

リニアスライダを用いた高速ステレオビジョンシステム ～左右カメラの見えの違いを考慮した最適基線長探索～

小出 智治[†], 竹内 義則^{††‡}, 中坊 嘉宏^{‡‡}, 向井 利春[‡], 大西 昇[†]

[†] 名古屋大学 大学院情報科学研究科 メディア科学専攻, ^{††} 名古屋大学 情報連携統括本部 情報戦略室
[‡] 理化学研究所 バイオ・ミメティックコントロール研究センター, ^{‡‡} 産業総合技術研究所

本稿では、高速ステレオカメラをリニアスライダ上に設置したシステムを用いて、奥行き計測誤差とステレオ対応誤りの双方を考慮した最適な基線長をリアルタイムに探索する手法を提案する。基線長制御は、奥行き計測誤差とステレオ対応誤りの見積り線の線形和からなるコスト関数を最小化することで行われる。コスト関数の最小化には最急降下法を用いる。実験により、提案した基線長決定方法によって最適な基線長が見つかることを確認した。また、見つかった最適な基線長においては、左右カメラの見えの違いによって増加すると考えられる非類似度 (SAD) の減少が見られた。

High-Speed Stereo Vision System Using Linear Slider - Optimizing Baseline Length Considering Difference Between Two Views -

Tomoharu KOIDE[†], Yoshinori TAKEUCHI^{††‡}, Yoshihiro NAKABO^{‡‡}, Toshiharu MUKAI[‡],
Noboru OHNISHI[†]

[†] Graduate School of Information Science, Nagoya University

^{††} Planning Office of Information and Communication Strategy, Information and Communications
Headquarters, Nagoya University

[‡] BMC Research Center, RIKEN

^{‡‡} AIST

In this paper, we present a method to optimize a baseline length of a stereo vision system considering both errors of depth measurement and failures of stereo matching. Baseline control is done by minimize cost function which consists of the errors of the depth measurement and the estimated failures of the stereo matching. Steepest descent method is used to minimize cost function. Experimental results show that the proposed method can find the optimal baseline length and the dissimilarity measure of the stereo matching (SAD) decreased at the optimal baseline length.

1 はじめに

近年、3次元計測に対するニーズが様々な分野で高まっている。この技術は、コンピュータビジョンの分野ではロボットの周辺視システムや動画画像解析で、またコンピュータグラフィックスの分野ではCADやバーチャルリアリティで必要とされている。

よく使われる3次元計測の手法として、ステレオビジョンがある。しかし、従来のステレオビジョンシステムでは、カメラ位置が固定されているものが多く、カメラ視野内で対応点の取れる範囲が限定される。そのため対象物体がその範囲を超えて移動する状況では、3次元計測は不可能である。しかし、現実世界では移動物体を含むシーンの方

が多く見られる。例えば、サッカーや陸上競技などのスポーツで、ボールや人物を追跡しつつ3次元位置も得たいといった、テレビ撮影の拡張やバーチャルリアリティで需要がある。よって、対象物体が移動している状況でも、3次元計測できるステレオビジョンシステムが必要である。

可変基線長でのステレオビジョンの研究として、Okutomiらの研究⁷⁾、William Klarquistらの研究⁸⁾がある。前者の研究では、複数のカメラを一行に並べていくつもの基線長の異なるカメラ対を作り出し、それらを統合することでステレオ視の奥行き精度を向上させた。後者の研究では、画像の空間周波数によって基線長を制御し、パターンテクスチャがあるシーンにもあいまいなく対応付けをし、精度のよい奥行きマップを生成している。

また、リニアスライダを用いた研究として、服部らの研究がある⁵⁾。この研究では、高速なステレオビジョンシステムとリニアスライダを用いて、対象に追従しながらその重心位置の三次元位置計測を行う。さらに、対象が画像面に大きく投影されている場合には加えて形状復元も行うシステムを開発した。形状復元時の基線長制御では、奥行き計測の誤差のみを考慮し、対象がカメラから外れない範囲で基線長を最大にとっている。

移動する対象を追跡する場合、対象の姿勢も変化していくと考えるのが普通である。その場合、カメラに対して対象が常に同じ面を向けているとは限らない。その時、基線長が大きいと、対象の形状によっては左右のカメラで見えの違いが大きくなり、ステレオ対応がとりにくくなる状況が生じる。そこで、本研究では、左右カメラの見えの違いによる対応誤りを考慮した基線長制御を行う。ステレオカメラをリニアスライダ上に設置することで、カメラの基線長を変更しながら、形状復元を行えるようにする。

2 システム構成

2.1 高速ステレオビジョンシステム

本研究で使用した超高速ビジョンシステムは石川らによって提案、開発され²⁾、浜松ホトニクス社と理化学研究所との共同研究で開発している超高速インテリジェントビジョンシステム(CPV-I:Column Parallel Vision System)である^{1) 4)}。CPVシステムの構成を図1に示す。

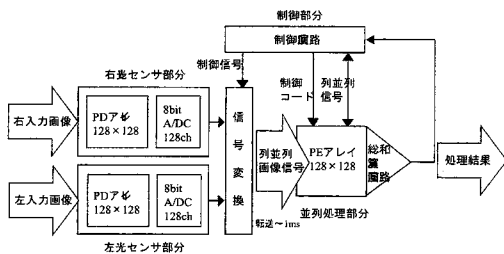


Fig. 1 CPVシステムの構成

システムは、光センサ部分・並列処理部分・制御部分から構成される。左右二つの光センサ部分からフレームレート1kHzで128pixels×128pixelsの8bit時系列デジタルデータが出力される。そして、128pixels×128pixels分用意された並列処理

部分に転送し、1フレーム毎に汎用画像処理を行い、その結果を出力する。制御部分ではシステム全体の制御、ホストPCとのインタフェースを行う。CPVシステムは、ホストPCで開発された実行プログラムを制御部分にアップロードし、その実行プログラムにより画像処理演算を行う。並列処理部分はS³PEアーキテクチャに基づいたSIMD型並列演算子(PE)から構成され、PDアレイの各画素に1対1に対応したPEを用意することで完全並列な演算を実現している。

3 高速リニアスライダ

本研究で使用した高速リニアスライダはTHK社製のリニアモータアクチュエータGLM形である。仕様を表1に示す。

Table 1 高速リニアスライダの仕様

項目	仕様
全長	3600mm
最高速度	2m/sec
リニアエンコーダ分解能	1μm
繰り返し位置決め精度	1μm

ホストPCとはINTERFACE社製PCIボードPCI-6205C(エンコーダボード)とPCI-3338(DAボード)で接続し、リニアスライダのエンコーダ値をカウントし、電圧指令値を出力する。

4 基線長制御

4.1 奥行き計測誤差の最小化⁵⁾

ステレオマッチングを行う時、得られる奥行き解像度と奥行き誤差は基線長により影響を受ける。ここでは、カメラ配置は標準ステレオで、物体を正面から見据える状況を想定し、カメラ1画素分の大きさによる誤差の影響を考える。

図2から三角測量による対象奥行き計算を行うと、

$$D = L \frac{\sin(\tan^{-1}(\frac{f}{x_l})) \sin(\tan^{-1}(\frac{f}{x_r}))}{\sin(\tan^{-1}(\frac{f}{x_l}) + \tan^{-1}(\frac{f}{x_r}))} \quad (1)$$

$$= \frac{fL}{x_l + x_r} \quad (2)$$

となる。

カメラ1画素分の大きさを p とすると、対象は図2上部の四角領域の内部に存在することになる。

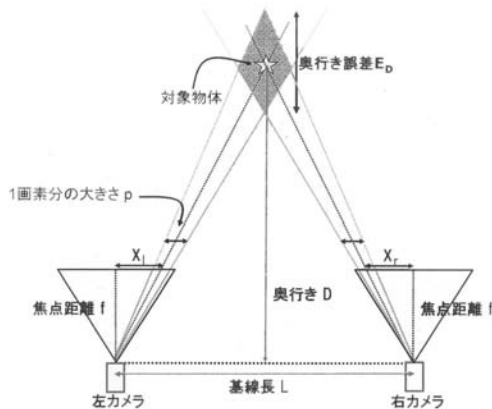


Fig. 2 カメラ 1 画素分の量子化誤差の影響

この領域内に対象がある限り、カメラ画像上では同じ画素に対象が映る。この四角領域の奥行き幅が、カメラ 1 画素分の奥行き誤差 E_D となる。

標準ステレオのカメラ姿勢で物体を正対視できる場合、 $x_l = x_r (= x)$ である。図 2 から、平面幾何より、 E_D を計算すると、

$$x = \frac{fL}{2D} \quad (3)$$

$$E_D = \frac{L}{2} \left(\frac{f}{x - p/2} - \frac{f}{x + p/2} \right) \quad (4)$$

と表すことができる。

式 (4) に $p = 0.04\text{mm}$ を代入してグラフにすると図 3 のようになる。このように、基線長が長いほど、奥行き誤差範囲は小さくなる。

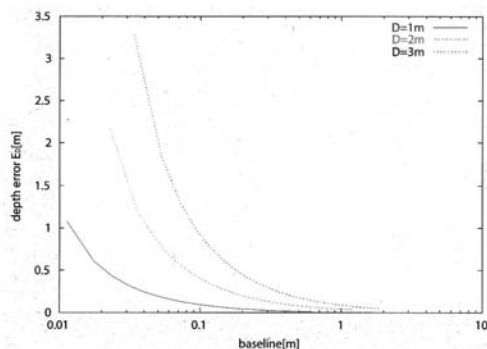


Fig. 3 基線長と奥行き誤差範囲の関係

4.2 ステレオ対応誤りの最小化

4.2.1 対応誤りの定義

ステレオマッチングを行う際、カメラと対象との位置関係や対象の形状によっては、「一方のカメラには写っているが、もう一方のカメラには写っていない領域」が生じる場合がある。このような領域では、対応する領域を見つけることができず、誤った対応付けが行われる。また、写っていたとしても、見えている面の画像上での大きさが左右のカメラで異なるときには、対応付けされない可能性もある。

ステレオマッチングの評価値としては、各点の輝度値から局所的な平均値を引き、分散値の類似度でマッチングを取る正規化相互相関 (NCC: Normalized Cross Correlation) や、輝度値の二乗残差の和である SSD (Sum of Squared Difference)、あるいはさらに単純に輝度の絶対残差の総和をとる SAD (Sum of Absolute Difference) などがある。

本稿の手法では、ビジョンシステムの特性を生かすため、ステレオマッチングの評価値として SAD (Sum of Absolute Difference) を用いる。探索ウィンドウ W 上の、SAD は次式で計算される。

$$\text{SAD}(x, y, d_x, d_y) = \sum_{i \in W} \sum_{j \in W} |f_l(x+i, y+j) - f_r(x+i+d_x, y+j+d_y)| \quad (5)$$

$\text{SAD}(x, y, d_x, d_y)$ は左画像座標 (x, y) の視差 (d_x, d_y) における SAD 値、 $f_l(x, y)$ は左画像の座標 (x, y) の画素値を、 $f_r(x, y)$ は右画像の座標 (x, y) の画素値を表している。すべての左画像点 (x, y) に対して、探索範囲内の各視差について SAD を求め、それが最小となる視差を、その座標点の視差として決定する。

SAD 値は、比較する領域同士の輝度が似ているほど値が小さくなり、逆に似ていなければ値が大きくなる。対応が取れていない領域では、対応する領域が存在しないために、探索によって見つかった最小 SAD 値が、対応が取れている領域よりも大きくなる傾向があると考えられる。そこで、以下のような手順で求められる特徴量 R を考える。

1. 探索によって見つかった SAD 値を各画素について記憶しておき、ある閾値 th_{SAD} より高い最小 SAD 値をもつ画素を数える。

2. 上の画素数が画像中の対象の画像上の面積に対して占める割合を R とする.

R は対応誤りとともに大きくなると考えられるため, 本稿ではこの R を $E_C(L)$ の定義として用いることにした.

$$E_C \equiv R \quad (6)$$

4.2.2 E_C の最小化

対応誤り E_C は基線長が広がるにつれて大きくなる傾向がある. これは, 基線長が広がることによって左右の画像の見えの違いが大きくなっていくからであると考えられる. これを減らすためには基線長を短くし, 左右カメラの見えの違いを小さくする必要がある.

4.3 制御則

4.1 節での最小化は基線長を広げることで行われるが, 基線長を広げると対応誤り E_C が増える可能性がある. 一方, 4.2 節の最小化は基線長を縮めることで行われるが, 基線長を縮めると E_D が増加する. これらはトレードオフの関係となっており, 二つを同時に考慮しようとする場合, 最適化の手法が必要となる.

本研究では, 基線長制御を E_D と E_C の線形和からなるコスト関数 $E(L)$ の最小化と位置づけ, $E(L)$ が最小となる基線長が最適であると考え.

$$E(L) = w_D E_D(L) + w_C E_C(L) \quad (7)$$

$$L_{opt} = \operatorname{argmin}_L E(L) \quad (8)$$

ここで, w_D と w_C は, それぞれ, 奥行き計測誤差とステレオ対応誤りの重みである.

最適化の手法としては最急降下法を用いる. 制御ループ内で以下の式に従って基線長を更新する.

$$L(t+1) = L(t) - \alpha \frac{d}{dL} E(L(t)) \quad (9)$$

基線長の更新のためには, $\frac{d}{dL} E(L)$ を計算する必要がある. $\frac{d}{dL} E(L)$ を計算すると,

$$\frac{d}{dL} E(L) = w_D \frac{d}{dL} E_D(L) + w_C \frac{d}{dL} E_C(L) \quad (10)$$

となる. この中の $\frac{d}{dL} E_D(L)$ は解析的に

$$\begin{aligned} \frac{d}{dL} E_D(L) &= \frac{d}{dL} \left\{ \frac{L}{2} \left(\frac{f}{\frac{fL}{2D} - p/2} - \frac{f}{\frac{fL}{2D} + p/2} \right) \right\} \quad (11) \\ &= \frac{d}{dL} \left\{ \frac{L}{2} \left(\frac{2Df}{fL - Dp} - \frac{2Df}{fL + Dp} \right) \right\} \quad (12) \\ &= \frac{1}{2} \left(\frac{2Df}{fL - Dp} - \frac{2Df}{fL + Dp} \right) \quad (13) \\ &\quad + \frac{L}{2} \left(\frac{2Df^2}{(fL + Dp)^2} - \frac{2Df^2}{(fL - Dp)^2} \right) \quad (14) \end{aligned}$$

と求められる.

しかし, E_C は対象依存の関数であり, 基線長が広がるにつれて大きくなるのが予想できるものの, $\frac{d}{dL} E_C(L)$ はあらかじめ求めることができない. そこで, $\frac{d}{dL} E_C(L)$ の代替となる量を考える必要がある.

4.2.1 節で求めた R を用いて, $\frac{d}{dL} E_C(L)$ を以下のように定義する.

$$\frac{d}{dL} E_C(L) \equiv \begin{cases} \beta \cdot (R - th_{rate}) & (R \geq th_{rate}) \\ 0 & (\text{otherwise}) \end{cases} \quad (15)$$

ここで β は正の定数で, th_{rate} は R の閾値である.

この定義は, $\frac{d}{dL} E_C(L)$ を「 R が閾値 th_{rate} を超えている場合, R が閾値からどれだけ離れているかに比例して大きくなる関数」とする. これは対応誤りに許容値を設け, それよりも誤りが小さい場合は奥行き計測誤差のみを考慮するように調整を加えるものである.

5 実験

本研究で提案した手法の有効性を検証するため, 基線長最適化実験を行った. 基線長を固定した状態で対象をカメラの前に正対視できるように置き, 提案した手法に従って基線長を最適化する. 基線長が最適化される様子を計測し, 検証した. 対象は直径 130mm で高さが 210mm の円柱, 高さ 290mm の鋭角な二面を持つ物体, 直径 125mm の球である. 実験時の配置は図 4 の通りであり, 鋭角な物体は角が二つのカメラの間点に向くように置いた. 開始時の基線長は最短の 185mm と最長 (見失わない範囲で) の 300mm の二つで実験を行った. 紙面の制約上, 結果は円柱と鋭角な物体のみを示す. カメラの配置は標準ステレオ配置とし, SAD はエピポーラ線 (水平線) 上を計算した. SAD のウィンドウサイズは 5×5 とした.

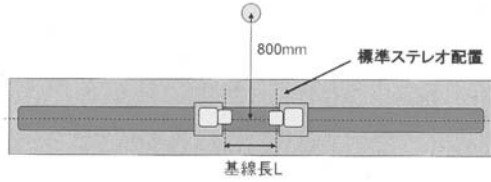


Fig. 4 実験時の配置

5.1 基線長最適化

最適化後の基線長とそのときの E_D , R の値を表 2 に示す. 最適化に用いたパラメータは表 3 に示す.

表 2 の min は最長基線長, max は最大基線長を表し, opt_{min} , opt_{max} は, それぞれ最小基線長から開始して得られた最適基線長, 最大基線長から開始して得られた最適基線長を表す.

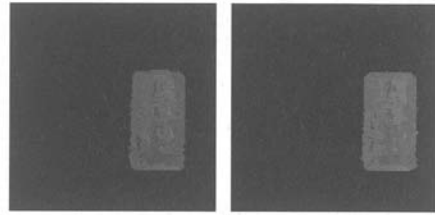
Table 3 最適化の各パラメータ

物体	円柱	鋭角な物体
w_D	0.5	0.5
w_C	0.5	0.5
th_{SAD}	550	500
th_{rate}	0.1	0.02
α	0.75	0.75
β	3.0	0.75

最適化された基線長においては, 最小基線長よりも E_D が小さく, 最大基線長よりも E_C が小さい. よって, E_D と E_C の両方を考慮した基線長が選択されたといえる. opt_{min} と opt_{max} が異なっているのは, $\frac{d}{dL}E_C(L)$ の定義で R という観測値を用いるようにしているため, 観測誤差が存在することが原因だと思われる.

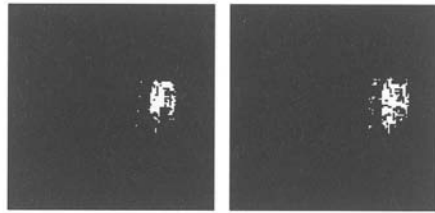
5.2 最適基線長の評価

最適とされた基線長における形状復元の結果を, 従来の奥行き計測誤差のみを考慮した場合における形状復元結果と比較, 検証した. 図 5, 図 7 に, 視差画像を示す. 本稿で用いたシステムは左画像基準のステレオマッチングを行っている. また, 対応点の持つ最小 SAD 値が最適化で用いた th_{SAD} を超えている画素を分布を図 6, 図 8 に示した. 図 7 から, 鋭角な物体の左面において視差の変化が



(a) 最適基線長 (b) 最大基線長

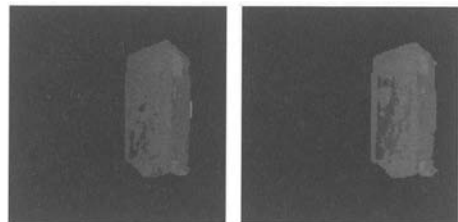
Fig. 5 円柱の視差画像



(a) 最適基線長 (b) 最大基線長

Fig. 6 円柱の最小 SAD 値が th_{SAD} を超えている画素

乱れていた部分が滑らかになっている箇所があった. これは基線長が短くなったことにより, 左右のカメラでの見えの違いが小さくなり, 正しく対応が取れる点が増えたことが原因として考えられる. また, 図 6, 図 8 を見ると, 高い最小 SAD 値を持つ画素の数は最適とされた基線長のほうが少ない. これは前節で最大基線長よりも R が小さくなっていることから明らかである. これは基線長が近づくことによって左右のカメラの見えの違いが小さくなったためであると考えられる.



(a) 最適基線長 (b) 最大基線長

Fig. 7 鋭角な物体の視差画像

Table 2 最適化後の各値

物体	円柱				鋭角			
	min	opt _{min}	opt _{max}	max	min	opt _{min}	opt _{max}	max
基線長 [mm]	185	242	250	300	185	228	248	300
E_D [mm]	42.5	31.7	30.7	26.0	55.3	45.4	41.1	33.0
R	0.082	0.140	0.142	0.174	0.059	0.086	0.075	0.163



(a) 最適基線長 (b) 最大基線長

Fig. 8 鋭角な物体の最小SAD値が th_{SAD} を超えている画素

6 おわりに

本稿では、高速ステレオカメラをリニアスライダ上に設置したシステムを用いて、奥行き計測誤差とステレオ対応誤りの双方を考慮した最適な基線長を探索する手法を提案した。実験により、提案した基線長決定方法によって最適な基線長が見つかることを確認した。また、探索された最適な基線長においては、ステレオマッチングによって見つかった対応点の持つ最小SAD値が減少する傾向が見られることがわかった。今後の課題としては、 $\frac{d}{dL}E_C$ の新たな計算方法の考案や、基線長制御と平行した対象の追跡の実装などがあげられる。

参考文献

- 1) Yoshihiro Nakabo, Idaku Ishii, Masatoshi Ishikawa, *3D Tracking Using Two High-Speed Vision Systems*, Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS2002), pp.360–365, 2002.
- 2) M. Ishikawa, K. Ogawa, T. Komuro, I. Ishii, *A CMOS Vision Chip with SIMD Processing Element Array for 1ms Image Processing*, IEEE Int. Solid-State Circuits Conference (ISSCC'99), pp.206–207, 1999.
- 3) 豊田 晴義, 向坂 直久, 水野 誠一郎, 中坊 嘉宏, 石川 正俊, 超インテリジェントビジュアルシステム (CPV-II) を用いた動画像特徴量抽出方法の検討, 第9回画像センシングシンポジウム講演論文集, pp.289–284, 2003.
- 4) 豊田 晴義, 向坂 直久, 水野 誠一郎, 中坊 嘉宏, 石川 正俊, 超インテリジェントビジュアルシステム:CPV(Column Parallel Vision system), 第6回画像センシングシンポジウム講演論文集, pp.213–216, 2000.
- 5) 服部 雄介, 中坊 嘉宏, 向井 利春, 竹内 義則, 大西 昇, 追従型高速ステレオビジョンシステムを用いた移動物体の形状復元, 電子情報通信学会技術研究報告, PRMU2003-283, pp.161-166, 2004.
- 6) Yoshihiro Nakabo, Toshiharu Mukai, Yusuke Hattori, Yoshinori Takeuchi, and Noboru Ohnishi, *Variable Baseline Stereo Tracking Vision System Using High-Speed Linear Slider*, IEEE Int. Conf. Robotics and Automation (ICRA 2005), pp.1579-1584, 2005.
- 7) Masatoshi Okutomi, Takeo Kanade, *A Multiple-Baseline Stereo*, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.15, No.4, pp.353–363, 1993.
- 8) William Klarquist, Alan Bovik, *Adaptive Variable Baseline Stereo for Vergence Control*, IEEE Int. Conf. on Robotics and Automation, pp.1952–1959, 1997.