the planned camera-work. If we control the camera so that the resultant camera-work is exactly the same as the planned camera-work, the camera motion are frequently occurred and the resultant video images are not comfortable for a human to see. Therefore, we should consider the timing and amount of camera motion adjustments. Considering this issue, we propose a method to control the camera so that the resultant camera-work is kept within an acceptable range of the planned camera-work with reduced number of camera motion adjustments.

Key words Automatic shooting, camera-work, pan-tilt-zoom camera

1. はじめに

講義、コンサート、試合などの人間の活動を自動的に撮影するために、様々な研究が提案されつつある [1] [2]。これらの撮影においては動物体の撮影が必要である。

首振りカメラを用いた撮影をどのように行うかはカメラワークによって指定できる。カメラワークとは、各フレームにお

る画像上での対象の位置・速度・大きさ・大きさ変化、および背景の速度からなるベクトルである。指定されたカメラワークを「標的カメラワーク (TCW)」と呼ぶことにする。一般に TCW の全要素を指定する必要はなく、たとえば、講師を画像内で中央にミドルショットで撮影したい場合には、対象の位置と大きさのみを指定すればよい。

カメラワークに基づく撮影を行う場合、対象の動きによって

は TCW と、実際に撮影されている画像上のパラメータとの間にずれが生じる。実際に撮影されている画像上のパラメータを「撮影状態パラメータ (SSP)」と呼ぶことにする。TCW と SSP とのずれを修正するためカメラを制御する。

ビジュアルサーボ [3] は画像を参照することによってカメラを制御する手法である。一般に、ビジュアルサーボでは TCW と SSP との差を減少するように常に制御が行われるため、カメラ動作の修正が頻発する。また、[4] のようにエピポーラ幾何を用いる場合には対応点を求める事が必要であり一般的な対象に適用する事は困難である。

このように、TCW と SSP のずれを修正する際、ビジュアルサーボを適用し、TCW の正確な実現を目指して制御を行うと、カメラ動作の修正が頻発し、映像が対象の動きに伴うずれに対して敏感に反応して見にくい映像となる。従って、見易い映像の実現のためには、ずれの修正に際して、カメラ動作修正のタイミングと修正量について考慮を払うべきである。

一方、カメラ動作の修正を抑制して動物体を撮影する手法として枠制御アルゴリズム [5] が提案されている。これは、対象の周囲に仮想的な枠を設定し、対象物がその枠内にある間カメラを固定する手法である。枠から対象物がでた場合枠制御モードを終了して追跡モードに入り、停留もしくは反復を検出することで再び枠制御モードに切り替える。

我々の手法では、位置だけではなく TCW の全ての要素に対して「許容範囲」を導入し撮影の指定の幅を広げる。さらに、制御を行うか否かの判断にのみ許容範囲を用いるのではなく、制御値の決定についても許容範囲を考慮し、一定時間先まで SSP の予測を行い、結果が許容範囲内になるように制御を行うことで、カメラ動作の修正を抑制する。許容範囲は、TCW で値が指定されている要素に対して上界と下界を与えることで決まる。TCW で値が指定されていない要素については、実現されているカメラワークにおいていかなる値をとったとしても、許容範囲内であるとする。全ての要素に対して SSP が許容範囲内であれば制約を満たしているとみなす。

なお、TCW は視聴者の興味と状況に合わせて決定されるべきものであり、ここでは TCW は他のシステム、あるいは人手によってあらかじめ与えられているものとする。TCW の全ての要素が指定されている必要はない。許容範囲についてもあらかじめ与えられていると仮定する。

本稿では TCW と許容範囲が与えられる例として講義の撮影をとりあげる。TCW については、例えば講師が黒板等を指示している時には、指示している内容も映像に含まれるように講師の位置が指定される。また講師が一定方向に移動している場合には、移動方向前方に余白があるように位置指定が行われる。許容範囲については、講師の移動傾向について事前知識が必要である。例えば、その場で小刻みに動いたり揺れることの多い講師については許容範囲を大きくとることで、揺れの映像に対する反応を抑えることができ、一度動き始めるとそのまま動く講師については許容範囲を小さくとることで追跡撮影の的確に移ることができる。

提案する手法では撮影カメラから入力画像が得られる度に以

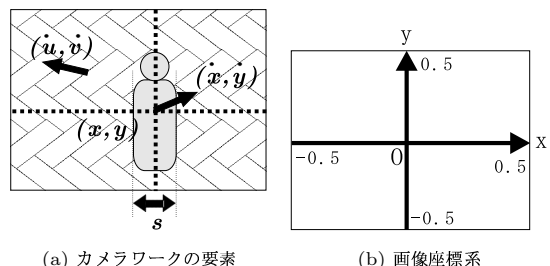


図 1 設定

下の処理を繰り返し行う。この処理を処理単位と呼び一回の処理単位をループと表記する。まず次ループ以降での SSP を予測する。予測された次ループ以降の SSP を「予測撮影状態パラメータ (ESP)」と呼ぶ。ESP が TCW に対して許容範囲内であれば、カメラ動作を修正しない。許容範囲外である場合には、修正後できるだけ長いループの間、ESP が許容範囲内となるようにカメラ動作を修正する。

以上の処理を実現するために、システムは、オプティカルフローに対して M 推定を適用することによって対象領域を分離する。また ESP を求めるためにカルマンフィルタを用いる。

この手法では逐次、カメラの制御を行うため、予測処理にかけられる時間は制限される。また、観測データを用いた予測は先の時刻になるほど誤差が増大していく。このため予測を何ループ先まで行うかを制限する。このループ数を最大予測ループ数 L とする。

本稿では、次章において提案するシステムの概要を述べた後、第 3 章では、アルゴリズムの詳細について述べる。第 4 章では、本システムの有効性を検証するために行った実験について述べる。第 5 章は本稿のまとめである。

2. 動作修正を抑制するカメラ制御

2.1 環境

単一の動物体を単一のカメラが撮影している場合を想定する。撮影カメラは投影中心と回転中心が一致するピンホールカメラとしてモデル化する。カメラは、パン・チルト速度 \dot{P} , \dot{T} と焦点距離 F からなるカメラ制御パラメータを指定することによって制御される。つまりカメラ制御パラメータの修正がカメラ動作の修正となる。これらのパラメータは離散値をとる。我々は、特に講義室での講義者のような屋内での人物の撮影に注目しているため、対象は未知の剛体運動をし、対象の見かけは大きくは変わらないと仮定する。

カメラワークは各フレームでの画像上の対象の位置 (x, y) 、速度 (\dot{x}, \dot{y}) 、大きさ (s) 、大きさ変化 (\dot{s}) 、および背景の速度 (\dot{u}, \dot{v}) からなるベクトルで表記される (図 1(a))。画像上での座標を図 1(b) のようにとる。

2.2 制御の流れ

提案手法を実現するための制御の流れを図 2 に示す。

繰り返し処理の 1 ループは観測、予測、制御からなる。観測段階では、SSP を観測する。観測された SSP には誤差が含まれるため「観測撮影状態パラメータ (OSP)」と別に名付けて

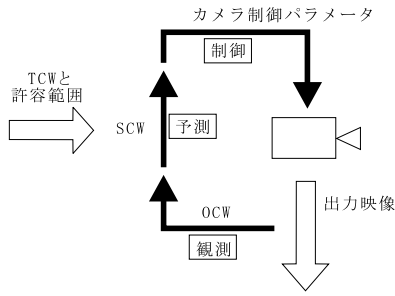


図2 制御の流れ

呼ぶ。予測段階では、現在とそれ以前の OSP を基にして、 L ループ先までの ESP を求める。 L ループ先までの ESP を安定して求めるために繰り返し処理に適したカルマンフィルタを用いる。制御段階では、ESP を基にして L ループにわたって許容範囲を逸脱せず、かつカメラ動作の修正を抑制するようにカメラ制御パラメータを修正する。

3. ループ内での処理

前章で述べた処理を実現するためには、OSP と ESP を求めることが必要である。本章ではこれらの実現方法とカメラ制御パラメータの決定方法について述べる。

3.1 OSP の算出

OSP を求めるためには対象領域を分離する必要がある。撮影カメラのパン・チルト・ズームにより背景が変化するため、背景差分は対象領域の分離に適していない。

そこでオブティカルフローを用いることで対象領域の分離を行う。オブティカルフローは水平方向成分、垂直方向成分、拡大・縮小成分からなるオブティカルフロー成分で特徴づけることができる。同一領域に属する点は同一の運動をするため同一のオブティカルフロー成分を持つ。領域は対象領域、背景領域の高々2つであるため、画像全体でオブティカルフロー成分が同一となるか、2種類のオブティカルフロー成分が得られるかのいずれかである。

2種類のオブティカルフロー成分が得られる場合は対象が背景に対して移動している場合である。この場合、オブティカルフロー成分を基にオブティカルフローを2つのクラスに分類し、一方を選択することによって対象領域を分離することができる。分類はオブティカルフローにM推定を適用し、オブティカルフロー成分を求めることによって行う。1度目のM推定でオブティカルフロー成分が得られる支配的な領域を第1領域、他方の領域を第2領域と呼ぶことにする。選択は、前のループにおけるカメラ制御パラメータから背景領域のオブティカルフロー成分を予測し、各領域から算出されたオブティカルフロー成分とを比較することによって行う(3.1.1節参照)。対象領域が分離できれば OSP を求めることができる。

画像全体でオブティカルフロー成分が同一である場合は、対象が背景に対して移動していない場合である。このときは対象領域を分離することができず速度以外は OSP を求められないが、求められない要素については直前のループで求められた現ループに対する ESP を用いる。

3.1.1 対象領域の分離

対象領域の分離は以下のように行う。

フローの算出 画像上の座標 (x_i, y_i) ($i = 1, \dots, N_{opt}$) でのオブティカルフローをブロックマッチングにより算出する。 N_{opt} はオブティカルフロー算出点の個数を表す定数である。

オブティカルフロー成分の推定 得られたオブティカルフローに対してM推定を適用し、第1領域のオブティカルフロー成分を推定する。

カメラは並進運動せず光軸方向の回転もしないため、点 (x_i, y_i) におけるオブティカルフロー (u_i, v_i) はオブティカルフロー成分 (H, V, S) を用いて次式のように表せる。

$$u_i = H + Sx_i, \quad v_i = V + Sy_i \quad (1)$$

ロバスト推定の手法の一つであるM推定を用いて第1領域のオブティカルフロー成分 (H_F, V_F, S_F) を得るためには以下の関数を最小化すればよい。

$$\sum_{i=1}^{N_{opt}} \rho(\|u_i - H_F - S_F x_i\| + \|v_i - V_F - S_F y_i\|) \quad (2)$$

ここで重み関数 ρ として Lorentzian 関数を用い、

$$\rho(z) = \log \left(1 + \frac{1}{2} \left(\frac{z}{\sigma} \right)^2 \right) \quad (3)$$

パラメータ σ として $\sigma = \frac{\Theta_{seg}}{\sqrt{2}}$ を用いた [6]。

第1領域の分類 $|H_F + S_F x_i - u_i| < \Theta_{seg}$ かつ $|V_F + S_F y_i - v_i| < \Theta_{seg}$ の場合、 (x_i, y_i) を第1領域として分類する。ここで、 Θ_{seg} は分類に関する閾値である。

第2領域の存在判定 N_F を第1領域に分類された点の数として、 $\frac{N_F}{N_{opt}}$ と閾値 Θ_{reg} を比較する。

(1) $\frac{N_F}{N_{opt}} > \Theta_{reg}$ である場合、他に異なる運動をしている領域がないため、対象は背景に対して移動していないものとみなして処理を終了する。

(2) (1) 以外の場合、第2領域のオブティカルフロー成分を推定するために、第1領域に分類されなかった点に対して再びM推定を適用する。第2領域のオブティカルフロー成分を (H_S, V_S, S_S) と表す。

対象領域の選択 背景領域として予期できるオブティカルフロー成分 (H_r, V_r, S_r) は前ループでのカメラ制御パラメータの値から以下のように計算する。

$$H_r = -F_{n-1} \dot{P}_{n-1} t \quad (4)$$

$$V_r = -F_{n-1} \dot{T}_{n-1} t \quad (5)$$

$$S_r = \frac{F_{n-1} - F_{n-2}}{F_{n-2}} \quad (6)$$

ここで、 n は第 n ループでの値であることを表し、 t は1ループにかかる時間である。

対象領域であるかどうかを判別するための評価基準として以下の式を用いる。

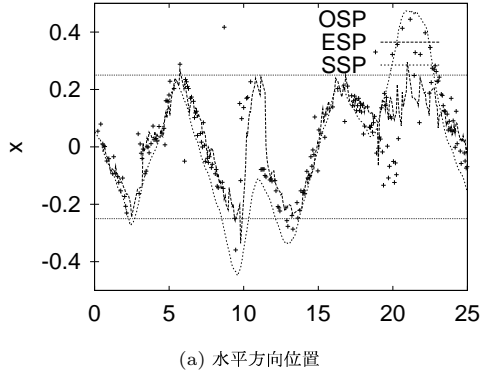
$$\left\| \frac{H - H_r}{H_r} \right\| + \left\| \frac{V - V_r}{V_r} \right\| + \left\| \frac{S - S_r}{S_r} \right\| \quad (7)$$

この評価基準が小さいオブティカルフロー成分を持つ領域の方を背景領域とみなし、そうでない方を対象領域として選択する。

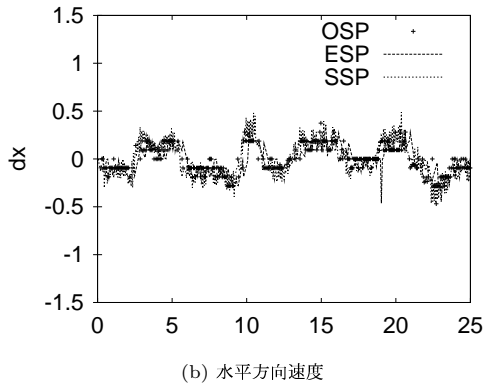


(a) 対象 小サイズ (b) 対象 中サイズ

図3 対象の大きさ



(a) 水平方向位置



(b) 水平方向速度

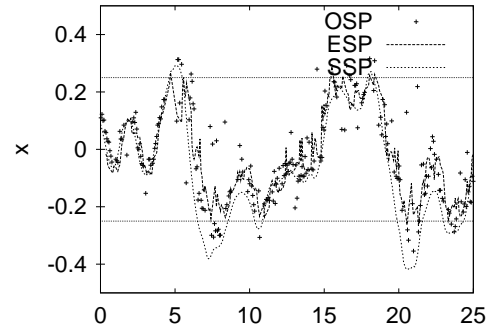
図4 位置許容範囲 ± 0.25 ・対象 小サイズ

的な制御は行わず、2種類の対象の大きさに対して F_n を固定し実験を行った。垂直方向の制御は水平方向の制御と同じであるため実験を省略した。対象の大きさを図3に示す。

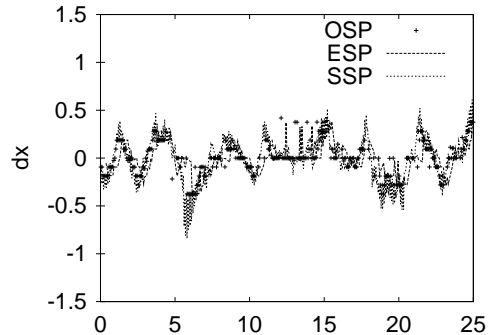
以下いずれのグラフにおいても OSP は OSP を、ESP は 0.33 秒 (5 ループ) 前に求められた ESP を表す。SSP は、色抽出によって求めた値を実際の対象の位置もしくは速度として採用した。横軸は時刻 (秒)、 x 座標については図 1(b) の座標系である。

図 4~7 は対象についてそれぞれ大きさを小、 x 座標の許容範囲を $-0.25 \sim 0.25$ とした場合、大きさを中、 x 座標の許容範囲を $-0.25 \sim 0.25$ とした場合、大きさを中、 x 座標の許容範囲を $-0.15 \sim 0.15$ とした場合大きさを中、 x 座標の許容範囲を $-0.25 \sim 0.25$ 、かつ x 方向速度の許容範囲を $-0.15 \sim 0.15$ とした場合の結果である。図中の横線は設定された許容範囲を表す。また、各々の場合について \dot{P}_n の修正回数と修正量の絶対値の総和を図 8 に示す。

ESP に基づくカメラ制御を行っているため、ESP ほどの場

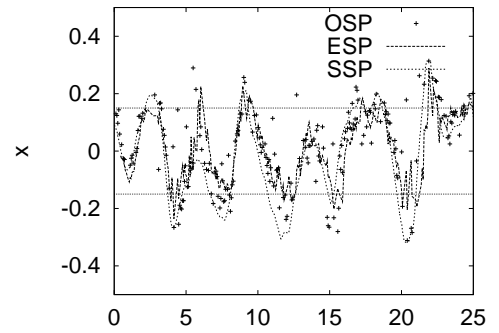


(a) 水平方向位置

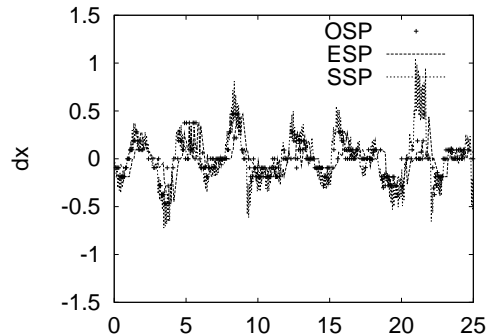


(b) 水平方向速度

図5 位置許容範囲 ± 0.25 ・対象 中サイズ



(a) 水平方向位置



(b) 水平方向速度

図6 位置許容範囲 ± 0.15 ・対象 中サイズ

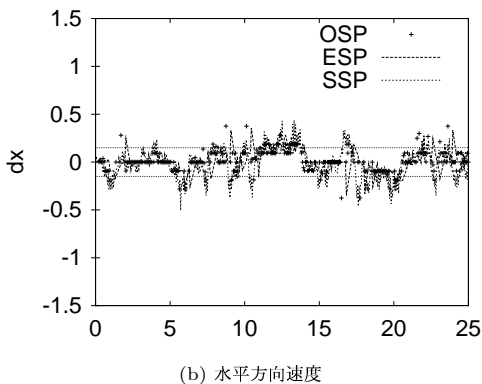
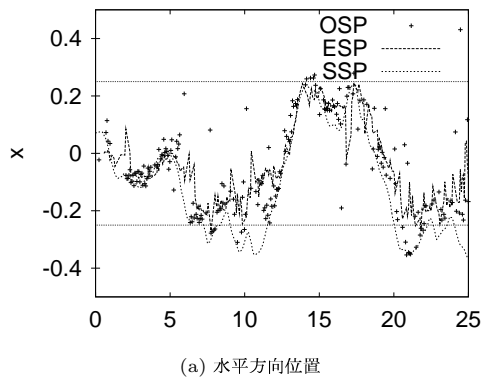


図7 位置許容範囲 ± 0.25 + 速度許容範囲 $\pm 0.15 \cdot$ 対象 中サイズ

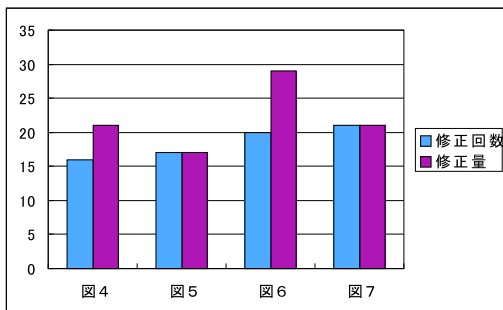


図8 カメラ動作の修正

合でもほぼ指定された許容範囲内におさまっている。しかし SSP については求められた ESP からはずれてしまい許容範囲外に出てしまっている部分がある。

図4と図5を比較した場合、対象が小さく映っている場合には、画像上での対象の速度も小さくなり制御が容易となるにも関わらず SSP が大きく外れている場合がある。これは対象の大きさが小さいため、対象領域の分離を誤判定し OSP も誤った値となったためだと考えられる。改善方法としては、オプティカルフローに加えて、オプティカルフロー計測点の位置関係や、他の画像特徴を用いることなどが考えられる。

また、どの場合でも等速度運動をしていると見なせる部分では、ESP と SSP とが近く、他の場合にずれが大きいことから、等速度運動によるモデル化のため速度変化に対応できていないことが考えられる。しかし、等加速度運動によるモデル化によって実験を行った場合、加速度項に対する誤差が他の項に

大きな影響を与え、予測結果が大きく変動し、結果として、カメラの急激な加速により対象がフレームアウトするなどの問題が発生した。詳細なモデルの導入には OSP の算出を改善し誤差を抑える必要がある。

図5と図6について許容範囲が狭い方が実際にカメラ動作の修正頻度が大きくなり、かつ、カメラ制御パラメータの修正量が大きくなる傾向があることが図8からわかる。また、図5と図7を比較すると、速度の許容範囲を指定した場合、実際に水平方向速度が制限されているが、制約が追加されたことによりカメラ動作の修正が増える傾向にあることが図8から分かる。カメラ動作修正の抑制とのトレードオフを考えて、許容範囲を指定する必要がある。

5. 結論

本稿では動物体を首振りカメラで撮影する際に、カメラ動作修正を抑制しつつ指定されたカメラワークを実現する方法について提案した。

我々のアプローチでは、カメラ動作の変更を抑制するために許容範囲を導入した。各ループにおいて、次ループの ESP を求め、許容範囲内であれば、カメラ制御パラメータを変更せず、許容範囲外である場合のみ、最大 L ループ先までの ESP が全て許容範囲内であり続けるようにパラメータを修正する。

提案手法の有効性を示すため2種類の対象の大きさに対して位置、および位置と速度の両方に許容範囲を設定した場合について実験を行った。実際に許容範囲の条件が緩くなるほどカメラ動作の変更が抑制されていることが示された。主に対象領域の分離の誤りによる OSP のばらつきや誤りのため SSP が許容範囲外に出てしまうことがあることが、問題点としてあげられる。

今後の課題として、対象領域の分離処理の改善、実際に対象の大きさの制御を行うことがあげられる。また、幾何以外の要素へ発展させ、例えば色や明るさ・コントラストなどの光学的な要素まで含めて制御を行うことも考えられる。

文献

- [1] Q. Cai and J.K. Aggarwal, "Real time tracking for enhanced tennis broadcasts", Proc. CVPR'98, pp.68-72, 1998.
- [2] Q. Liu, Y. Rui, A. Gupta and J.J. Cadiz, "Automating camera management for lecture room environments", ACM CHI 2001, pp.442-449, 2001.
- [3] S.A. Hutchinson, G.D. Hager, and P.I. Corke, "A tutorial on visual servo control", IEEE Trans. Robot. Automat, vol.12, no.5, pp.651-670, 1996.
- [4] 足立淳, 佐藤淳 "未校正カメラによる未校正ロボットの視覚サーボ", 信学論 (D-II), Vol.J83-D-II, no.4, pp.1110-1118, January 2000.
- [5] 尾関基行, 中村裕一, 大田友一 "プレゼンテーションの知的撮影システム—手元作業を対象とした適応的カメラワーカー", 信学技報, PRMU2000-104, pp.31-38, 2000
- [6] M.J. Black, "The robust estimation of multiple motions", Computer Vision and Image Understanding, vol.63, no.1, pp.75-104, 1996.
- [7] 谷崎久志, "状態空間モデルの経済学への応用", 日本評論社, 1993