

空間共有通信における透過型ビデオアバター

See-Through Video Avatar for Shared Space Communication

松下 剛士† 新田 拓哉† 苗村 健‡ 原島 博†
†Takeshi Matsushita †Takuya Nitta ‡Takeshi Naemura †Hiroshi Harashima
東京大学工学部† スタンフォード大学‡
†School of Eng., The Univ. of Tokyo ‡Stanford Univ.

Abstract: The concept of shared-space communication has been attracting much attention in the field of virtual reality. A person in the real world can appear in the shared virtual space as his/her avatar. The concept of video avatar is one of the promising approaches to realize a photo-realistic avatar of a real person. For this purpose, the authors have proposed and developed the “See-Through Video Avatar” system, which is a kind of clipping/composition method using alpha-blending technique. The alpha value is designed to relate with the probability of each pixel to belong the regions those correspond to human face and body. Our implementation has been a thermal-vision based system, in which the alpha value is controlled by the temperature of human region. In this paper, some improvements for our system are presented. First, the background subtraction technique is applied to thermal images in order to improve the quality of synthesized images. Secondly, a new capturing system, which can avoid the parallax between thermal and color images, is introduced. Thirdly, we propose some new applications of the translucent video avatar by considering how to control the transparency. Experimental results show that the proposed method is effective and useful for the shared space communication.

Keywords: video avatar, shared space communication, thermal vision, alpha blending, translucent and transparent expression

1 はじめに

昨今、大画面表示装置 (IPT:Immersive Projection Technology) などを用いた空間共有通信が関心を呼んでいる。筆者らは、実世界の人物を仮想空間に登場させる技術として「透過型ビデオアバター」を提案し、サーマルビジョンカメラを用いるシステムを開発してきた [1-3]。本稿では、そのシステムの改良の報告及び「透過型」であることを利用した新しいコミュニケーションの提案を行う。

2 従来の研究

空間共有通信とは、例えば相互接続した IPT に同じ仮想空間を構築するなど、様々なレベルにおける空間の共有を実現する通信システムのことである。この際には、ユーザの代わりにアバターを分身として登場させる。文献 [4] や文献 [5] などは CG をベースとしたアバターの例であり、立体表現を比較的容易に実現できるという特長がある。

一方、より高い写実性を求める観点からビデオアバターが注目され、その実現のために実写画像が

ら人物を切り出し仮想空間内に合成する技術が要求されている。これには、従来から 2 値判定に基づくものが多く研究され、最近ではクロマキーと距離画像を併用する手法 [6] や同期式クロマキーを用いる手法 [7] などが提案されている。しかし 2 値の切り出し手法では、精緻な切り出しと実時間性の両立が困難な問題や、撮影時の背景に制約がある問題などが存在する。

そこで筆者らは、あえて 2 値ではなく、例えば 256 階調の重み付けによる、多値の領域指定を行う「透過型ビデオアバター」を提案している [1]。これは、「人物領域である確からしさ」に応じて実写画像と仮想空間画像を加重合成するものであり、人物であることが「確かな」領域は透過させずにはっきりと表現し、逆にあいまいな領域は半透明に、背景が透けて見えるように表現する。

「人物領域である確からしさ」としては、人間の体温に着目し、サーマルビジョンカメラで得られる熱画像を用いたシステムを提案している [1, 2]。文献 [2] のシステムでは、サーマルビジョンカメラと高感度の CCD カメラを並べて配置し、得られた熱画像に Fig.1 に表される階調処理（熱画像輝度値を入力、 α 値を出力とした写像）を施す。

この結果得られる画像を α マップとしたアルファブレンディングを、グラフィクスエンジンを用い

連絡先: 苗村 健
東京大学工学部電子情報工学科
〒113-8656 文京区本郷 7-3-1
TEL : 03-5841-6668
E-mail: naemura@hc.t.u-tokyo.ac.jp

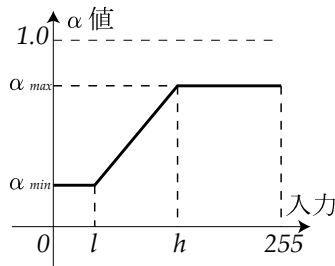


Fig.1: From thermal data to α -value

て高速処理することによって実時間で行う。さらに、Fig.1の写像是プログラム上で変更することができるため、完全透過とする温度範囲やアバタのコントラストなどの設定が容易に可能である。

しかし、この従来システムには次のような問題点がある。

1. 温度の高い領域を人物領域としているため
 - 温度の高い物体（お湯の入ったポットなど）は映りこみ、温度が低い着衣の一部や眼鏡などは透けてしまう
 - 撮影環境の気温が高い場合には人体との温度差が小さくなり、人物だけを切り出すことが困難になる
2. 2台のカメラの光軸が一致していないため視差が生じ、奥行き異なる被写体に対して α マップとカラー画像がずれてしまう
3. 視点位置やアバタの登場位置が固定のため臨場感に乏しい

1.の問題は文献[2]の階調処理によってある程度改善されているが、まだ不十分である。

3 提案手法

3.1 α マップの効果的生成—熱画像背景差分

文献[2]の階調処理では画像全体に同じ写像を適用するため、暖かい物体や温度分布のむらがある場合には、人物領域のみの切り出しが困難になる。

そこで、熱画像に対して背景差分を適用することを考える。すなわち、無人状態における熱画像を背景熱画像として予め撮影しておき、現在の熱画像との差分値を求めて α 値の設定に利用する。この手法を用いれば、背景の温度と異なる温度分布を持つ領域のみを容易に切り出すことができる。こうして得られる差分熱画像に対してさらに文献[2]同様に階調処理を適用することにより、アバタの透明度やコントラストも自由に設定することが可能である。Fig.2はそのブロック図である。

CCDカメラで撮影された画像は、1画素あたり

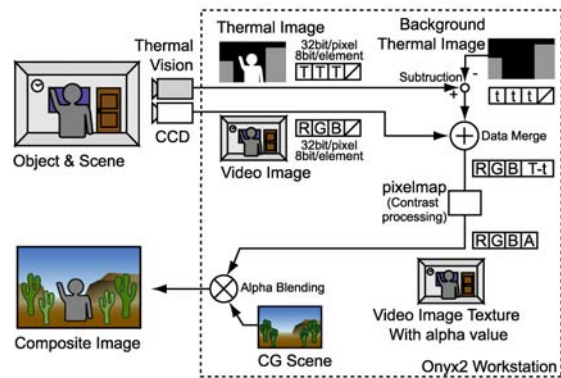


Fig.2: Block diagram

RGBA4要素のデータ列として扱われる。入力熱画像・背景熱画像の輝度値差とFig.1の写像とから得られる透過係数 α をAの部分に1画素毎に挿入することで、 α 値に応じた混合処理がなされ目的の合成画像を得ることができる。

3.2 光軸一致の実現

一般のハーフミラーは赤外線の透過・反射特性が悪いため熱画像に十分なコントラストを得ることができず、2つのカメラの光軸一致の用途には適していない。そこで、赤外反射ミラー（赤外光反射・可視光透過）とシリコンミラー（可視光反射・赤外光透過）とを用い、カメラの位置・姿勢の調節機構を備えた撮像系を構築した（Fig.3）。

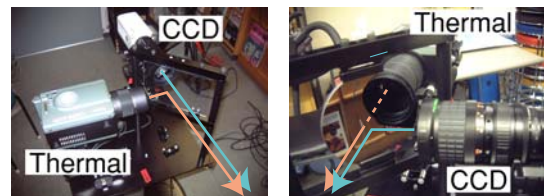


Fig.3: Optical system(Ir. mirror and Si mirror)

3.3 空間共有通信に向けて

アバタは、観察者の高さにはほぼ等身大で現れるように位置と大きさを調節した書割として、実物大で構築した仮想空間内に配置することで登場させる。より高い臨場感を得るために、本研究室所有の2面ディスプレイRICUEを用い、両眼立体視によって視点位置に応じた立体映像を提示する。また、空間共有通信でのアプリケーションとして、「透過表現を用いたコミュニケーション支援」を提案する。

多くの場合、複合現実環境においては、現実世界で可能な（生じる）事象を忠実に再現する、あるいは状況の限定された空間を構築するというアプ



(a) Thermal(bg.) (b) Thermal(input) (c) Color

Fig.4: Original images



Fig.5: Synthesized by method of [1]

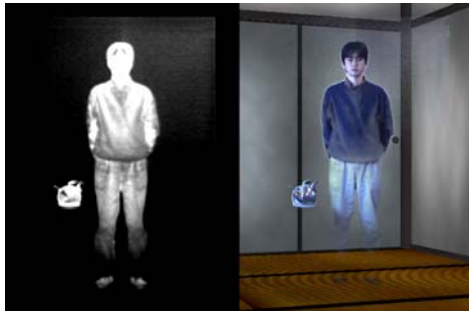


Fig.6: Synthesized by method of [2]

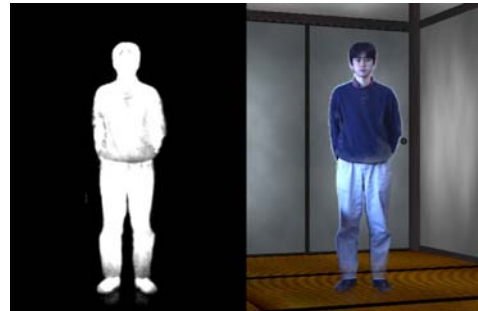


Fig.7: Synthesized by proposed method

ローチから臨場感を高め、それによってスムーズなコミュニケーションを図っている。しかしここで提案したいのは、非現実的事象によって逆に表現力を増大させるという考え方であり、そのために透過表現が利用できると考えている。

ここでは、「関心・注目の度合い」をアバタをはじめとする仮想オブジェクトの透過率に対応させる。つまり、関心がある、あるいは注目すべきものははっきりと登場させ、関心の薄い、それほど重要でないものは半透明に登場させるのである。さらに、この透過率をインタラクティブに変化させることにより、参加者同士のコミュニケーションがスムーズになると期待できる。

たとえば複数のアバタが登場する場面において、普段は半透明にしておき、動いたり喋ったりしているアバタだけをはっきりと登場させれば、どのアバタが動いたり喋ったりしているのかがよくわかる。あるいは指差し先や視線の先にあるオブジェクトのみをはっきり登場させれば、どれに注目して欲しいのかが相手に明確に伝わるであろう。

4 実験結果

4.1 α マップの効果的生成—熱画像背景差分

赤外反射ミラーを用いた場合の実験結果を、Fig.4 から Fig.7 に示す。Fig.4 はそれぞれ、背景熱画像、入力熱画像、入力カラー画像である。Fig.5

は入力熱画像をそのまま α マップとする文献 [1] の手法、Fig.6 は入力熱画像に対して適当な階調処理を施す文献 [2] の手法、そして Fig.7 は、入力熱画像と背景熱画像の差分をとり、さらに階調処理を施す本手法による合成結果である。それぞれ左側が α マップ、右側が合成画像である。

Fig.6 においては、Fig.5 に比べ人物領域を比較的良好に切り出し・合成しているが、画面下部は全体に温度が低く、透過してしまっている。また、画面左にあるお湯の入ったポットは温度が高く、はっきりと現れてしまっている。一方 Fig.7 においては、背景と温度の異なる部分のみに反応して、ポットを切り出すことなく人物のみを鮮明に合成することに成功している。さらに、背景熱差分画像によって僅かな温度の違いを検出できるので、温度の低い眼鏡部なども透過しないようにできる。

背景差分を行う場合は行わない場合に比べ更新周期が低下するが、実時間動作が確認できている。

4.2 光軸一致の実現

カメラを光軸が平行になるように配置し拡大縮小・平行移動の補正を加える従来手法 [1, 2] と、Fig.3 の撮像系を用いた本手法とで比較を行った。光軸あわせは手作業で、合成結果の目視で行った。また厳密には焦点位置も一致する必要があるが、ここではずれが目立たない程度まで調整した。



Fig.8: Synthesized from parallax capturing system



Fig.9: Synthesized from non-parallax capturing system

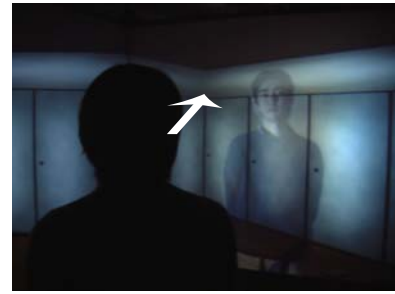


Fig.10: The direction of observer controls appearance of avatars

従来手法 (Fig.8) では、手前の人物の切り出しがずれているが、本手法 (Fig.9) では奥行き異なる被写体も概ね正確に切り出すことに成功している。レンズ歪等の問題もあり、実際には画像が完全には一致しないが、原理的には光軸と焦点位置が一致していればソフトウェアによる補正が可能である。本実験では赤外反射ミラーを用いたが、シリコンミラーを用いた場合も同様の結果が得られる。

4.3 空間共有通信へ向けて

RICUE 内でアバタを立体位置に表示することにより、より高い臨場感が得られるようになった。

透過表現を用いたコミュニケーション支援については、プロトタイプとして動いているとき、喋っているとき、そして視線の先にあるときにアバタをはっきり登場させるシステムを実装した。

Fig.10 は視線の先にあるアバタをはっきり表した例である。視線方向や動きの度合い、喋る声の大きさをキーとして取得し、そのキーによってアバタの最大 α を決定することで実現した。さらにキー値と最大 α の対応関係に時間遅れを持たせることで、人が動くなると次第にアバタが現れ、止まると次第に消えるという振舞いを実現できた。

5 まとめ

従来システムの問題点に対して解決策を提示し、より良好な合成結果を得た。また、IPT 内での高

度な表示を可能にするとともに、半透明アバタによる簡便かつ効果的なコミュニケーション支援を提案し、その可能性を示した。

今後の研究課題として以下の点を挙げる。

- より正確で効果的な、撮影条件に左右されない α マップの生成
- 手間がかからず確実に光軸合わせを行うための装置や手順、カメラ個体の特性 (レンズ歪みなど) を適切に補償する手段の開発
- 透過表現の適用対象を仮想オブジェクトに広げること及び、豊かな表現力を持たせるためのさらに高度なインタラクションの実現

参考文献

- [1] 川原, 苗村, 原島: サーマルビジョンカメラを用いた透過型ビデオアバタ, 信学総大, A-16-22, 2000.
- [2] 川原, 松下, 新田, 苗村, 原島: 透過型ビデオアバタの..., VRSJ 大会, Vol.5, pp333-336, 2000.
- [3] 松下, 新田, 苗村, 原島: 透過表現を用いたビデオアバタの改良と応用, 信学総大, A-16-17, 2001.
- [4] 本田, 木村, 大沢, 大田, 岡田, 松下: 人間の現実動作の..., VRSJ 論文誌, Vol.4, pp377-388, 1999.
- [5] 前田, 岡田: ネットワーク型 VR システムによる..., VRSJ 大会, Vol.5, pp267-268, 2000.
- [6] 廣瀬, 小木, 玉川, 山田: 没入型コミュニケーション..., HIS 学会誌 Vol.2, No.2, pp161-168, 2000.
- [7] 廣瀬, 小木, 加納, 山田: 没入型ディスプレイ間通信..., 映メ技報, Vol.24, No.34, pp49-52, 2000.