

ステレオ画像からのアフィン座標系に基づく 任意視点からの見え方生成の検討

Kuntal Sengupta 坂口竜己 大谷淳

(株) ATR 知能映像通信研究所
〒619-02 京都府相楽郡精華町光台2-2
TEL:0774-95-1401 E-mail:{kuntal,tatsu,ohya}@mic.atr.co.jp

本報告では、ステレオ画像から得られる2次元アフィン座標系に基づき、任意視点から見た画像を合成する手法について述べる。3次元空間中のシーンを撮影した2次元の画像を基に任意視点から見た画像を合成するためには、撮影されている対象の3次元構造を再構築する手法が一般的である。しかし、3次元シーンの再構築には正確なカメラキャリブレーションが必要となり、この時点で発生する誤差が生成されるシーンの品質に大きく反映してしまう。

そのため本報告では、より直接的な手法として、生成する画像と基になるステレオ画像の対応関係を推定するために、カメラキャリブレーションとして与える5点の対応点のアフィン座標値を用いた。この手法は実画像とCGの融合や、画像のシームレスな結合等の他のアプリケーションにも応用可能である。

A Study of Novel View Generation from Stereo Images Using 2D Affine Space

Kuntal Sengupta Tatsumi Sakaguchi Jun Ohya

ATR Media Integration & Communication Research Laboratories
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto, 619-02, Japan
TEL: +81-774-95-1401 E-mail: {kuntal,tatsu,ohya}@mic.atr.co.jp

We present an unified theoretical framework for novel view synthesis and other applications. To start with, we have a set of real images from weakly calibrated cameras, for which we compute the dense point match correspondences. For applications like novel view synthesis, one may first solve the 3D scene reconstruction problem, followed by a view rendering process. However, errors in 3D scene reconstruction usually gets reflected in the quality of the new scene generated, so we seek a more direct method. In this report, we use the knowledge of dense point matches and their affine coordinate values to estimate the corresponding affine coordinate values in the new scene. Our technique of reprojection is extended for other applications like merging real and synthetic worlds.

1. はじめに

近年、遠隔地間映像通信の重要性が増加しているが、現存するテレビ会議システムのような映像通信システムでは、ユーザ間での距離の隔たりを意識させないことが困難である。

この問題は、遠隔地間のユーザが一つの実世界に共存しているように感じる環境を創り出すことで解決できると考えられる[1]。このためには、任意の視点から見た実空間の視野映像を生成することが必要である。また、仮想現実感のアプリケーションにおいても、実物体と仮想物体を含んだシーンを構成することが求められている。本報告では、一つの理論的な枠組みの中でこれらの問題を取り扱う。

既に報告されているステレオ画像からの任意視野画像生成手法[2-5]では3次元再構築問題を、強カメラキャリブレーションを用いる手法、もしくは画像間の対応関係を3次元アフィン変換または透視投影変換と仮定する手法のいずれかにより解決している。その後3次元座標系で表されたオブジェクトを新しい画像に投影することで、任意視点から見た画像を生成するのである。しかし3次元構造再構築の過程における画像の測定ミスや、その他のあらゆる要因に対して不安定要素が多い。

また、シーンの再投影や新しい視野画像合成のためのより直接的なアプローチであるエピソードライン交差法等では、強カメラキャリブレーションと少なくとも3枚の画像間で8点の対応点が必要である[6,7]。

最近では2つの画像が与えられた場合、モーフィングを用いて、2つのカメラの光学中心を結ぶ線上にある任意の位置からの画像を再合成する手法[8]が示されたが、このモーフィングの考え方は任意位置の仮想カメラからの画像合成へと拡張することはできない。

本報告で述べるアフィン座標に基づいた新しい視野画像の生成法は非常に単純であり、物体認識手法として提案された手法[9]に基づいている。この文献では3次元剛体モデルを撮影することによって得られる2次元画像のペアは、高次 α - β アフィン空間上の2本の直線として最適に表現できると述べられている。しかしここではカメラの投影モデルを正射影として仮定しており、再合成された画像には劣化が見られる。

本報告では正射影に替えて、射影にピンホールカメラ系を採用し、ステレオカメラから撮影された2枚の画像間で対応する点におけるアフィン座標の特性を調べた。我々はこの特性を新しい視野画像の合成、実世界と仮想世界の融合等のアプリケーションのためにシーン再投影手法として用いている。

2. カメラキャリブレーションモデル

ここでは、強カメラキャリブレーションと弱カメラキャリブレーション[10,11]について簡単に解説する。

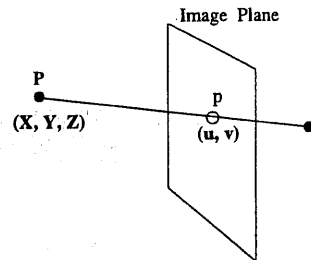


図1 絶対標定のカメラモデル

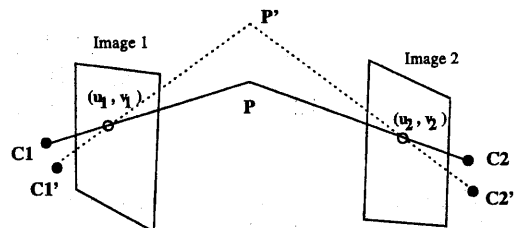


図2 3次元再構築

2. 1 強カメラキャリブレーション

図1において強カメラキャリブレーションに基づき空間中の点 P を画像平面中の点 p に投影する場合には、 3×4 の変換マトリクス C を計算する必要がある。点 P の同次座標系での座標が $[X Y Z 1]$ であり、 p の座標が (u, v) であるならば、

$$\begin{bmatrix} tu \\ tv \\ t \end{bmatrix} = C \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

ここで t はパラメータである。

変換マトリクス C の全体のスケーリングファクタは重要でない、キャリブレーションでは C の11個の要素を推定することとなる。すなわち、少なくとも6個の3次元座標中の点のマッピング情報を知識として、強カメラキャリブレーションパラメータは推定できる[12]。

図2に示すステレオペアが与えられ、双方のカメラでの強カメラキャリブレーションパラメータが既知のものとするならば、画像1中の点 (u_1, v_1) が画像2中の点 (u_2, v_2) に対応する場合には、その点の3次元座標を標準的な三角測量法を使ってキャリブレーションパラメータから推定することができる。これにより、この点は仮想カメラの投影行列を使って任意の画像中に再投影することができる。もしカメラが適切にキャリブレーションされていない場合には、これはつまりキャリブレーション過程における測定誤差となり、3次元再構築の際には図中点線の様に誤ってしまう。これは新しい視点からの画像を生成する際の再投影処理の誤りをも導く。

2. 2 弱カメラキャリブレーション

弱カメラキャリブレーションとは、2つのズレたステレオ画像間でのエピソード位置関係を求める手法である。

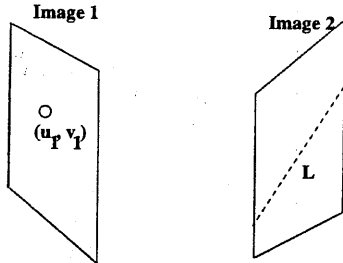


図3 エピポーラ・ライン

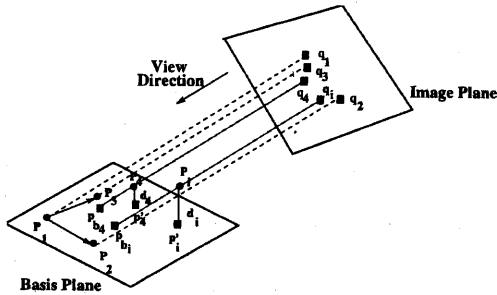


図4 平行投影の場合のアフィン座標

図3に示すように、 (u_1, v_1) が画像1中の点 p の座標を表すとする、 3×3 の行列 F を用いて、画像2中のエピポーラライン L を計算することができる。 p に対応した画像2中の点はこのエピポーラライン上に位置する。もし、エピポーララインが $h_1u_2 + h_2v_2 + h_3 = 0$ で表されるならば、その関係は以下の式で表される。

$$[h_1 \ h_2 \ h_3]' = F[u_1 \ v_1 \ 1]' \quad (2)$$

スケーリングファクタまで含めて F を解くためには、2枚の画像中で8点の対応関係が必要になる[13]。この行列 F を用いて、画像間の点の対応を見つけるためのステレオマッチング問題を解くことができる。本手法では、エピポーラライン上に探索範囲を限定した相關マッチングによってあたえられた画像間の点の対応を得ている。また、新しい視点からの画像を生成する際にもこの弱カメラキャリブレーションの結果を用いている。

5. では、この手法を使って算出した点の再投影後の座標値が、参照点の座標値測定時における誤差に対して、強カメラキャリブレーションに基づいた再投影法と比べ、より安定した性能を持っていることを示す。

3. アフィン座標の特性

ここではアフィン座標に基づいた再投影の考え方を以下の2つの場合に分けて検討する。

- (i) 平行投影の場合：シーン中の点は何処にあっても良く、カメラの投影法を正射影と仮定する。
- (ii) ピンホール射影の場合：シーン中の点は何処にあっても良く、カメラの投影法をピンホール射影とする。

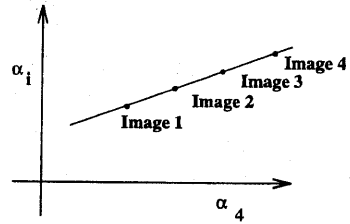


図5 α_4 と α_i の関係

3. 1 平行投影の場合

カメラ投影のモデルを平面への平行投影とそれに続くアフィン変換とで単純化する手法が既に提案されている[9]。点の組 $(R, B, B, \dots, B_i)$ があったとして、 R, B, B を通る仮想の面（基準面）を考える（図4）。 R から基準面に下ろした垂線の足を p_i' とする。 p_i' の (R, B, B) を基準としたアフィン座標は (a_i, b_i) である。同様に、 i 番目の点から基準面へ下ろした垂線の足を p_i' 、そしてアフィン座標は (a_i, b_i) とする。更に、 d_i と d_i をそれぞれ R, B の点から基準面までの距離とする。

アフィン座標 (α_i, β_i) とは、視点から R を見たときに基準面に投影される点のアフィン座標を示している。この (R, B, B) に対してアフィン座標 (α_i, β_i) を持つ点を p_{ik} とする。

p_{ik} と R を結ぶラインが視線方向となる。この線は画像面（この面の法線は視線と平行）と点 q_i で交わる。これで、 q_i は R の像となる。同様に (R, B, B) をそれぞれ画像面の (q_1, q_2, q_3) に投影する。この (q_1, q_2, q_3) を基準に選択した場合、画像面に対して（平行移動、回転、拡大縮小、などの）アフィン変換を施した場合にも q_i はアフィン座標 (α_i, β_i) を持つ。

次に、残った投影点のアフィン座標 (α_i, β_i) もこの与えられた視線方向から (α_i, β_i) と同様に計算する。 p_{ik} を基準面と R を通る視線方向からの平行光線との交点とするならば、 q_i はその画像面への写像となる。同様に、 p_{ik} と q_i は (R, B, B) 、 (q_1, q_2, q_3) をそれぞれ基準としたときに、アフィン座標 (α_i, β_i) を持つ。 $R p_{ik} p_i'$ と $R p_{ik} p_i'$ の三角形の相似を用いるならばその関係式は、

$$(p_{ik} - p_i) = \frac{d_i}{d_4} (p_{ik} - p_4) \quad (3)$$

の様に表され、これをアフィン座標で表すと、

$$(\alpha_i, \beta_i) - (a_i, b_i) = \frac{d_i}{d_4} ((\alpha_4, \beta_4) - (a_4, b_4)) \quad (4)$$

となり、 α だけを考慮すれば、

$$\alpha_i - a_i = \frac{d_i}{d_4} (\alpha_4 - a_4) \quad (5)$$

となる。この式中の a_i, a_4, d_i, d_4 は画像内の全ての点で同一の定数で、これらは与えられた3次元上の点から生成できる。すなわち、 $(R, \dots, B_i)$ から生成されるすべての可能な像において、 α_4 と α_i は直線にプロットされ、その傾きは図5の様に d_i/d_4 となる。この傾きは R が基準面からどれだけ離れているかを示す。同様に β_i, β_4 も

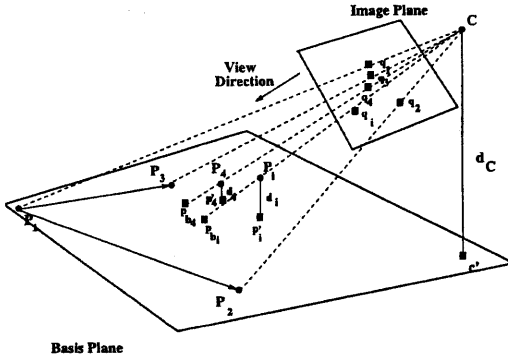


図6 ピンホール射影系のアフィン座標

α と同じ傾きで直線となる。

3. 2 ピンホールの場合

ここでは、3次元上の点を平面上に投影するために標準的なピンホールカメラの座標系を仮定する。その後これらの投影した点のアフィン変換を行う。 $(R, B, B, \dots, R)$ を3次元上の点とし、 R, B, B を通る仮想平面(基準面)を図6の様に定義する。 R から基準面に下ろした垂線の足を p'_i とし、その (R, B, B) を基準としたアフィン座標を (α_i, β_i) とする。同様に p'_i を作成する。 (α_i, β_i) は視点位置から基準面へ投影した点 p_{hi} に対して (R, B, B) を基準としたアフィン座標である。 p_{hi} と R を通る直線の延長上にカメラの光学中心 C がある。この位置は任意に定めたものである。視線についても任意に選んだ。直線 Rp_{hi} は q_i で画像平面と交わる。 q_i は R の像である。同様に、 R, B, B を画像平面にそれぞれ q_1, q_2, q_3 として写像する。 q_i は、この (q_1, q_2, q_3) を基準として、アフィン座標 (α_i, β_i) を持っている。

R を画像面に投影した点を q_i とする。前述の通り、 p_{hi} と q_i はそれぞれ (R, B, B) 、 (q_1, q_2, q_3) を基準とした際に共にアフィン座標 (α_i, β_i) を持つ。また、 c' と d_c をそれぞれ、カメラ中心 C から基準面に垂直に下ろした点、そこからカメラ中心までの距離とする。 $Cp_{hi}c'$ と $Rp_{hi}p'_i$ の2つの三角形の近似を用いて、

$$p_{hi} - p'_i = \frac{d_i}{d_c} (p_{hi} - c') \quad (6)$$

同様に $Cp_{hi}c'$ と $Rp_{hi}p'_i$ の近似を用いて、

$$p_{hi} - p'_i = \frac{d_i}{d_c} (p_{hi} - c') \quad (7)$$

(6)と(7)式から、

$$\frac{p_{hi} - p'_i}{p_{hi} - p'_i} = \left(\frac{d_i}{d_c} \right) \left(\frac{p_{hi} - c'}{p_{hi} - c'} \right) \quad (8)$$

アフィン座標で上記の式を書き直せば、

$$\frac{(\alpha_i, \beta_i) - (a_i, b_i)}{(\alpha_i, \beta_i) - (a_i, b_i)} = \left(\frac{d_i}{d_c} \right) \left(\frac{(\alpha_i, \beta_i) - (a_i, b_i)}{(\alpha_i, \beta_i) - (a_i, b_i)} \right) \quad (9)$$

α だけを取り出すと、

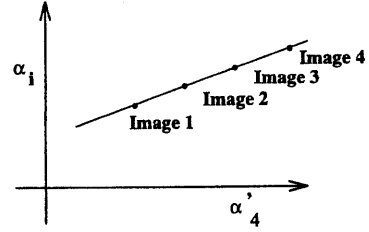


図7 α'_i と α_i の関係

$$(\alpha_i - a_i) = \frac{d_i}{d_c} \alpha'_i \quad (10)$$

ここで、

$$\alpha'_i = \left(\frac{\alpha_i - a_c}{\alpha_i - a_c} \right) (\alpha_i - a_i) \quad (11)$$

となる。式中、 a_i, a_c, d_i そして d_c は生成可能なすべての画像に対して同じ値であり、 a_c はカメラパラメータに依存する。つまり、 a_c と a_i が分かれば、与えられた画像に対して簡単に α'_i を計算することができる。すなわち、生成可能なすべての画像に対して、 α'_i, α_i は傾き d_i/d_c の直線上にプロットされる。この特性を、図7に示した。 β 空間に描かれる直線も α 空間に描かれる直線と同様の傾きである。

番目の点に対応した直線の傾きはその点から基準面までの距離に正比例している。

4. 新しい視野画像の合成と他のアプリケーション

4. 1 新しい視野画像の生成

4. 1. 1 平行投影の場合

2つの画像 I_1, I_2 が与えられたとする。新しい視野画像の生成のためには、2つの画像の間の点の対応関係を必要とするため、 I_1 上の点 p'_i に対応する I_2 上の点は p'_i と定義する。 α や β 空間で直線を生成するためには、4つの参照点が必要であるため(3. 2)、ここでは画像 $I_j (j=1, 2)$ 上で p'_1, p'_2, p'_3, p'_4 を定める。簡単にするために、 RR と RR および RR が直角で、かつ $|RR|=|RR|=|RR|$ となるように定めた。この構造をあらかじめ実験の前に2台のカメラで撮影し、その投影後の座標値 p'_1, p'_2, p'_3, p'_4 を記録しておく。画像 I_j 中の点 p'_i と p'_i のアフィン座標はそれぞれ (α'_i, β'_i) 、 (α'_i, β'_i) である。 α 空間中の直線は2つの画像に対応した点 (α'_i, α'_i) と (α'_i, α'_i) を通る。

ここで R, B, B, R のワールド座標を $(0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1)$ とそれぞれ仮定する。これらの点は仮想画像中の p'_1, p'_2, p'_3, p'_4 に 3×3 の仮想カメラの透視変換行列をつかって投影される。 (α_i, β_i) は p'_i の (p'_1, p'_2, p'_3) を基準としたアフィン座標である。次に、 I_1 中の p'_i を仮想画像に再投影するためには仮想画像中のそのアフィン座標を p'_i に対応する α 空間上の直線から計算すればよい。直線の方程式が $\alpha_i = \kappa_0 \alpha_4 + \kappa_1$ ならば、 $\alpha'_i = \kappa_0 \alpha'_4 + \kappa_1$

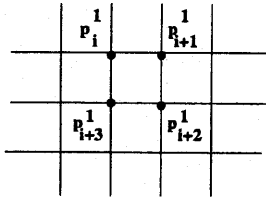


図8 ピクセルを頂点と考えたポリゴン

として求めることができる。同様に、 β_i' についても計算する。画像中の座標 p_i' は式(12)を使って求めることができる。

$$p_i = p_i + \alpha_i'(p_i - p_i) + \beta_i'(p_i - p_i) \quad (12)$$

ここで画像 I_i 中の複数の点が仮想画面上の1点にマッピングされてしまうことに注意しなければならない。この問題は、マッピングすべき点を選択する指標としてアフィン空間上の直線の傾きを利用することで解決できる。傾きの大きい直線上の点は、基準面からの距離が離れている、すなわち、仮想カメラが基準面をのぞき込んでいる形になっているとすれば、仮想画面との距離がより近いことになる。

また、このようにして選択、再投影された画像には画素抜けが生じてしまう恐れがある。なぜならこの再投影処理はすべての画素に画像 I_i 中の点がマッピングされることを保証してはいないからである。ここではこれらの問題をコンピュータグラフィックスの技法であるZバッファ、テクスチャマッピングを使って解決した。

まず図8に示すように、画像 I_i を一辺が1ピクセル長の四角形に分割する。図中の p_i' の x, y 座標は、それぞれ p_i' の x, y 座標に投影される。z座標の値はアフィン空間での直線の傾きの逆数とし、テクスチャは画像 I_i のものを使う。この処理を残った四角形の3点、およびグリッド上のすべての点にも繰り返す。次にこのポリゴンにテクスチャをマッピングする。この処理によって、新しく生成された画像を滑らかに補間することができる。

4. 1. 2 ピンホールの場合

2つの画像 I_1, I_2 が与えられたとする。新しい視野画像の生成のためには、2つの画像の間の点の対応関係を必要とするため、 I_1 上の点 p_i^1 に対応する I_2 上の点は p_i^2 と定義する。 α や β 空間で直線を生成するためには、4つの参照点が必要であるため(3. 2)、ここでは画像

上で $p_1^1, p_2^1, p_3^1, p_4^1$ を定める。簡単にするために、 RA_1 と RA_2 および RA_3 が直角で、かつ $|RA_1|=|RA_2|=|RA_3|$ となるように定めた。この構造をあらかじめ実験の前に2台のカメラで撮影し、その投影後の座標値 $p_1^1, p_2^1, p_3^1, p_4^1$ を記録しておく。画像 I_i 中の点 p_i' と p_i' のアフィン座標はそれぞれ (α_i', β_i') 、 (α_i', β_i') である。 α 空間中の直線は2つの画像に対応した点 (α_i', α_i')

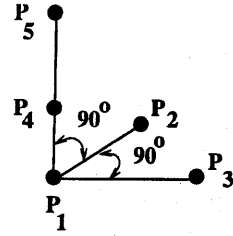


図9 5つの参照点の配置

と (α_i', α_i') を通る。

j 番目の画像において、式(11)を使って α_i' を計算するためには、 α_i' と α_i' を知っている必要がある。選んだ $R_1, \dots, R_5$ では α_i' の値は0である。 α_i' を計算するためには、5番目の制御点 R_5 が必要である。便宜上 R_5 を RA_5 の延長上に選び、 $|RA_5|=k|RA_1|$ と定義した(図9)。点 R_5 の j 番目のカメラへの写像は p_5^j であり、そのアフィン座標は (α_5^j, β_5^j) である。 $\alpha_5 = 0$ であるから式(10)(11)を用いて、

$$\alpha_i' = \frac{\alpha_i'(1-k)}{1-k \frac{\alpha_i'}{\alpha_i'}} \quad (13)$$

と書ける。仮想画像を生成するために、 R_1, R_2, R_3, R_4, R_5 のワールド座標をそれぞれ $(0,0,0), (1,0,0), (0,1,0), (0,0,1), (0,0,k)$ とそれぞれ仮定する。これらの点は仮想画像に 3×4 の仮想カメラの透視変換行列を使ってそれぞれ $(p_1^j, \dots, p_5^j)$ へ投影される。 i 番目の点 p_i^j の再投影は (p_i^j, p_i^j, p_i^j) を基準としたアフィン座標 (α_i', β_i') を計算することで実現できる。 α 空間中の直線の方程式が $\alpha_i = \kappa_0 \alpha_i' + \kappa_1$ とすると、(10)(11)式により、

$$\alpha_i' = \frac{\frac{\kappa_0 \alpha_i' \alpha_i'}{1 - \frac{\kappa_0 \alpha_i'}{\alpha_i'}} + \kappa_1}{1 - \frac{\kappa_0 \alpha_i'}{\alpha_i'}} \quad (14)$$

となる。ここで、 α_i' と α_i' はそれぞれ p_i^j のアフィン座標、仮想カメラの中心の写像である。同様に、 β_i' についても計算する。画像 I_i 中の全ての点は仮想画面上にこの手法を使って投影することができる。新しい画像中に発生する画素抜けや、z値の問題は、平行投影の場合で述べた方法で解決する。

4. 2 実物体と仮想物体の融合

VRのアプリケーションにおいては、実シーン中にCADモデルなどのCGオブジェクトを配置することが求められる。新しい視野画像の合成と同様に、現実のシーンや物体の3次元構造を計算し、その中に配置することが一つの解決法となるが、本報告では、3. で述べた手法を用いてこれを解く。

4. 1 で述べたのと同様の R を原点とした軸 $R_1, \dots, R_5$ の系を作り、その座標をそれぞれ $(0,0,0), (1,0,0), (0,1,0), (0,0,1)$ と定義する。図10に示すように、CADモデルのポリゴン表現が与えられており、その座標値が我々の軸のシステムに対応していると仮定するならば、こ

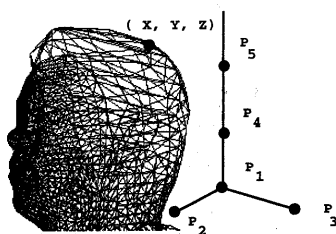


図10 CGオブジェクトと座標系

のCGオブジェクトをレンダリングし、4.1で述べた z 値を考慮した再投影法を用いて、融合したシーンを生成することが可能となる。

5. 本手法と他の直接的な手法との比較

1. で触れたように、再投影手法の一つであるエピソードライン交差法[6,7]は強カメラキャリブレーションもしくは3枚の画像(2枚は実画像, 1枚は新しい画像)中の8つの対応点が必要である。強カメラキャリブレーションパラメータを正確に推定することは困難であり、これは再投影時の誤差に通じる。また3枚の画像中の8つの参照点座標に含まれる測定誤差は、5点の参照点が必要なアフィン座標に基づいた方法と比較して、再投影時のより大きな座標誤差を導く。また、4.1で述べたように、シーンの再投影は“希薄な”画像を生成するだけである。アフィン座標を基にした方法では、アフィン空間中の直線の傾きをそれに対応する平面との距離の指標として使い、“穴埋め”が行えることを示した。エピソードライン交差アルゴリズムでは、再投影した点に対応する深さ(画像面に対する z 値)は明示的には計算されない、ゆえに新しい視野の生成へ拡張することは困難である。

文献[8]でのモーフィング手法は、厄介なステレオマッチング問題を避けることができる反面、2つの実カメラのカメラ中心を結んだ直線上に仮想カメラがある場合にしか利用できない。これに対し我々の手法では、まず点の対応問題を解く必要があるが、仮想カメラは空間中の何処にあっても良いという利点がある。

5.1 誤差に対する性能の実験的な比較

まず、アフィン座標に基づく再投影アルゴリズムが3次元再構築手法と比較してより安定していることを実験的に検証する。このため、図11に示すようなチェック模様の箱を含んだシーンを考え、カメラのキャリブレーションのために箱から6点を選んだ。手動で点の対応を取り、その点の3次元座標を計算し、新しい画像へと投影した。そのカメラマトリクスは任意に定めた。定めた6つの参照点それぞれの x 座標値の測定誤差が-1から1の間に均一に分布していると仮定し、再投影後の点の座標値の分布をシミュレートする。



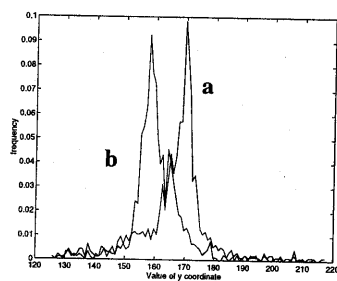
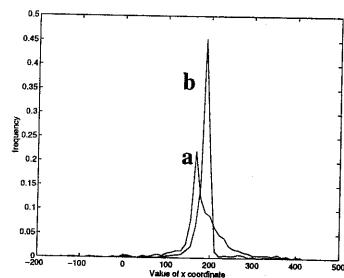
図11 実験に用いたステレオ画像

5点の参照点をアフィン座標を基にした再投影アルゴリズムのために選択し、測定誤差には同一の性質を仮定し、再投影点の座標値に与える影響を調べた。3つの異なった点を任意の画像の組から選び、3次元再構築手法と、アフィン座標を基にした手法の比較結果を考察する。この3点(i)(ii)(iii)に対して、 x, y 座標のヒストグラムを、3次元再構築手法(a)、アフィン再投影(b)それぞれについてプロットしたものを図12に示す。3点共に、アフィン座標を元にした再投影法の x 座標の値は明確に鋭いピークを持っており、かつ分散も3次元再構築手法に比べて非常に小さいことが分かる。 y 座標については、(i)(ii)において3次元再構築手法の方により鋭いピークが表れている。しかし、対応するヒストグラムはアフィン座標を基にした手法に比べて、分散が大きいたことが分かる。投影点の正確な x, y 座標という一つの根拠が欠如しているため、この手法の正確さをピークの高さに対応するヒストグラムの広がり度で判定するとすれば、我々の直接的再投影手法では3次元再構築による手法に比べてより安定した結果が得られると言える。

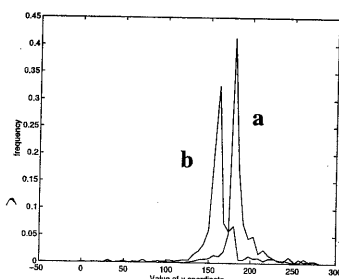
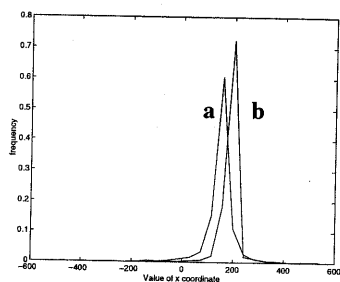
5.2 アプリケーション

アフィン座標を用いたシーン再投影の手法を検証するために、SGIのワークステーションに接続されたINDYCAMで、あるシーンの3枚の画像を撮影した。その画像を図13に示す。(画像(c)と(d)は同じもの)前節では、3枚目の画像は常に仮想画像だったが、3番目の画像が実画像の場合にもこの手法は利用できる。画像上のチェック状のブロックから、3枚の画像中の対応点5点を選び出し、 x' でマークした。また、7つの対応点を最初の二つの画像から選び出し、図13(a)(b)に黒いドットで示した。図13(c)(d)はそれぞれピンホールモデル、平行投影モデルによりこれらの点が3枚目の画像にどのように投影されるかを示した図である。ピンホールモデルの方が平行投影モデルに比べて良い再投影結果を示していることが分かる。

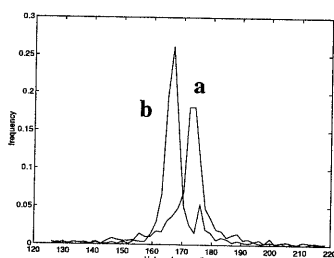
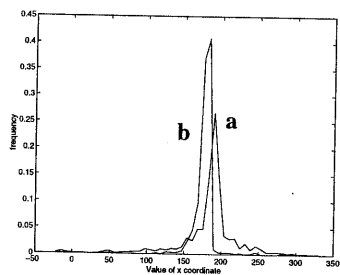
次に、図14(a)に示すような2枚の画像を基に新しい視野画像の合成を行う。新しい視野画像の生成結果を図14(b)(c)に示した。図14(b)では、画像は左側の実カメラより更に左にある仮想カメラから生成したものであり、図14(c)は極端に左上から見た仮想カメラにより生成したものである。特にこのような例では前に挙げた例に比べて、点の対応誤差が明白に再投影結果に反映されてしまう。これは主に入力画像の質の悪



(i)



(ii)



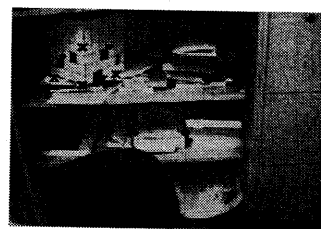
(iii)



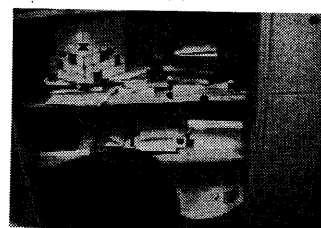
(a)



(b)



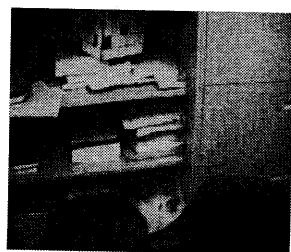
(c)



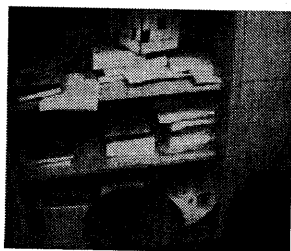
(d)

図 1 2 再投影結果に含まれる誤差の分布

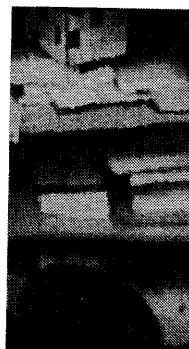
図 1 3 投影法の比較



(a)



(b)



(c)



図 1 4 新しい視野画像の生成例

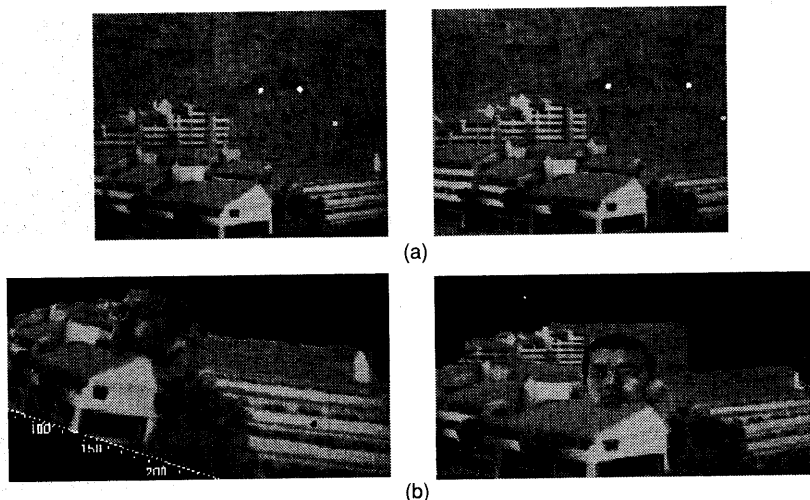


図15 実物体と仮想物体の融合例

さによるステレオマッチングの誤りに起因している。この新しい視野画像の生成アルゴリズムはSGI ONYXワークステーションにおいてほぼリアルタイムに実行することができる。

図15(b)は、人間の頭部のCADモデルを画像(a)に融合した結果を示している。前に触れたようにCADモデルの頂点の座標値は参照点の軸の設定と同様に正規化した。

6. まとめ

本報告では、新しい視野画像の生成、仮想物体と実空間画像の融合等のアプリケーションのための、ステレオ画像におけるアフィン座標の特徴を使った統一的な考え方を示した。シーン再投影のための計算は(14)式に示されているとおり、非常に単純であることが特徴である。このため、新しい視野画像の合成は、グラフィックスワークステーションを使うことによってほぼリアルタイムに実現できる。

また本手法は、文献[13]に述べられているような標準的なアプローチに比較して、5点という少ない対応点数で実現可能であり、ステレオマッチング時の誤差の再投影画像への影響を最小限に抑えることができる。

本手法を動的なシーンへ応用する場合の最大の問題点は、フレームレートでのステレオマッチングである。標準的なステレオマッチングのアルゴリズムは専用に作られたハードウェアを使用することで実現可能であるが、より効率的な手法との組み合わせが今後の課題である。

参考文献

- [1] J. Ohya, Y. Kitamura, H. Takemura, H. Ishii, F. Kishino and N. Terashima: "Virtual Space Teleconferencing: Real-Time Reproduction of 3D Human Images", Journal of Visual Communication and Image Representation, Vol.6, No.1, pp.1-25, 1995
- [2] O.Faugeras: "What can be seen in three dimensions with an uncalibrated stereo rig?", Proceedings of ECCV, pp.563-578, 1992
- [3] T.Kanade, P.J.Narayanan, and P.Rauder: "Virtualized Reality: Concepts and Early Results", Proceedings IEEE Workshop on the Representation of Visual Scene, pp.69-76, 1995
- [4] T.Kanade, A.Yoshida, K.Oda, H.Kano and M.Tanaka: "A stereo machine for video rate dense depth mapping and its new applications", Proceedings of CVPR, pp.196-202, 1996
- [5] K.Kutulakos and J.Vallino: "Affine Object Representation for Calibration-Free Augmented Reality", Proceedings of VRAIS, pp.25-36, 1996
- [6] O.Faugeras and L.Robert: "What Can Two Images Tell Us About a Third One?", IJCV, vol.18, pp.6-19, 1996
- [7] J.Mundy and A.Zisserman: "Appendix-Projective geometry for Machine Vision", Geometric Invariances In Computer Vision, MIT Press, 1992
- [8] S.Seitz and C.Dyer: "Toward Image-Based Scene Representation Using View Morphing", Proceedings of ICPR, pp.84-89, 1996
- [9] D.W.Jacobs: "Space Efficient 3D Model Indexing", Proceedings of CVPR, pp.439-444, 1992
- [10] O.Faugeras: "Three-Dimensional Computer Vision: A Geometric Viewpoint", MIT Press, 1993
- [11] 江尻, 大田, 池内: "マシンビジョン", 昭晃堂
- [12] D.Ballard and C.Brown: "Computer Vision", Prentice Hall, 1982
- [13] Q.Luong, R.Deriche, O.Faugeras and T.Papadopoulos: "On Determining the Fundamental Matrix: Analysis of Different Methods and Experimental Results", Tech Report No.1894, INRIA Sophia Antipolis, France, 1993