

多眼カメラを用いた任意視点人物像の実時間生成システム

國田 豊[†] 稲見 昌彦[†] 前田 太郎^{††} 館 暲[†]

Real-Time Rendering System of Human Figures from Arbitrary Viewpoints
Using Multiple Cameras

Yutaka KUNITA[†], Masahiko INAMI[†], Taro MAEDA^{††}, and Susumu TACHI[†]

あらまし 遠隔面談コミュニケーションのために、写実的な人物像を実時間でバーチャル空間に投影する試作システムを開発した。このような用途では特に実時間性が要求されるために、本試作機は次のような特徴をもつ。まず任意視点位置の画像生成法として、バーチャル空間上に配置した平面へ、多眼カメラで得た画像を切り出して投射するというシンプルな手法を用いる。この手法では、人物の幾何的な形状の取得が不要であるため、必要となる計算機パワーは極めて少ない。この際、バーチャル平面付近の被写体でない生成画像の誤差が大きくなるが、これが一定の範囲内に収まる被写体の存在範囲を「等価被写界深度」として定式化した。また実装では、多眼カメラで得られる映像信号のうち、画像生成に必要なもののみを走査線単位で選択してコンピュータに取り込んでいる。このように早い段階で情報を取捨選択することで、取込み用のデバイスが少なくても、画像の劣化なく高精細な画像を生成できる。このような特徴により、コンシューマ向けの低コスト PC で実時間の画像生成に成功し、本手法の有効性を示した。

キーワード 相互トレイグジスタンス、バーチャルリアリティ、臨場感通信、実時間、Image-based Rendering

1. ま え が き

近年、パソコンにカメラを接続すればインターネット経由で遠くにいる人と手軽に映像つきのチャットを楽しむことができる。また、小型カメラを搭載した携帯通信端末の製品が相次いで発表されるなど、個人が「テレビ電話」を利用できる状況は整いつつあるといえる。しかし、そのようなシステムは、お互いのカメラでとらえた画像をそのままやり取りしているだけで、「同じ空間を共有している」という臨場感を提示するまでには至らない。その主な理由として、相手のいる実環境がそのまま提示されると、自分のいる実環境との継ぎ目が目立ってしまい、窓を隔てているかのような感じを受けてしまうことや、提示される相手の画像はある位置にあるカメラから撮影された固定視点のものであり、自らの視点移動に応じて見える相手の姿が変化しないということなどが挙げられる。

臨場感通信 [1] ~ [3] では、これらの問題を解決した「場」ともいふべき 3 次元共有空間を提示することを目指した。すなわち、コンピュータグラフィックにより生成したバーチャルな環境の中に、遠隔地にいる人々の姿を合成し、「一堂に会している」という状況を提示する。ただし、その際の人物像の表現としては、あらかじめ個人ごとの幾何的な 3 次元モデルを用意してはならない。また、視線の方向、表情、身振りなどを実時間で検出するために、膨大な計算機パワーを必要とするが、生成される人物像はそれに見合うだけの自然さをもつに至っていないという印象を受ける。その一方で、時間的に変化する人物の 3 次元形状を多眼カメラのビデオシーケンスから取得し、任意の視点からの映像を生成する研究が近年盛んであるが [4], [5], 特徴点の対応付けといった複雑な処理が必要であり、現在の計算機スピードでは実時間での処理はできていない。

我々の研究している相互トレイグジスタンス (Mutual Telexistence) [6] ~ [15] においても、離れた相手と臨場感をもって対面コミュニケーションを行うことを目的としてきた。臨場感通信と理念として共通

[†] 東京大学大学院工学系研究科, 東京都
Graduate School of Engineering, The University of Tokyo,
Tokyo, 113-8656 Japan

^{††} 東京大学大学院情報学環, 東京都
Interfaculty Initiative in Information Studies, Graduate
School of The University of Tokyo, Tokyo, 113-8656 Japan

する部分が多いが、人物の表現に3次元モデルを用いず、実際に多視点で撮影された画像をもとに、新たな視点位置からの「見え方」を生成する点が大きく異なる。このような画像生成法に関する技術は近年数多くなされ [16] ~ [21]、一般に Image-based Rendering (IBR) と呼ばれる。IBR はモデルを生成することなしに実物体の写実的な任意視点画像を生成できるという特徴をもつが、空間的に密な多視点データを扱うため、比較的データ量が多くなりがちである。そのため、データを収集するのにオフラインで時間をかけるか、個々のカメラの解像度を犠牲にして取り込むデータ量を減少させていた。一方、我々のシステムでは画像の取込みの段階でカメラの走査線単位で必要となるデータのみを選択して取り込むため、カメラの解像度を犠牲にすることなく実時間で任意視点位置の画像を生成することに成功している。

このように、IBR のアプローチは写実的な任意視点画像を生成するのに有効である。しかし、バーチャル空間との合成を考えた場合、幾何的な構造をもつバーチャル空間の中に、どのように「見え方」でしかない IBR の生成画像を投影するかという問題が生ずる。物体の存在領域の幾何的な情報を付加することにより解決した研究例 [22] もあるが、領域情報をあらかじめ指定する必要があり、我々が必要とする実時間での処理には適さない。

本論文では、バーチャル空間中に透明なバーチャル投影面を配置し、そこへ人物像を投影するというシンプルな手法を紹介する。このバーチャル投影面はレンダリング時にはテクスチャマッピングの張付け面としての役割を果たす一方で、その後でのバーチャル物体との遮へい関係を表現することができる。そしてこのような手法を用いた際、生成される画像に生じる誤差について明らかにし、この誤差がある基準以内に収まるような被写体の存在範囲を「等価被写界深度」として定式化する。更に、試作システムの設計と特徴について述べたあと、これによって写実的な人物像を実時間でバーチャルな空間に投影できることを示す。

2. 基本原理

2.1 近似画像の生成法

本節では、複数カメラでとらえた被写体の画像をもとに、新たな任意視点位置における近似的な画像を生成する方法について述べる。その際の前提として、被写体の形状情報は未知、被写体のおよその位置を既知

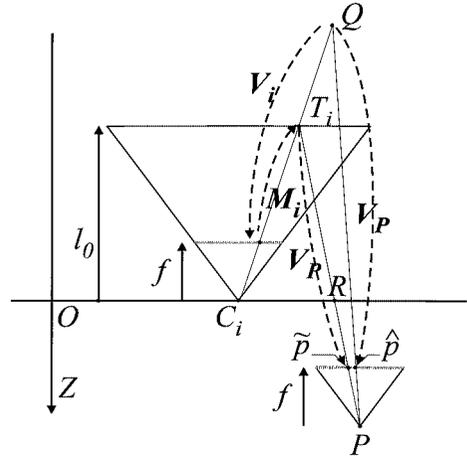


図1 バーチャル投影面を用いた画像
Fig. 1 Mapping using a virtual projection plane.

とする。そして、コンピュータ上の3次元空間中で、被写体があるおよその位置に透明なバーチャル平面を置き、生成する視点位置に応じて各カメラで得られた画像の一部をその平面に投影する、というのが大まかな画像生成の手順となる。以下、その具体的な手順を示す。

まず、図1のように3次元空間をデカルト座標系にとる。この際、水平方向を X 、鉛直方向を Y 、奥行方向を Z とする。ここで、カメラは平面 $Z = 0$ 上にあり Z 軸を負の方向に向き、 $Z < 0$ の領域にある被写体の像を得る。そして、近似画像を生成する視点位置を $P(X_P, Y_P, Z_P)$ とする。なお、次節で紹介する試作システムでは、カメラを水平方向に一列に並べているが、この節では一般化のためにカメラが2次元に配列してあるとする。

ここで、被写体の1点 $Q(X, Y, Z)$ の視点 P における本来の像 \hat{p} は、

$$s \begin{bmatrix} \hat{x}_P \\ \hat{y}_P \\ 1 \end{bmatrix} = V_P \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (1)$$

のように与えられる。ここで V_P は透視変換行列で、

$$V_P = \begin{bmatrix} -f & 0 & 0 & f \cdot X_P \\ 0 & -f & 0 & f \cdot Y_P \\ 0 & 0 & 1 & -Z_P \end{bmatrix} \quad (2)$$

となる。ところが (X, Y, Z) は未知であるため、 \hat{p} を

直接求めることはできない．そこで以下のような手順で近似的な像を得る．

まず，あるカメラ i に注目する． $C_i(X_i, Y_i, 0)$ をカメラ i の投影中心とすると，空間上の点 $Q(X, Y, Z)$ は，

$$s \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = V_i \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3)$$

$$V_i = \begin{bmatrix} -f & 0 & 0 & f \cdot X_i \\ 0 & -f & 0 & f \cdot Y_i \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (4)$$

によりカメラ i の像面上の点 (x_i, y_i) に写される．これがカメラで得られる像である．

そしていったんこの像を，投影中心を C_i として， $Z = -l_0$ なるバーチャル投影面に透視投影し， $T_i(X_{T_i}, Y_{T_i}, Z_{T_i})$ に写像する．すなわち，

$$s \begin{bmatrix} X_{T_i} \\ Y_{T_i} \\ Z_{T_i} \\ 1 \end{bmatrix} = M_i \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (5)$$

$$M_i = \begin{bmatrix} l_0/f & 0 & X_i \\ 0 & l_0/f & Y_i \\ 0 & 0 & -l_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6)$$

ここで，この l_0 を等価物体距離と呼び，被写体が存在するおよその距離に設定する．そして，未知の Q の代わりにこの T_i を視点位置 P で透視投影し， P での近似像 $\tilde{p}(\tilde{x}_P, \tilde{y}_P)$ を得る．すなわち，

$$s \begin{bmatrix} \tilde{x}_P \\ \tilde{y}_P \\ 1 \end{bmatrix} = V_P \begin{bmatrix} X_{T_i} \\ Y_{T_i} \\ Z_{T_i} \\ 1 \end{bmatrix} \quad (7)$$

結局，式 (3)，(5)，(7) により，被写体 $Q(X, Y, Z)$ は

$$s \begin{bmatrix} \tilde{x}_P \\ \tilde{y}_P \\ 1 \end{bmatrix} = V_P M_i V_i \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (8)$$

のように点 P での近似像 \tilde{p} に写される．ここで，被写体 Q の位置 (X, Y, Z) は未知であるが，既知であるカメラ i での像 (x_i, y_i) を用いて，

$$s \begin{bmatrix} \tilde{x}_P \\ \tilde{y}_P \\ 1 \end{bmatrix} = V_P M_i \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (9)$$

のように， \tilde{p} が得られる．

ここで被写体の像は，本来，式 (1) によって与えられる (\hat{x}_P, \hat{y}_P) であり，式 (8) により得られる $(\tilde{x}_P, \tilde{y}_P)$ とは異なる．そこで，両者の誤差ベクトルを Δp とすれば，

$$\begin{aligned} \Delta p &\equiv \begin{bmatrix} \tilde{x}_P \\ \tilde{y}_P \end{bmatrix} - \begin{bmatrix} \hat{x}_P \\ \hat{y}_P \end{bmatrix} \\ &= \frac{f(l_0 + Z)}{Z(l_0 + Z_P)(Z_P - Z)} \\ &\quad \cdot \begin{bmatrix} Z(X_P - X_i) - Z_P(X - X_i) \\ Z(Y_P - Y_i) - Z_P(Y - Y_i) \end{bmatrix} \end{aligned} \quad (10)$$

となる．

一方，投影面上の点 $T(X_T, Y_T, -l_0)$ と P を結ぶ直線とカメラのある平面 $Z = 0$ との切片を $R(X_R, Y_R, 0)$ とすれば，

$$s \begin{bmatrix} X_R \\ Y_R \\ 1 \end{bmatrix} = V_R M_i V_i \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (12)$$

$$V_R = \begin{bmatrix} -Z_P & 0 & X_P & 0 \\ 0 & -Z_P & Y_P & 0 \\ 0 & 0 & 1 & -Z_P \end{bmatrix} \quad (13)$$

更に平面 $Z = 0$ 上での R とカメラ i の投影中心の差分ベクトルを Δr とおくと，

$$\begin{aligned} \Delta r &\equiv \begin{bmatrix} X_R \\ Y_R \end{bmatrix} - \begin{bmatrix} X_i \\ Y_i \end{bmatrix} \\ &= \frac{l_0}{Z(l_0 + Z_P)} \\ &\quad \cdot \begin{bmatrix} Z(X_P - X_i) - Z_P(X - X_i) \\ Z(Y_P - Y_i) - Z_P(Y - Y_i) \end{bmatrix} \end{aligned} \quad (14)$$

ここで被写体のカメラ平面からの距離を $l \equiv -Z$ とおき，式 (11)，(15) により，

$$\Delta p = \frac{f(l_0 - l)}{l_0(Z_P + l)} \Delta r \quad (16)$$

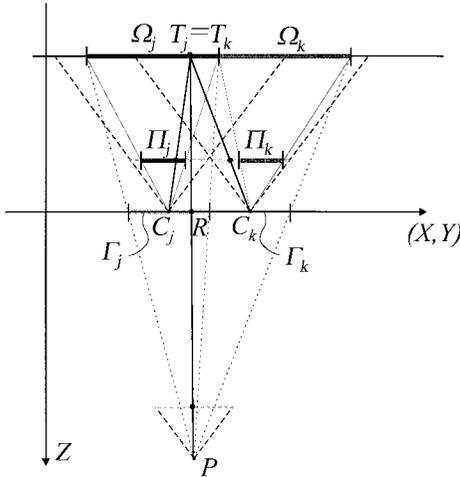


図2 各カメラの分担領域
Fig.2 Mapping regions of each camera.

が成り立つ。すなわち、像面上の誤差ベクトルは、カメラ平面上での直線 T_iP との切片とカメラの投影中心の差分ベクトルに比例する。

次に、カメラが複数ある場合を考える。カメラ j 、カメラ k ($j \neq k$) におけるバーチャル投影面への写像は

$$M_j : (x_j, y_j) \rightarrow T_j$$

$$M_k : (x_k, y_k) \rightarrow T_k$$

のように表される。このとき図2のように、お互いのカメラの画角が重なる領域において、 $T_j = T_k$ となる (x_j, y_j) 、 (x_k, y_k) が存在する。このような際、 R と投影中心が最も近いカメラ j を選び、バーチャル投影面への写像を行えば、式(9)より像面上での誤差は最小となる。すなわち、カメラの投影中心 C_j を母点としたボロノイ領域

$$\Gamma_j : \{(X_R, Y_R) | d(R, C_j) < d(R, C_k), j \neq k\} \quad (17)$$

に R がくるときに、カメラ j を用いる。ただし $d(P, Q)$ は2点 P, Q のユークリッド距離である。

ここで、

$$(x_i, y_i) \rightarrow (X_T, Y_T) \rightarrow (X_R, Y_R) \quad (18)$$

の写像は1対1なので、 (X_R, Y_R) 上の領域 Γ_i に対応する (x_i, y_i) 上の領域 Π_i 、 (X_T, Y_T) 上の領域 Ω_i を求めることができる。

まとめると、

(1) カメラ平面上で各カメラの分担するボロノイ領域 Γ_i を求める。

(2) 各 Γ_i に対応する像面上の領域 Π_i とバーチャル投影面上の領域 Ω_i を求め、投影行列 M_i により Π_i から Ω_i に画像を投影する。

(3) 投影行列 V_P により、バーチャル投影面の P における透視変換像を得る。

という手順により P における近似画像を生成できる。

2.2 等価被写界深度

ここで δ を視点 P で許容される最大誤差とする。そして、視点 P における生成像のすべての点に対して $|\Delta p| \leq \delta$ となる被写体の存在範囲を以下で示す。

式(16)のとおり、像面上での誤差の絶対値 $|\Delta p|$ は、 $|\Delta r|$ に比例する。ここで、カメラが等間隔 ε で並んでいるとすれば、 $|\Delta r| \leq \frac{\varepsilon}{2}$ となるカメラ C_i を必ず選ぶことができる。よって、

$$|\Delta p| \leq \frac{f|l_0 - l|}{l_0(l + Z_P)} \cdot \frac{\varepsilon}{2} \quad (19)$$

また、[式(19)の右辺 = δ] を l について解くと、以下に示す解 $l = l_1, l_2$ ($l_1 \leq l_0 \leq l_2$) を得る。

$$l_1 = l_0 \cdot \frac{1}{1 + \frac{l_0 + Z_P}{f \cdot \frac{\varepsilon}{2} - Z_P}} \quad (20)$$

$$l_2 = l_0 \cdot \frac{1}{1 - \frac{l_0 + Z_P}{f \cdot \frac{\varepsilon}{2} - Z_P}} \quad (21)$$

ここで、[式(19)の右辺] は $l = l_0$ のときに最小値0をとり、 $l < l_0$ で単調減少、 $l > l_0$ で単調増加である。よって、視点 P で生成された画像のすべての点は、その被写体が $l_1 \leq l \leq l_2$ の範囲に存在すれば、必ず $|\Delta p| \leq \delta$ なる誤差の範囲内で生成することができる。そしてこのとき δ を視点 P における解像度の1画素分とすれば、実質的に正しい位置に生成しているといえる。ここで、このように式(19)を満たす l の範囲の大きさを「等価被写界深度」と呼ぶことにする。

図3は ε に対して l_1 と l_2 をプロットしたグラフである。この図において、等価被写界深度は通常、 $\Delta l = l_2 - l_1$ により与えられる。ただし、 $\varepsilon < \frac{2\delta}{f} \cdot (l_0 + 2Z_P) \equiv \varepsilon_2$ においては $l > l_0$ なる l で必ず式(19)を満たす。更に、被写体は仮定のとおり $l > 0$ 、つまりカメラの前にあるが、 $\varepsilon < \frac{2\delta}{f} \cdot Z_P \equiv \varepsilon_1$ においては $l_1 < 0$ となるため、すべての $l > 0$ において式(19)を満たす。よって一般には、図中にグレーで示された領域を直線 $\varepsilon = \varepsilon_0$ で切った長さが $\varepsilon = \varepsilon_0$ における等価被写界深度となる。

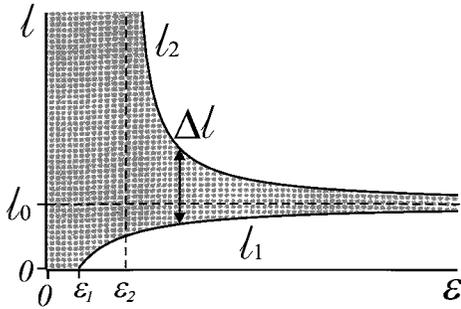


図 3 等価被写界深度のグラフ．等価被写界深度は图中， Δl で表されている
 Fig. 3 Graph of equivalent depth of field (EDOF). EDOF is expressed by Δl .

我々が通常のカメラで何か被写体を撮影するとき、ある被写体距離に焦点を合わせる。そして、この被写体距離にない被写体の像はぼけるが、このぼけが像面の解像度と比べて十分に小さいときには許容することができる。そのときの被写体の存在範囲を被写界深度というが、「等価被写界深度」は、この被写界深度と等価な考えであることから名づけた。

3. 試作システム

3.1 概要

前章で述べた画像生成法をもとに、人物のように動く被写体の任意視点画像を実時間で生成するシステムを試作した。図 4 と図 5 は、その概観とブロック図をそれぞれ示している。システム全体は大きく分けて、カメラ部、レンダリング用 PC、制御用 PC で構成されている。

カメラ部では、小型カラー CCD カメラ（東芝 IK-C40、画素の 1 辺 $\delta \approx 0.01\text{mm}$ ）に超広角レンズ（焦点距離 $f = 3.5\text{mm}$ ）を搭載し、これを水平に 12 台、54 mm 間隔で並べている。ここで、すべてのカメラに同期をかけている。また、図 4 の囲みのようにカメラは光軸を中心に 90° 回転させて台座に固定してある。こうすることで、カメラの走査線の向きは鉛直向きになる。これは、後程述べるように、各カメラで必要となる画像領域のみを効率良く選択するためである。

レンダリング用 PC のスペックは CPU: Intel Pentium II/400 MHz \times 2、メモリ: 128 MB、更に 3 次元グラフィックボード、映像取り込みボードを搭載している。そして、カメラ部からの NTSC 信号を取り込みつつ、それをもとに画像の描画を行う。ただし、こ

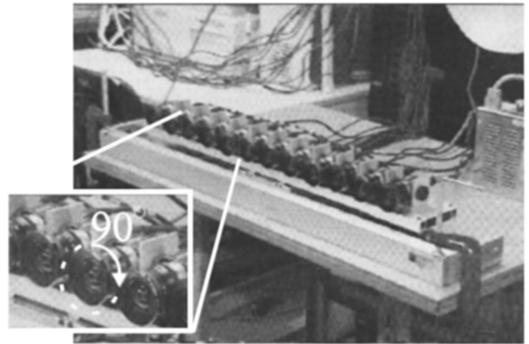


図 4 試作機の概観
 Fig. 4 Overview of our prototype.

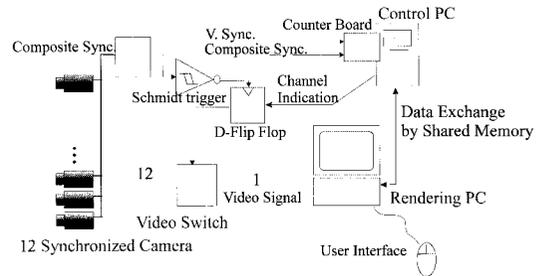


図 5 試作機のブロック図
 Fig. 5 Block diagram of our prototype system.

ここで注意したいのは、もしカメラ部からくる 12 本すべての映像信号を同時に取り込もうとすれば、映像取り込みボードも 12 枚必要となる点である。また、たとえ多数のボードを用いてもバスの転送レートがボトルネックとなり、すべてのデータをメインメモリに転送することは困難である。このため、カメラ部とレンダリング用 PC の間に映像スイッチを設置し、12 本の映像信号のうちで画像の生成に必要なものだけを選択するようにしている。この映像スイッチの切換タイミングは、外付けのラッチ回路を用いてカメラの水平同期信号に合わせてあり、カメラはすべて同期がとれているため、走査線単位での映像の選択が可能である。

制御用 PC は、カウンタボードを搭載している。このカウンタボードにはカメラからの複合同期信号を単相入力し、同期分離回路により抽出された垂直同期信号を Z 相に入力している。このようにして、現在の走査線の番号をカウントしつつ、適切なタイミングで映像スイッチへチャンネルの指示を行う。また、スイッチの切換パターンは、レンダリング PC でバーチャル視

点 P を移動するたびに变化するため、制御用 PC 上にある共有メモリを介してお互いに同期をとる。

3.2 実時間描画の手順

次に、我々の試作システムでの実時間描画の手順について述べる。基本的には、前章で述べた描画アルゴリズムに従っている。ただし、実時間処理を実現するために、必要な映像の選択をレンダリング PC のメモリに取り込む前の段階で行っていることが大きな特徴となっている。

まず、操作者によりバーチャルな視点の位置が与えられると、レンダリング PC は、各カメラでレンダリングに必要な画像領域 Π_i (前章参照) を計算する。カメラは水平に並んでいるので、この画像領域は画像全体を水平に切り分けたもの、すなわち垂直に長い形になる。このシステムでは走査線が垂直になるようにカメラを設置しているため、 Π_i に走査線の組が対応する。そこで、各カメラについて必要となる走査線の組を計算し、これをテーブルの形で制御用 PC 上の共有メモリに書き込む。

制御用 PC では、カウンタボードで走査線の番号をカウントしつつ、前述のテーブルに基づき、走査線ごとに映像スイッチにチャンネルを指示する。こうすることで、レンダリング PC はカメラからの映像のうち描画に必要な画像領域 Π_i を選択的に、しかも空間的な解像度の劣化なく取り込むことができる。スイッチの代わりに画像合成機を用いた実装例 [21] もあるが、そうすると必要のない画像を取り込むばかりではなく、必要な画像の空間的解像度も犠牲になるという欠点をもつ。ただ、我々の手法を用いても、領域 Π_i に対応する走査線がカメラ間で重複してしまうときがあり、そのときにはフィールドごとに時間を交互にずらしてスイッチの切替を行った。

最後に、レンダリング PC は、取り込んだ画像を用いてバーチャル視点からの映像を生成する。バーチャル空間の構築には 3 次元画像ライブラリ (OpenGL) を使い、この空間内にあるバーチャル投影面上に取り込んだ画像をテクスチャマッピングする。この処理は、前章で述べたとおりプログラムとしては単純なメモリ転送であり、グラフィックアクセラレータがハードウェアで高速にスケーリングフィルタリングの処理を行う。

4. 実験

4.1 カメラの設置間隔と等価被写界深度の関係

まずは予備実験として、カメラ間隔と等価被写界深

度の関係を確認する。そのため、静止した被写体に対してカメラ間隔 ε を変化させて画像を生成してみる。ここで、カメラは実際のシステムと同様に、水平方向にのみ移動させて撮影する。このような場合、生成した画像の垂直方向の誤差は、水平方向の誤差に比べて大きくなり、カメラのある位置よりも大きく上下にバーチャル視点を移動できない。ただし、我々の用途である遠隔コミュニケーションでは、十分であると考えられる。なぜならば、通常の会話においては上や下からのぞき込むことはないし、人間の両眼は水平についているため、両眼立体視で必要なのは水平方向の精度だからである。

図 6 の左上部は、カメラと画像を生成するバーチャル視点 P 、そして二つの被写体の位置関係を示している。カメラは $Z = 0$ 上に並べられ、 P は $Z = 50$ に位置する。また、二つの被写体は $Z = -350$ と -70 [mm] 付近にそれぞれ置かれている。

カメラのパラメータ $f = 3.5$ [mm]、 $\delta = 0.01$ [mm]、更に等価物体距離を l_0 を前側の被写体付近の $l_0 = 350$ [mm] にした際、水平方向の等価被写界深度のグラフは図の右上部のようにになる。

図の左下部の画像は、カメラ間隔 $\varepsilon = 4$ [mm] で生成したものである。このときの等価被写界深度は

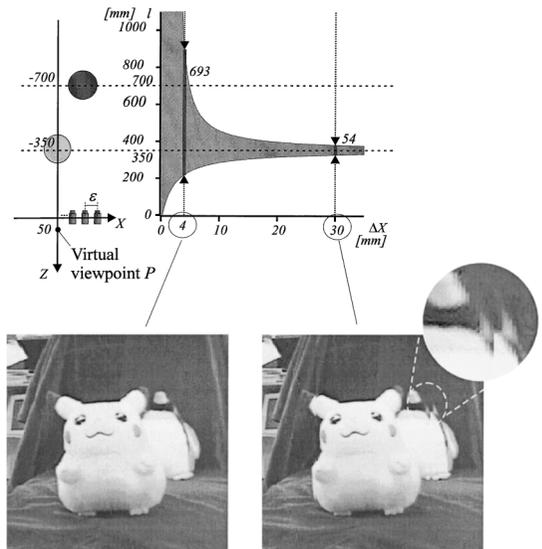


図 6 静止する被写体において、カメラの設置間隔 ε を変化させた場合の画像生成結果

Fig. 6 Image synthesis of static objects with different camera intervals ε .

693 [mm] となり、グラフのとおり、後ろの被写体のある $Z = -700$ まで含む．そして実際に、 l_0 付近にある前の被写体だけでなく、後ろの被写体の画像も誤差が実質的になく生成されている．

一方、 $\varepsilon = 30$ [mm] のときの等価被写界深度は 54 [mm] で、後ろの被写体には足りない．そして実際、図の右下部に示すように、生成された画像では後ろの被写体（と前の被写体の後ろに張り出した耳）が 2 重となるような箇所が生じている．生成される像は式 (16) のように、 $l < l_0$ のとき Δr と同方向、つまり各カメラの分担領域の境界に近づく方向にずれ、 $l > l_0$ のとき Δr と同方向、つまり各カメラの分担領域の境界から遠ざかる方向にずれるが、そのとおりになっている．

このように、理論的に導かれた等価被写界深度の考えが、実際によく合致することを確かめることができた．

4.2 バーチャル空間への人物像の投影

次に、今回の試作システムを用いて、動きのある人物を様々な視点位置から見た映像を実時間で生成する．

まず、本論文のシステムのパラメータとして像面上の許容誤差 $\delta = 0.01$ [mm]・カメラの焦点距離 $f = 3.5$ [mm]・カメラの設置間隔 $\varepsilon = 54$ [mm]、更に被写体の想定距離 $l_0 = 1000$ [mm] としたときの等価被写界深度を計算してみる．本実験では視点の位置を $-200 \leq Z_P \leq 1000$ の範囲で移動させるが、そのとき被写界深度は 167 [mm] $\leq \Delta l \leq 501$ [mm] の範囲

で変化する．これは、人物の厚みをカバーするには十分であるが、背景画像を正しく構成するには足りない．また、遠隔コミュニケーションでは、相手のいる実環境の背景画像が提示されると、自分のいる実環境との境界が目立ってしまい、臨場感を損なうことが知られている．このような理由から、人物像のみを実環境の背景から切り出し、コンピュータにより生成したバーチャルな空間に投影することにした．

図 7 の左部は、被写体となる人物とカメラ、撮影用のブルーバックがある実環境と、人物を投影するバーチャル投影面とバーチャルな物体を配置したバーチャル環境を重ねて表している．ただし図中、実環境を点線で、バーチャル環境を実線で示している．実環境では、人物はブルーバックを背にカメラからおよそ 1000 [mm] の位置に座る．バーチャル環境は、三方を壁で囲み、 $(X, Z) = (0, -300)$ 付近に茶瓶と台を設置している．そして、実環境で人物がいるおよその距離と同じく $l_0 = 1000$ [mm] の位置に透明なバーチャル投影面を設置する．

レンダリング時には、クロマ・キーを用いて人物が映っている画像のみをバーチャル投影面に投影し、背景部分は透明のままとしておく．ここで注意したいことは、固定のバーチャル投影面に人物像を投影しているが、この像は視点位置に応じて変化するということである．よって、左右視点位置の映像を生成して両眼に提示すれば、平面としてではなく、視差により立体として認識される．これはちょうど、液晶シャッタ式

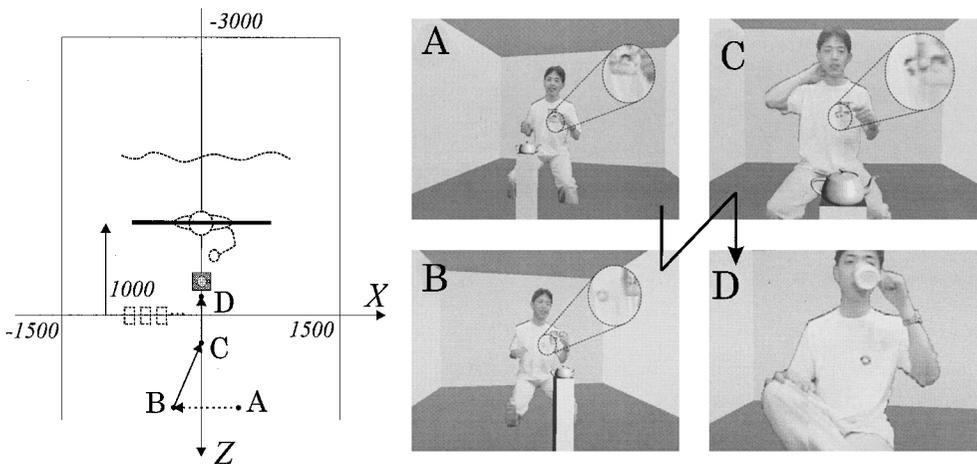


図 7 バーチャル空間への人物像の投影
Fig. 7 Projection of human figures into a virtual space.

の立体ディスプレイのディスプレイ面が固定平面であるにもかかわらず、立体を提示できるのと同じ理由である。

図 7 の左部は、視点位置 $P(X_P, Z_P)$ を $A(400, 1000) \rightarrow B(-300, 1000) \rightarrow C(0, 300) \rightarrow D(0, -200)$ と動かし、そこからの映像を実時間で生成した結果である。ここで、 $A \rightarrow B \rightarrow C \rightarrow D$ の時間間隔はそれぞれ、 $1/30, 3.5, 5.0$ [s] である。

視点位置 A, B, C における生成画像の図中の囲みは、人物がもつカップ付近の拡大図である。視点位置 A, B はほぼ同時刻であるが、 B からの画像では衣服の左胸に付いているマークが見えるのに対し、 A からはコップに遮られて見えない。そして、真ん中から見た C では、マークがちょうど半分ほどカップに遮られ、カップを上げた D の状態ではマークが完全に見えている。この結果から、動く被写体内で正しい遮へい関係が表現できていることがわかる。

また、バーチャル投影面の手前に設置した茶瓶は人物を隠し、背後に配置された壁は人物により隠されている。また、各画像における茶瓶と人物及び背景の位置関係も、視点位置に応じたものとなっている。すなわち、平面という単純な幾何的構造によってもバーチャル空間内での人物、茶瓶、壁の間の遮へい関係が適切に表現できているということがわかる。ただし、実際の人物の厚みよりもバーチャル投影面に近い位置にバーチャルな物体を配置したときには、正しい遮へい関係が表現できないという限界がある。しかし、我々の遠隔コミュニケーションシステムにおいては、利用者は一般に社会距離と呼ばれる 1 m 以上の距離は保って会話することを想定しており、お互いに重なり合うという状況は考えにくい。また、バーチャルな物体を操作する場合には、各操作者の実空間において「オブジェクト指向型ディスプレイ」[23] のように遮へい関係の表現に適したディスプレイを用いるのがよいと考える。

また実際には、等価被写界深度内の人物像に対して、若干の画像の継ぎ目が認められる。これは、各カメラのキャリブレーションが十分でなかったことと、レンズの球面収差の影響によるものであると考える。これらについては、他の多眼カメラを用いたシステムと共通する問題であり、我々のシステムも例外ではない。ただ、上記の不具合は各フレームを静止画として見たときに認められる程度であり、動画像として見たときの画質は満足いく結果となっている。ただし、

その心理物理的な評価を今回は行わなかったため、今後の課題としたい。

5. む す び

本論文では、我々の目指す遠隔コミュニケーションシステム「相互トレイグジスタンス」のために、人物像をバーチャル環境に投影することを試みた。その際、バーチャル空間内に透明な投影面を配置し、複数カメラからの画像を選択的にテクスチャマップすることで、視点位置に応じた人物像を生成するという手法を用いた。この手法は非常にシンプルである一方で、人物像に平面という最低限の幾何的構造をもたせることで、バーチャル空間内での遮へい関係も表現できる。ただし、幾何的な構造をもたないために、被写体が投影面付近に存在しない場合、その像では誤差が生じる。そこで、この誤差の定量的な評価基準として「等価被写界深度」を導入した。被写体がこの「等価被写界深度」の範囲内にあれば、生成画像の誤差が定めた範囲内にあることが保証され、システムの設計において有用な指針となる。

また、試作システムを用いて実時間で任意視点の人物像が生成できることを示した。この試作システムは、早い段階でカメラからの映像のうち必要となるもののみを選択するという、従来システムにない特徴をもつ。前述のシンプルな画像生成法との組合せにより、現時点のコンシューマ向け低コスト PC で、十分に実時間描画を実現している。

そして将来的には、本論文のシステムにより遠隔間で利用者の視点位置に応じた映像を生成して通信する実験を行うことを考えている。ただし、現在はカメラを直線状に配置しているため、背面の人物像を生成できないなど、視点位置に制約がある。そこで、カメラを円形に配置して視点移動が可能な範囲を広げ、その際に今回のバーチャル投影面の画像生成法を拡張することが考えられる。更に、本論文では主に画像の生成法について述べたが、相互トレイグジスタンスのシステム全体としては、その画像を視覚提示する必要がある。よって、このような用途に適したディスプレイに関する研究も併せて行う予定である。

文 献

- [1] 岸野文郎, “ヒューマンコミュニケーション—臨場感通信” テレビ誌, vol.46, no.6, pp.698–702, June 1992.
- [2] 竹村治雄, “仮想環境の共有による協調作業” 知能情報メディア, 田村秀行, 池田克夫 (編), pp.127–147, 総研出版, 東京, 1995.

[3] K. Singh, J. Ohya, and R. Parent, "Human figure synthesis and animation for virtual space teleconferencing," Proc. VRAIS '95, pp.118-126, 1995.

[4] T. Kanade, P.J. Narayanan, and P. Rander, "Virtualized reality: Being mobile in a visual scene," Proc. ICAT '95, pp.133-142, 1995.

[5] S. Moezzi, A. Katkere, D.Y. Kuramura, and R. Jain, "Immersive video," Proc. VRAIS '96, pp.17-24, 1996.

[6] 小柳雅明, 横山裕記, 前田太郎, 柳田康幸, 館 暉, "実動物体の仮想空間への投射システムに関する研究," 日本バーチャルリアリティ学会大会論文集, vol.1, pp.27-30, Oct. 1996.

[7] S. Tachi, T. Maeda, Y. Yanagida, M. Koyanagi, and Y. Yokoyama, "A method of mutual tele-existence in a virtual environment," Proc. ICAT '96, pp.9-18, 1996.

[8] 國田 豊, 稲見昌彦, 前田太郎, 館 暉, "相互トレイグジスタンスのための画像構成法," 日本バーチャルリアリティ学会大会論文集, 第2巻, pp.350-351, Sept. 1997.

[9] 國田 豊, 稲見昌彦, 前田太郎, 館 暉, "相互トレイグジスタンスのための視覚ディスプレイ設計," 日本バーチャルリアリティ学会大会論文集, 第2巻, pp.352-353, Sept. 1997.

[10] 國田 豊, 稲見昌彦, 前田太郎, 館 暉, "相互トレイグジスタンスのための画像構成法(第二報)," 信学技報, MVE98-43, June 1998.

[11] 國田 豊, 稲見昌彦, 前田太郎, 館 暉, "相互トレイグジスタンスのための画像構成法(第三報)," 日本バーチャルリアリティ学会大会論文集, 第3巻, pp.203-206, Aug. 1998.

[12] 國田 豊, 稲見昌彦, 前田太郎, 館 暉, "三次元物体の実時間描画システム," 3次元画像コンファレンス'99, pp.13-18, June 1999.

[13] Y. Kunita, M. Inami, T. Maeda, and S. Tachi, "Real-time rendering system of moving objects," Proc. 1999 IEEE Workshop on Multi-View Modeling & Analysis of Visual Scenes (MVIEW'99), pp.81-88, June 1999.

[14] Y. Kunita, M. Inami, T. Maeda, and S. Tachi, "Prototype system of mutual tele-existence," ACM SIGGRAPH '99 Conference Abstracts and Applications, p.267, Aug. 1999.

[15] 國田 豊, 稲見昌彦, 前田太郎, 館 暉, "三次元バーチャル空間への写実的な人物像の実時間投影システム," 第5回知能情報メディアシンポジウム論文集, pp.207-212, Dec. 1999.

[16] L. McMillan and G. Bishop, "Plenoptic modeling: An image-based rendering system," SIGGRAPH'95 Conference Proceedings, pp.39-46, Aug. 1995.

[17] S.J. Gortler, R. Grzeszczuk, R. Szeliski, and M.F. Cohen, "The lumigraph," SIGGRAPH'96 Conference Proceedings, pp.43-54, Aug. 1996.

[18] M. Levoy and P. Hanrahan, "Light field rendering," SIGGRAPH'96 Conference Proceedings, pp.31-41, Aug. 1996.

[19] 片山昭宏, 田中宏一郎, 押野隆弘, 田村秀行, "多視点画

像の補間・再構成による視点追従型立体画像表示法," 信学論(D-II), J79-D-II, no.5, pp.803-811, May 1996.

[20] 亀井克之, 丸山 稔, 瀬尾和男, "実環境再現を目的とした部分画像の検索転写による情景画像の生成," 信学論(D-II), J81-D-II, no.2, pp.331-341, Feb. 1998.

[21] 苗村 健, 原島 博, "多眼ビデオ入力を用いた実時間IBRシステム—Video-based rendering," 日本バーチャルリアリティ学会論文誌, vol.4, no.4, pp.639-646, Dec. 1999.

[22] 内山晋二, 片山昭宏, 山本裕之, 田村秀行, "幾何形状モデルを持たない多眼実写データの仮想環境における配置と操作," 日本バーチャルリアリティ学会論文誌, vol.3, no.3, pp.149-157, Sept. 1998.

[23] 川上直樹, 稲見昌彦, 柳田康幸, 前田太郎, 館 暉, "オブジェクト指向型ディスプレイの研究," 情処学論, vol.40, no.6, pp.2725-2733, June 1999

(平成12年3月21日受付, 7月31日再受付)



國田 豊 (学生員)

平8東大・工・計数卒。平10同大学院修士課程了。現在、同博士課程在学中。人工現実感などの研究に従事。平8日本工業新聞社主催先端学術学生論文文部大臣賞、平10日本バーチャルリアリティ学会学術奨励賞各受賞。計測自動制御学会、映像情報メディア学会、日本バーチャルリアリティ学会学生員。



稲見 昌彦

平6東工大・生命理工卒。平8同大学院修士課程了。平11東大先端学際工学専攻博士課程了。博士(工学)。同年より、東大国際・産学共同研究センターリサーチ・アシエイト。ロボット及びバーチャルリアリティの研究に従事。平10情報処理学会山下記念研究賞、平11同学会論文賞、日本VR学会学術奨励賞各受賞。IEEE Computer Society, 情報処理学会、日本ロボット学会、計測自動制御学会各会員。



前田 太郎 (正員)

昭62東大・工・計数卒。同年通産省工技院機械技研入所。平4東大先端科学技術研究センター助手。平6同大学院・工・計数工助手。平9同専攻講師。平12同大学院情報学環講師。工博。人間の知覚特性と神経回路網モデル, マン・マシンインタフェース, トレイグジスタンスの研究などに従事。平2計測自動制御学会論文賞, 平9同学会学術奨励賞, 平3日本ロボット学会技術賞各受賞。計測自動制御学会, IEEE, 神経回路学会, 日本ロボット学会, 日本バーチャルリアリティ学会各会員。



館 障 (正員)

昭 43 東大・工・計数卒．昭 48 同大大学院博士課程了．工博．同年同大助手．昭 50 通産省工技院機械技研研究員．主任研究官，遠隔制御課長，バイオロボティクス課長，マサチューセッツ工科大学 (MIT) 客員研究員を経て，平 1 東大・先端科学技術研究センター助教授．平 4 同センター教授．平 6 同大大学院・工・計数工教授．盲導犬ロボット，テレイグジスタンス，人工現実感などの研究に従事．IEEE/EMBS 学会賞，通産大臣賞，国際計測連合 (IMEKO) 特別勲功賞など各受賞．IMEKO ロボティクス会議議長，SICE フェロー，日本バーチャルリアリティ学会会長．