JAIST Repository
https://dspace.jaist.ac.jp/

Title	ViewFrame : 画像処理による位置検出法を用いた「借景」
Author(s)	河原塚,有希彦; 高橋,誠史; 宮田,一乗
Citation	情報処理学会研究報告 : ヒューマンインタフェース研究会報告, 2003(111): 45-51
Issue Date	2003-11
Type	Journal Article
Text version	publisher
URL	http://hdl.handle.net/10119/3328
Rights	社団法人 情報処理学会, 河原塚有希彦／高橋誠史／宮田一乗, 情報処理学会研究報告 : ヒューマンインタフェース研究会報告, 2003(111), 2003, 45-51. ここに掲載した著作物の利用に関する注意: 本著作物の著作権は（社）情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。 The copyright of this material is retained by the Information Processing Society of Japan (IPSJ). This material is published on this web site with the agreement of the author (s) and the IPSJ. Please be complied with Copyright Law of Japan and the Code of Ethics of the IPSJ if any users wish to reproduce, make derivative work, distribute or make available to the public any part or whole thereof. All Rights Reserved, Copyright (C) Information Processing Society of Japan.
Description	

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

社団法人 情報処理学会 研究報告
IPSJ SIG Technical Report

2003－ＨＩ－106 (7)
2003－MUS－ 52 (7)
2003／11／7

ViewFrame

－ 画像処理による位置検出法を用いた「借景」

河原塚 有希彦　　高橋 誠史　　宮田 一乗

本報告では，画像処理による位置検出技術を用いた新たな画像提示の環境，”ViewFrame”システムを提案する．本手法は，「窓越し現実感」というコンセプトの下に，窓の代わりに設置したディスプレイ上に，ユーザの視点移動に伴い変化する窓越しに見える景色をシミュレーションして提示するものである．本システムはリアルタイム性が強く求められるために，グラフィックボードのピクセルシェーダ機能を用いて，ハードウェアによる画像処理の高速処理を行った．本システムにより，自然なインタラクションで景観の奥行き知覚や３次元物体の形状認識が容易になる．また，本システムの実装例として「借景」を制作したのであわせて報告する．

ViewFrame

– An interactive image viewer by means of image-based position sensitive detector

Yukihiko Kawarazuka　Masafumi Takahashi　Kazunori Miyata

An interactive image viewer, “ViewFrame”, by means of image-based position sensitive detector is proposed. In this paper, the new concept, “window-though reality”, is introduced. A window in a room is replaced with a LCD or another flat screen, and this system displays a view-depended CGI scenery that would be seen from an observer. This system is required to react to user’s action in real time; therefore the image processing is implemented by means of pixel-shader functions for GPU to accelerate the process. This system makes it possible to recognize the depth of image and three-dimensional shapes easily by natural man-machine interaction. An application of this system, “Shakkei”, is also reported.

北陸先端科学技術大学院大学
知識科学研究科
School of Knowledge Science
Japan Advanced Institute of Science and Technology

1. はじめに

ヒューマンインターフェイスの研究において，物理的なオブジェクトを用いた計算機とのインタラクションや，現実世界における人間の状況を計算機の入力として利用する実世界志向インタフェースに関して，精力的な研究がなされている．また，科学技術計算の可視化や都市景観の検討，芸術品のディジタルアーカイブなど，さまざまな用途にVR環境が利用され始めている．VR環境を利用することで，立体物の形状把握や構造，景観の検討などを高臨場感で行うことが可能になる．しかし，これらの環境には，一般的に大型の映像投影装置が必要とされ，結果として大規模なシステムになってしまうという問題点がある．

本報告では，標準的なディスプレイを利用した，手軽で，かつ，自然な身体動作で映像コンテンツに擬似3次元的にアクセスするための手法を提案する．

2. 本研究の概観

本章では，本研究のコンセプトと関連研究例について述べる．

2.1 コンセプト：「窓越し現実感」

本研究では，手軽な環境で高いリアリティを示し，かつ，自然な身体動作を通して，デジタルコンテンツにアクセスするための概念として，「窓越し現実感」を提案する．

従来のVRでは，現実感を高めるためには「没入感」が重要であるとされ，CAVE[1]を代表とした，視野を覆いつくすための大型の大掛かりな映像投影の設備を要する．これに対し本手法では，ディスプレイ装置を「窓」と仮定し，ユーザが窓越しに見るであろうと予測されるシーンを，ユーザの視点の位置に連動しながらインタラクティブに描画することで，現実感を高めるアプローチをとる．

図1に「窓越し現実感」の概念図を示す．これはユーザが窓越しにポットを見ている図である．A地点からは，ポットの注ぎ口は見えず，ポットは左に寄って見える．一方，B地点からは，ポットは中央に見え，A地点と比較して窓から遠ざかった分，画面の中にポットが占める割合が大きくなる．以上のように，窓とユーザの位置関係によりポットの見え方は異なる．

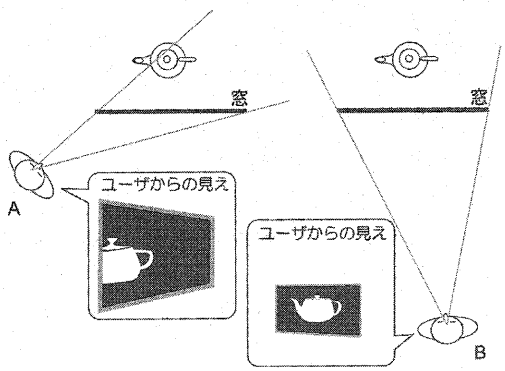


図1「窓越し現実感」の概念図

このような物体の見え方の違いを，ユーザの視点の位置に連動しながらディスプレイに表示することで，現実感のあるCG鑑賞環境が実現できる．

本手法は全方向から対象物を確認したい場合には不向きではあるが，実際のスケール感を保ったまま表示することが可能である．このため，建物の窓からの景色や，建築・都市計画，インテリア，自動車などのように，ある程度のスケールがあり，人の目にする角度が限定的な3次元物体に対して，形状の検討や鑑賞のための手法として有用性が高いと考える．

2.2 関連研究

位置検出技術としては，磁気や光学式のモーションキャプチャ装置，RFIDタグを用いたもの[2]などがあるが，手軽なものとは言えない．一方，画像認識による非接触式の姿勢・位置検出法もいくつか提案されている[3, 4]．本手法

では画像認識を採用し，ピクセルシェーダを用いることで既存の手法とは異なるアプローチで高速処理を行った．

人間の位置に応じたディスプレイの研究として，Lumisight[5]や，Illusion Hole[6]がある．Lumisightは，視界制御フィルムという特殊な光学特性をもつフィルムをスクリーンとし，視線方向による映像の切り替えを行い，多方向からの映像鑑賞を可能とした．この方式では，表示の変化は離散的になる．それに対し我々の方式では，身体動作に対して表示する映像を連続的に変化させることが可能である．Illusion Holeは，PDPを水平に配置し，その上部を中心に穴のあいたマスクで覆うものである．これにより，観測者の視野がユーザの位置に依存し，その位置に応じた形状を表示する仕組みである．この方式では，複数人で利用することが可能であるが，構造上表示できる像のサイズはPDPサイズの数分の1となってしまう．

また，複数台のカメラで撮影された人物画像に対してコンピュータビジョンを応用して3次元モデル化し，その3次元人物モデルとリアルタイムで対話できるようなシステムも報告されている[7]．さらに，複数台のカメラを配置し，多対多のビデオによるグループミーティングのシステム[8]も提案されているが，これらの研究は遠隔コラボレーションを主たる目的としており，我々と目指す方向が異なる．

3. システムの構成と実装

本章では，はじめにシステムの概要を述べ，つづいてシステムの各構成要素を述べる．

3.1 システムの概要

本システムは，図2に示すように，IEEE1394接続のDVカメラ（Sony DCR-PC110）2台とPC，そして大型ディスプレイ（LCDやPDPなど）により構成される．

DVカメラは，体験者の正面と側面に設置した．ただし，設置場所に関しては任意である．DVカメラの画角は，水平画角41度 垂直画角28度 （35mm フィルムレンズ換算 48mm）である．正面に設置したカメラには，死角を減少させるために0.5倍ワイドコンバージョンレンズ（吉田産業株式会社製 レイノックス DVR-5000）を取りつけており，水平画角74度 垂直画角53度となっている．

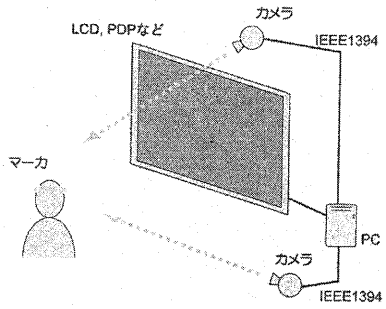


図2 システム構成

3.2 頭部の位置検出法

本システムでは頭部位置検出を画像処理により行う．

3.2.1 処理の概要

体験者には図3に示すように，頭部に黄色のマーカ付きの帽子を装着させ，2台のカメラから取得されるフレーム画像を解析してマーカの位置を算出する．その手順を以下に示す．

1. DVからのフレーム画像の切り出し
2. ピクセルシェーダによる色解析
3. ビデオメモリに解析結果を描画
4. ビデオメモリからシステムメモリへデータ転送
5. メモリから解析結果を読み出して認識色の重心を計算
6. 重心座標がカメラから実世界のどの方向にあるかを割り出す
7. 割り出した2つの方向ベクトルの交差点

を頭部位置とする.

ここで，1-5 の手順が色検出，6-7 の手順が位置検出に関する処理である.

色検出処理の手順は大きく 2 つにわけられる．1-3 までの処理は GPU 上で処理を行い，4-5 を CPU にて行う．このように，GPU と CPU とで負荷分散させることで，DV カメラ(解像度 720×480 ピクセル)2 台の解析と，映像コンテンツの描画を 1 台の PC でリアルタイム処理させる事が可能となった.

図 3 頭部位置検出に利用したマーカ付き帽子

3.2.2 GPU による前処理

本システムでは，頭部移動に対し,映像がリアルタイムで応答することが要求される．しかし，動画像処理はリアルタイム処理のボトルネックとなる.そこでグラフィックカードに搭載されているピクセルシェーダを用いることで,画像処理の高速化を試みた.

ピクセルシェーダを画像解析に用いることで,これまでピクセルごとに処理していたものを画像バッファに対して 1 度での処理が可能になる.

本手法では,画像内のマーカの色にあたる黄色の抽出にピクセルシェーダを用いた.抽出後には黄色と判断されたピクセルは緑の単色(R:0 G:255 B:0)で塗る.これは後にピクセルを走査する際の負担軽減のための処理である.

黄色の判定は，式(1)の条件をすべて満たすものとした．ここで，ピクセルの色成分を(R,G,B)と表現し，それぞれ 0.0～1.0 の範囲の値を持つ.

$$\left.\begin{aligned} &(R^2+G^2+B^2)^{1/2} < 0.8 \\ &R/B > 1.7 \\ &G/B > 1.7 \\ &|R-G| < 0.35 \end{aligned}\right\} \quad (1)$$

ピクセルシェーダによる画像解析は高速であるが，処理内容を直接 CPU に返す事ができない．すなわち，色解析後の座標などを CPU 側で取得できないことになる.

そこで本手法ではピクセルシェーダで処理するサーフェイスを,体験者に見せるシーンとは別の独立したバッファにレンダリングすることでピクセルシェーダの処理結果を読み取れるようにした．CPU ではそのサーフェイスを読み取ることにより,ピクセルシェーダで解析した結果を取得することにした．ここで，ピクセルシェーダで処理したサーフェイスは，CPU からアクセスできないビデオメモリ上に置かれるので,システムメモリ上に転送させる必要がある．この処理では AGP バス経由でデータを転送するので一定の時間を要するが,次のステップで CPU がサーフェイスのバッファを読み込む際にはビデオメモリを読み込むより高速に処理できるという利点がある.

3.2.3 CPU による処理

ピクセルシェーダでは色認識後の座標取得が出来ないため,その部分は CPU で処理する．CPU 側は処理後のサーフェイスのピクセルを読みとり,緑の単色で塗り替えられたピクセル集合の重心を求めた．そして，この重心座標を 1 台のカメラから見えるマーカの位置とした.

3.2.4 位置の算出

体験者の頭の位置は 2 台のカメラから取得した重心座標値から算出する.

まず，図4に示すように，カメラの視線のベクトルとカメラから見た重心座標へのベクトルのなす角θをそれぞれのカメラについて求める．このなす角は，式(2)に示すように，画像の中心からの水平方向の差分と水平画角とで求められる．

$$\theta = \frac{\text{画像の中心からの重心の水平方向の差分}}{\text{画像の水平解像度 } /2} \times \text{水平画角} \qquad (2)$$

ここで求めた，それぞれのカメラから重心への直線の交点から頭部位置のX座標とZ座標が求まる．同様に，垂直画角を元にX座標とY座標が求まる．以上で求めた頭部位置の座標値を，3.3で述べる表示法で利用する．

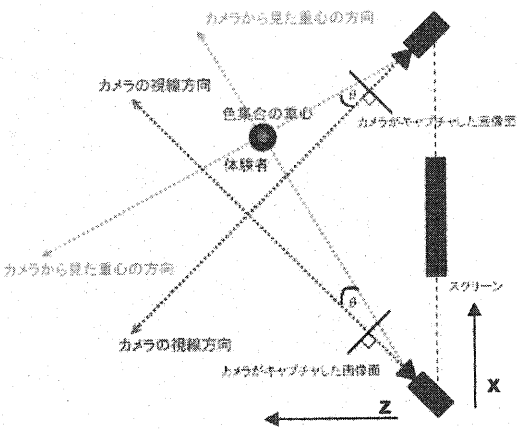


図4　頭部位置検出

3.3 映像の表示法

本システムでは，窓から見える外の風景を3DCGで表示する．3DCGの描画の際には，どこから描画対象物を見るかという視点の情報,および3次元空間から2次元平面への射影行列を求められればよい．本システムでは，観測者の頭部にマーカを装着し，前述した動画像処理で求められたマーカの位置情報をもとに視点位置と射影行列を求め，シーンの描画を行う．

図5(a)のような場合では一般的なパースペクティブ射影行列を用いても問題が無い．しかし，図5(b)のような場合には，不自然なシーンが描画されてしまう．これは，パースペクティブ射影行列においては，前方クリップ面が視線のレイと直交しており，視線のレイに対して左右対称な視野を作成するという前提条件のためである．したがって，前方クリップ面として想定されるスクリーンと視線が直行しない場合には，一般的なパースペクティブ射影行列では正しい透視射影を実現する事ができない．

そこで，図5(c)に示すような非対称なパースペクティブ射影のための射影行列を，式(3)で与える．

$$\begin{vmatrix} \frac{-2zn}{r-l} & 0 & 0 & 0 \\ 0 & \frac{-2zn}{t-b} & 0 & 0 \\ \frac{l+r}{r-l} & \frac{t+b}{t-b} & \frac{zn+zf}{zn-zf} & -1 \\ 0 & 0 & \frac{-2\cdot zn\cdot zf}{zn-zf} & 0 \end{vmatrix} \qquad (3)$$

但し，

zn：前方クリップ面までの距離

zf：後方クリップ面までの距離

l：クリップ面の左端のx座標

r：クリップ面の右端のx座標

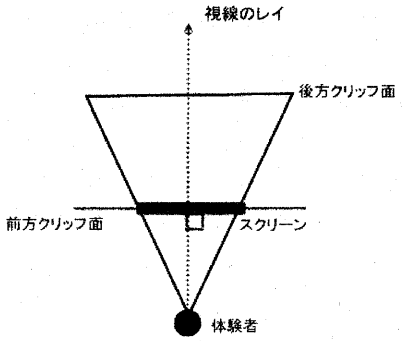
t：クリップ面の上端のy座標

b：クリップ面の下端のy座標

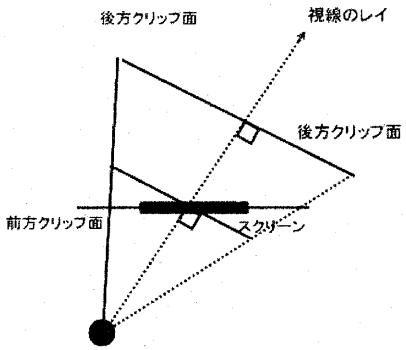
ここで，l < r および b < t が成り立つものとする．

式（3）を用いることで，スクリーンに正対せずに覗き込むような視点からでも，正しい射影を実現できる．システム上では，znは体験者からスクリーン平面への垂線の足の長さである．*l,r,t,b*の各値は体験者の位置とスクリーンの幅，高さの実測値を用いて算出する．また，

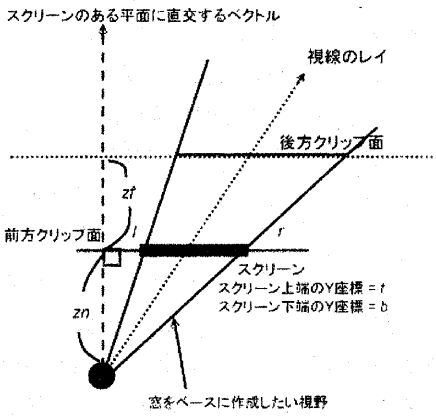
zfは 1000m の固定値にした.



a.体験者の視線がスクリーンの中央に直交する場合
パースペクティブ射影行列でも視野を作成できる



b.体験者の視線がスクリーンと直交しない場合
パースペクティブ射影行列では望ましい視野を作成する事ができない



c.スクリーンを窓と見立てたときの正しい視野
非対称なパースペクティブ射影のための行列で作成

図 5 体験者と視野の関係

4. 映像コンテンツについて

ViewFrame を応用したコンテンツとして「借景」を制作した. 借景とは, 古寺にみられる造園技法のひとつで, 近景としての庭園に, 自然に広がる遠景を借りいれることで作り出される風景の事を指す[10,11]. 本コンテンツでは, 自然の景観を ViewFrame の設置場所に借りてくること事を「借景」と呼び, 無機的な室内環境に潤いを与えることを目的のひとつとする.

本映像コンテンツは, 近景としての紅葉樹, 中景としての庭園風景, そして遠景としての山々の 3 層から構成されている. 素材は実写画像を利用し, 図 6 に示すように, ポリゴンにテクスチャとして貼ることで 3 次元空間に配置した. 画面スクリーンショットを図 7 に示す.

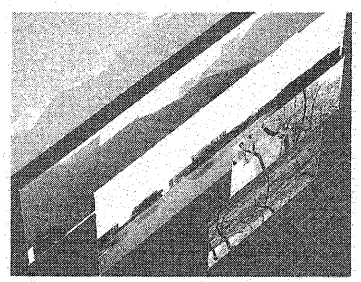


図 6 三次元空間での配置

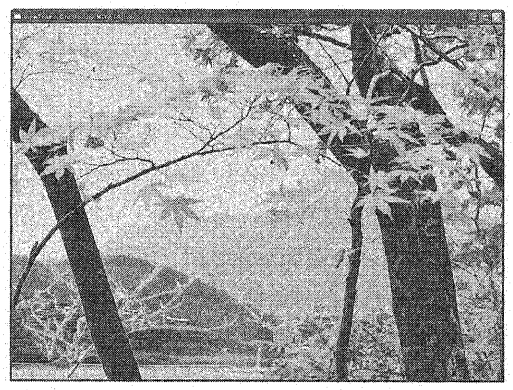


図 7 「借景」のスクリーンショット

5. 結果と考察

「借景」の鑑賞中の様子を図８に示す．システムに用いた PC のスペックは，Pentium4 2.8GHz, RAM 1GB, RADEON 9700 Pro 128MB であり，60 フレーム/秒で動作した．鑑賞を行った結果，滑らかに映像が変化し，高い現実感が確認された．

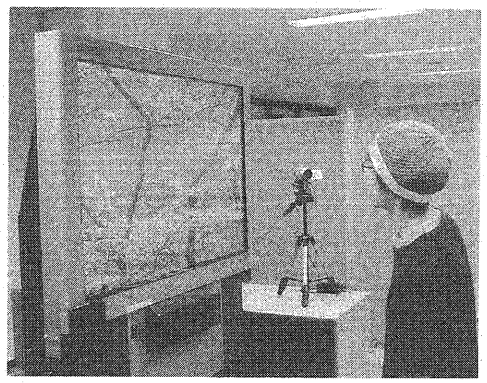


図８ 体験の様子

本システムの新規性として，ピクセルシェーダを動画像処理に用いたことが挙げられる．これによりビデオメモリ上に展開した配列データを高速に扱えるようになり，リアルタイム画像処理に有用なアプローチであると考える．

最後に，本作品の今後の発展性について述べる．人体検出法については，マーカを用いない方法が期待される．身体動作を用いたインターフェースでは，高いリアルタイム性が求められる．本手法では，処理速度と精度の向上をねらうため頭部検出用のマーカを用いたが，結果 60 フレーム/秒の映像表示を実現できたため，今後は肌色検出などのマーカを用いない手法の検討を行いたい．

実写画像を 3 次元モデル化する方法としては，NPR 技術[9]の利用が考えられる．今回は借景の特徴である，平面的な景色が重なりあい，切り取られた後景が見る位置によって移り変わるありさまを表現するため，３次元モデルの採用を見送った．しかし映像コンテンツの性質によっては効果的であると考えられる．

また，液晶シャッター眼鏡を利用した立体視化や，半透明ガラス 投影を用いた実世界への重畳化による新しいコンテンツも期待される．

今後，より多くの人間によって本システムの評価実験を行い，本アプローチの有効性について検証していく予定である．

謝辞

本研究の一部は，(財)コニカ画像科学振興財団および文部科学省科学研究費補助金(基盤研究（ｃ)) の助成により行われた．

参考文献

[1]C. Cruz-Neira, et. al, “Surround-ScreenProjection-Based Virtual Reality : The Design and implementation of the CAVE, Computer Graphics, pp. 135-142 (1993)
[2] 椎尾一郎, "RFID を利用したユーザ位置検出システム", 情報処理学会研究会報告 00-HI-88, pp. 45-50
[3] 米元ほか, “多視点動画像処理による非接触式実時間モーションキャプチャシステム”, 情処研報 CVIM, 119-10, pp. 71-78, 1999
[4] http://www.mic.atr.co.jp/~mao/midas/
[5] 川上 玲, 筧 康明, 苗村 健, 原島 博: ``視界制御フィルムを用いた方向依存透明ディスプレイ Lumisight'', 信学技報 HIP2003-15, pp. 1-6, 2003
[6] 北村喜文ほか, “多人数共有型立体ディスプレイ IllusionHole”, 映像情報メディア学会誌, Vol. 57, No. 10, pp. 1320-1327, Oct 2003.
[7] Chen, Wei-Chao, et. al., "Toward a Compelling Sensation of Telepresence: Demonstrating a portal to a distant (static) office," Proceedings of IEEE Visualization 2000.
[8] Yang, Ruigang, et. al., "Creating Adaptive Views for Group Video Teleconferencing -- An Image-Based Approach," International Workshop on Immersive Telepresence (ITP2002).
[9] Ken Anjyo, “non-photorealistic rendering: "Tour into the picture" as a non-photorealistic animation "
, ACM SIGGRAPH, Vol33, Issue1, pp54-56, 1999.
[10]
http://landship.sub.jp/stocktaking/archives/000016.html (2003.10.8 現在)
[11] http://web.kyoto-inet.or.jp/people/yasugi/diary47.htm (2003.10.8 現在)