

透視変換下における仮想視点からの画像合成の一手法[†]

尾崎 信之*・館 暲*

Image Synthesis from Virtual Viewpoints under Perspective Transformation

Nobuyuki OZAKI* and Susumu TACHI*

This paper proposes a new algorithm to create virtual world by utilizing 2D scenery images in virtual reality. The algorithm based on Ullman's algorithm deals with perspective transformations, which is essential. We can construct only with several real images stored in advance without making 3D model by computer graphics.

Key Words: Virtual Reality, Tele-existence, Perspective Transformation, Virtual Viewpoint, Image Synthesis

1. はじめに

人工現実感において、仮想世界を構築するには、実世界をモデリングしてから、3次元CADなどのコンピュータグラフィックスを用いるのが一般的である。一方、仮想世界へのテレイグジスタンス^{1),2)}は対象が実世界であっても、構築の方法は同じである。そのため、現実感あふれるものとはいいがたい。そこで、仮想世界そのものをリアリティの高いものとする1つの方法として、実画像を用いて構築することを考える。このとき、仮想世界とのインタラクション、特に視点の移動によって提示する映像を変更していくが必要になる。

一方、Ullmanらは、複数枚の既知の2次元投影像を用いて、新規の位置からの投影像を線形結合により算出するアルゴリズムを考案し、物体の認識を行っている³⁾。このアルゴリズムは、カメラ撮影のモデル式である透視変換を平行投影で近似し、議論を進めている。さらに、近似の精度がよい平行透視変換に拡張してUllmanらのアルゴリズムを適用した研究もある⁴⁾。しかし、これらは、対象物体の大きさに比して、カメラと対象物体までの距離が大きいことを前提としているため、相対距離が近い場合には、近似の精度が悪くなり、合成される画像はかなり変形してしまうことになる。

そこで、Ullmanらのアルゴリズムを基本に、実用上避けて通れない透視変換を考慮して、画像を合成する

アルゴリズムを考案し、その正当性を計算機シミュレーションで確認したのでここに報告する。

2. Ullman らのアルゴリズム

カメラと仮想の視点⁵⁾が Fig. 1 のような位置関係にあると考える^(注1)。

得られる像の座標はつぎのように計算される。

$$P_i = (x_i, y_i) = (P_0 \cdot m_1^i, P_0 \cdot m_2^i) \quad (1)$$

ただし、 P_0 は対象の3次元の座標、 P_i はカメラ $i(i = 1, 2, 3)$ または、仮想視点 ($i = N$) で得られる P_0 の2次元画像の xy 座標、 m_j^i はカメラまたは仮想視点 i までの回転要素のみを考慮した 3×3 の座標変換行列 $M^i = (m_1^i, m_2^i, m_3^i)$ の j 列目の列ベクトルを表わす。ここで、 $(m_1^j, m_2^j, m_3^j), (j = 1, 2)$ が線形独立ならば、下式を満たす定数 $a_{ij}(j = 1, 2)$ が存在する。

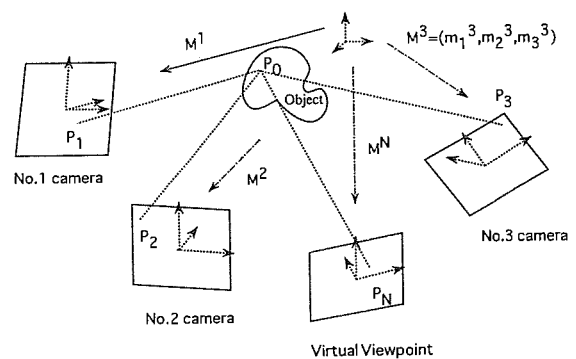


Fig. 1 Concept of image synthesis

(注1) 原論文では、物体に回転などの変換を与えた後、 xy 平面に投影する形で、議論を進めているが、本論文では、等価的にカメラを動かし、回転行列の変換に対応した姿勢方向からの撮像と考える。

[†] 第33回計測自動制御学会学術講演会で発表 (1994.7)
 * 東京大学先端科学技術研究センター 東京都目黒区駒場 4-6-1
 * Research Center of Advanced Science and Technology, The University of Tokyo, Meguro-ku, Tokyo
 (Received May 15, 1995)

$$m_j^i = a_{1j} \cdot m_1^j + a_{2j} \cdot m_2^j + a_{3j} \cdot m_3^j \quad (2)$$

求めたい像 P_N^i は, (1) 式と (2) 式と P_1, P_2, P_3 の定義を代入しつぎのようになる.

$$P_N^i = P_0 \cdot m_1^i + a_{11}(P_0 \cdot m_1^i) + a_{21}(P_0 \cdot m_2^i) + a_{31}(P_0 \cdot m_3^i) \quad (3)$$

y_N も同様.

$$y_N = a_{12} \cdot y_1 + a_{22} \cdot y_2 + a_{32} \cdot y_3 \quad (4)$$

透視変換への拡張における問題点

カメラの透視変換を考慮すると, (1) 式に相当する像 P_1^i までの変換は同次座標を用いてつぎのようになる.

$$P_1^i = (x_i, y_i, z_i, w_i) = P_0 \cdot M_1 \cdot S_1 \quad (5)$$

$$P_1^i = \begin{pmatrix} x_i \\ y_i \\ z_i \\ w_i \end{pmatrix} \quad (6)$$

ただし, M_1 はカメラの撮像中心までの並行移動を含めた座標変換行列で, S_1 は, 焦点距離 λ_1 による透視変換行列を表わす. (3), (4) 式を同次座標に拡張する.

$$x_N = \frac{(a_{11} \cdot x_1 + a_{21} \cdot y_1)w_1 + a_{31} \cdot x_2 \cdot w_2 + a_{41} \cdot x_3 \cdot w_3}{(a_{12} \cdot x_1 + a_{22} \cdot y_1)w_1 + a_{32} \cdot x_2 \cdot w_2 + a_{42} \cdot x_3 \cdot w_3} \quad (7)$$

$$y_N = \frac{(a_{14} \cdot x_1 + a_{24} \cdot y_1)w_1 + a_{34} \cdot x_2 \cdot w_2 + a_{44} \cdot x_3 \cdot w_3}{(a_{14} \cdot x_1 + a_{24} \cdot y_1)w_1 + a_{34} \cdot x_2 \cdot w_2 + a_{44} \cdot x_3 \cdot w_3} \quad (8)$$

(7), (8) 式は, w_i なる項が存在し, 対応する対象物の点 P_0 がわからないと算出できない.

3. 提案するフルコリスム

像の合成のフルコリスムは, Fig. 2 に示すように, 3つの手順により構成される.

3.1 ステップ-1: P_V^i を求める

前章で指摘した問題は, 透視変換 S_1^i の存在により, $w_i \neq 1$ となることにより生じる. そこで, (5) 式を $P_1^i \cdot S_1^{-1} \cdot P_0 \cdot M_1$ の2つに分割し, P_V^i を以下のよ

$$P_V^i = P_1^i \cdot S_1^{-1} \quad (9)$$

$$= P_0 \cdot M_1 \quad (10)$$

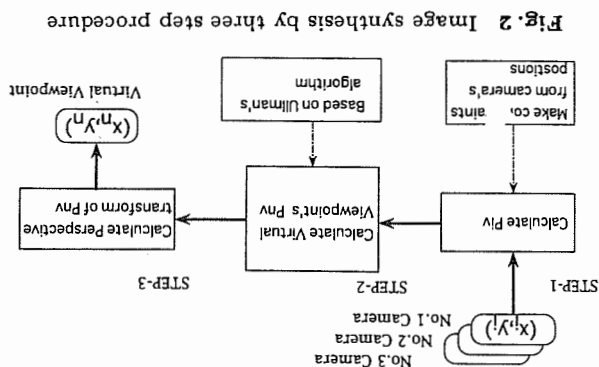


Fig. 2 Image synthesis by three step procedure

(10) 式において, M_1 は 4 番目の列ベクトルが

$$(0, 0, 0, 1)^T \text{ であるので, } P_V^i = (a_i, b_i, c_i, 1) \text{ と表現}$$

できる. つぎに, (9) 式を, $P_V^i \cdot S_1^i = P_1^i$ と変形し, カメラの撮像面上の既知の座標 (x_i, y_i) と λ_i を用いる

と, $P_V^i = \begin{pmatrix} x_i(1 - \frac{c_i}{\lambda_i}) \\ y_i(1 - \frac{c_i}{\lambda_i}) \\ c_i \\ 1 \end{pmatrix}$ とする. こ

ここで, 未知数は c_i である. 一方, カメラの位置関係がわかっているので, カメラ間の座標変換 U^i (カメラ i の座標原点を世界原点と見立てたカメラ j への座標変換) より, つぎのような拘束

$$P_V^j = P_V^i \cdot U^i \quad (11)$$

ここで, カメラ間の座標変換は, $U^i = (M_1^i)^{-1} \cdot (M_j^i)$ と表現できる. (11) 式はカメラが 3 台存在するときは, 未知数 3 に対して, $P_V^2 = P_V^1 \cdot U^1, P_V^3 = P_V^1 \cdot U^1$ なる拘束式が得られる. これは, 未知数を $C = (c_1, c_2, c_3)$ とおくと, $A \cdot C = B$ と表現され, 一般化逆行列を用いて P_V^i の座標が求まる.

3.2 ステップ-2: Ullman のフルコリスムを適用し, 仮想視点からの仮想像を求める

前節の計算により, つぎの 2 つの前提条件の下で Ullman のフルコリスムを適用することになる. 第 1 の条件は, Ullman のフルコリスムを適用する対象の変換 M_i はカメラの座標変換行列であり, 第 2 の条件は, Ullman の平行投影後に相当する点を, P_V^1, P_V^2, P_V^3 とすることである.

Ullman のフルコリスムの適用

ここで, カメラ 1, 2, 3 の変換行列から適当な列ベクトルを抜き出し, $N = \begin{pmatrix} m_1^1, m_2^1, m_3^1 \\ m_1^2, m_2^2, m_3^2 \\ m_1^3, m_2^3, m_3^3 \end{pmatrix}$ なる行列を生成する. この行列 N が一次独立 ($\det \neq 0$) ならば, M_N は $A = [a_{ij}]$ なる定数を用いて, 次のように表現できる.

$$M_N = N \cdot A \quad (12)$$

$$A = N^{-1} \cdot M_N \quad (13)$$

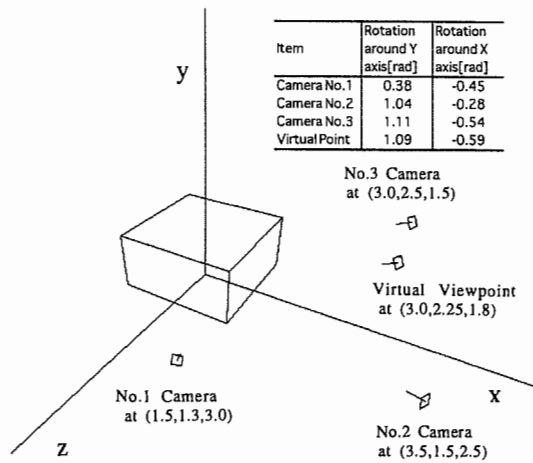


Fig. 3 Whole scenery including cameras' positions

上記の係数 \mathbf{A} を用いて、各要素に付いて、(3), (4) 式と同様に計算すると、 $\mathbf{P}_N^v = (x_N^v, y_N^v, z_N^v)$ が得られる。

$$x_N^v = a_{11} \cdot x_1^v + a_{21} \cdot y_1^v + a_{31} \cdot x_2^v + a_{41} \cdot x_3^v \quad (14)$$

$$y_N^v = a_{12} \cdot x_1^v + a_{22} \cdot y_1^v + a_{32} \cdot x_2^v + a_{42} \cdot x_3^v \quad (15)$$

$$z_N^v = a_{13} \cdot x_1^v + a_{23} \cdot y_1^v + a_{33} \cdot x_2^v + a_{43} \cdot x_3^v \quad (16)$$

3.2.1 ステップ-3: 透視変換を用いて、仮想視点上の点を求める。

(x_N^v, y_N^v, z_N^v) に透視変換を作用させ、求めたい仮想視点からの撮像面上の点の座標 (x_N, y_N) が求まる。

$$x_N = \hat{x}_N / \hat{w}_N, \quad y_N = \hat{y}_N / \hat{w}_N \quad (17)$$

$$(\hat{x}_N, \hat{y}_N, \hat{z}_N, \hat{w}_N) = (x_N^v, y_N^v, z_N^v, 1) \cdot \mathbf{S}_N \quad (18)$$

4. 計算機シミュレーション

高さが1で、幅・奥行きが2の直方体を対象として、カメラの位置関係を Fig. 3 のような条件にしてシミュレーションを実施した。ただし、焦点距離はすべて同じとした。既知の画像である Fig. 4 より仮想視点から見える像を合成すると Fig. 5 (a) となる。Fig. 5 (b) は、仮想視点にカメラを置いたときの画像である。また、Fig. 5 (c) は Ullman らのアルゴリズムを元に合成した像で、Fig. 5 (d) は平行透視変換による合成結果で、いずれの場合も、実際より変形していることがわかる。

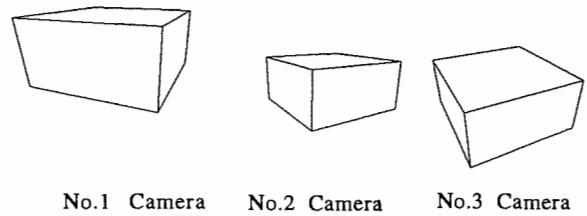


Fig. 4 Images obtained from cameras

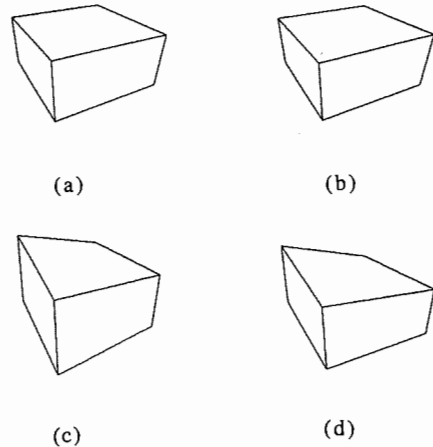


Fig. 5 Simulation results. (a) Synthesis based on the proposed method, (b) Image which should be obtained, (c) Synthesis based on the Ullman's algorithm, (d) Synthesis based on the Ullman's algorithm under paraperspective projection.

5. おわりに

本稿では、実画像を用いて仮想世界を構築するときの像を合成するための変換方法について述べた。端点などの特徴点を本方式により変換し、特徴点により構成される閉領域を対応する参照画像からテキスチャマッピングすることにより、仮想的な視点からの映像を合成することができる。したがって、人工現実感 (VR) において、リアリティの高い仮想世界を構築することができる。

参考文献

- 1) 館, 前田: 人工現実感を有するトレイグジスタンスロボットシミュレータ, 電子情報通信学会論文誌 D-2, **J75-D-2-2**, 179/189 (1992)
- 2) 大山, 常本, 前田, 館: 仮想環境へのトレイグジスタンスのための一手法, 日本ロボット学会誌, **12-8**, 1175/1182 (1994)
- 3) Ullman and Basri: Recognition by Linear Combinations of Models, IEEE Trans., **PAMI-13-10**, 992/1006 (1991)
- 4) 杉本, 室田: 半透視射影下での3次元物体認識, 信学技報, PRU92-96, 55/62 (1992)