

マーカレス顔部検出手法を利用した”ViewFrame”

河原塚有希彦 高橋誠史 宮田一乗

北陸先端科学技術大学院大学・知識科学研究科

1. はじめに

「窓越し現実感」というコンセプトのもとに、ディスプレイを「窓」と想定し、ユーザが観察すると予想される風景を、ユーザの位置にあわせて対話的に描画するシステムが提案されている[1] 本報告では、ユーザにマーカを装着させるという制約を設けず、手軽かつ自然な身体動作で映像コンテンツに疑似3次元的にアクセスするための手法を提案する。

2. システムの構成

本システムは、図1に示すように、IEEE1394 接続のDVカメラ1台とPC、赤外線測距センサ、そしてディスプレイにより構成される。DVカメラを体験者の正面に設置し、赤外線測距センサとの組み合わせでユーザ顔部の位置検出を行い、その位置に応じた画像を実時間で生成する。



図1 システムの概要

類似のシステムに、3DV System社のZCam™[2]が挙げられる。ZCamでは、赤外線による奥行き情報とカメラからの画像情報を同時に取得し、テクスチャ情報を持った3次元形状を計測する。また、同様のセンサを用いたEyeToy™[3]という名称のゲーム用デバイスも市販されている。本質的な原理の違いはないが、本手法では、測距センサは画像処理の精度向上に用いるという点でコンセプトを異とする。

3. 顔部の位置検出法

本システムでは、DVカメラからの入力画像と赤外線測距センサからの入力データを用い、顔部位置を算出する。赤外線測距センサアレイは奥行きと横方向位置の検出に用い、DVカメラからの入力画像は、顔部の横、縦方向の位置の検出に用いる。ここで、測距センサアレイから取得される横方向の位置情報を用いることで、画像の認識率の向上が可能となる。

3.1 赤外線測距センサによる位置取得

本システムでは、図2に示すように、赤外線測距センサ(Sharp GP2D12^{*1})をディスプレイ上部に横方向に4個並べ設置した。そして、センサから得られる情報に基づいて、三角測量の原理で、センサから人体までの奥行きと横方向位置を算出した。



図2 赤外線測距センサの設置

(*1:測距範囲 10-80cm, 測距結果を電圧(リア)出力)

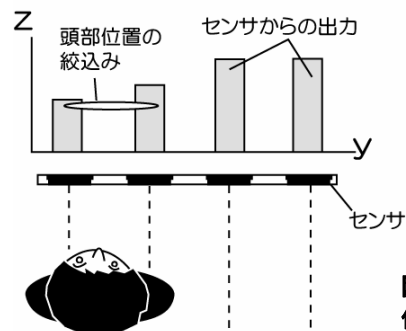


図3 測距センサによる位置検出の絞り込み

赤外線測距センサから取得した横位置情報は、後述する画像認識による顔部位置の検出の際に、位置の絞り込みのために用いる。これにより、画像認識の際に、体験者以外の人間や背景等の影響を受けにくくすることが可能となる。

Marker-less "ViewFrame"
Yukihiko Kawarazuka, Masafumi Takahashi, Kazunori Miyata, School of Knowledge Science, JAIST

3.2 DVカメラによる位置取得

DVカメラからの入力画像を利用して、横、縦方向の顔部位置を検出する。

本システムではカメラから取得されるフレーム画像を解析し、以下の手順で顔の位置を算出する。

- | |
|-----------------------------|
| 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. 取得した色重心と測距センサの値から、頭部位置決定 |

ここで、1-5の手順が色領域検出、6の手順が位置検出に関する処理である。色領域の検出処理においては、1-3までの処理はGPU上で、4-5までの処理をCPUにて行い、負荷分散させた。

本システムでは、GPU上のピクセルシェーダ機能を用いて画像処理の高速化を試みた。画像内の顔領域にあたる肌色の判定条件は、式(1)をすべて満たすものとし、 $L*a*b*$ 色空間で処理を行った。

$$\left. \begin{array}{l} 20 < L < 75 \\ 5 < a < 25 \\ 0 < b < 20 \\ 0 < b-a < 15 \end{array} \right\} (1)$$

ピクセルシェーダによる画像解析は高速であるが、処理結果を直接CPUに返す事ができないため、解析結果を別バッファにレンダリング（緑の単色で塗る）し、後の処理で利用することとした。

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、この重心座標値を1台のカメラから見た顔の位置とした。

4. 画像の表示法

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の描画には、式(2)で与えられる非対称なパースペクティブ射影のための射影行列を用いて映像を表示する。

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

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（1000mに固定）

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 : クリップ面上端のy座標

b : クリップ面下端のy座標

ここで、 $l < r$ および $b < t$ が成り立つものとする。式(2)により、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。

5. 結果と考察

「借景」の鑑賞中の様子を図4に示す。システムに用いたPCのスペックは、PentiumM 1.4GHz, RAM 768MB, NVIDIA® GeForce™ FX Go5200 64MB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4 体験の様子

本システムでは、ピクセルシェーダを動画処理に用いたことで、リアルタイムでの処理が可能となった。また、赤外線測距センサを画像認識の補助として用い、認識率の向上につながった。

今後は、本システムを応用し、液晶シャッター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。

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] 河原塚 他, "ViewFrame - 画像処理による位置検出法を用いた「借景」", 情処ヒューマンインタフェース研究会, 2003-HI-106 (7), pp.45-51
- [2] http://www.3dvsystems.com/technology/Gabi_and_Giora_Article.PDF
- [3] <http://www.eyetoy.com/>