

リアルタイム生成3D人物顔を用いたコミュニケーションシステム

箕浦 大祐 松浦 宣彦 菅原 昌平 石橋 聡
NTTサイバースペース研究所

あまし 利用者が自分に似せた3次元人物顔モデルをリアルタイムに作成することができ、かつこのモデルを使って自分の顔を持つ分身を介したコミュニケーションを行うことができる3次元仮想空間共有システムを開発した。本システムは、3次元仮想空間で広角度・遠距離から他の参加者の分身を見て誰であるかわかること、かつ汎用的な装置を用いて誰でもその場で簡単に顔モデルを作成できることを目的としている。そのため3次元モデルとしての正確さは相手を識別可能な程度にとどめ、リアルタイムに顔モデルを作成できることが最重要であると考えた。実際に複数の利用者について顔モデルを作成した結果、良好な処理時間を得たと同時に、両目と口の位置の指定だけの簡単な操作で十分自然に見えるモデルを作成できることがわかった。

A communication system using a 3-D face model real-time creating tool

Daisuke MINOURA Norihiko MATSUURA Shohei SUGAWARA Satoshi ISHIBASHI
NTT Cyber Space Laboratories

Abstract: We have developed a multi-user communication platform where users show their faces as three-dimensional (3-D) models. Users can create their 3-D face models instantly and show the models as their avatars in the 3-D virtual world. The aims of the system are that a user is able to identify other users from wide angle at a distance in the virtual environment and that they can create their face models quickly before participation. We do not recognize that exact accuracy of shape of the models is important for the aims. We have got a short image-processing time to create a face model as a result of applying the creation method to many users. Also we have observed natural appearance of the model with the creation method in which a user only indicates the positions of his or her eyes and mouth on an input image.

1 はじめに

計算機のグラフィックス表示性能の向上にともなって、デスクトップPCやCAVE等の表示端末を使った3次元仮想空間を利用する研究がさかんに行われている。このような端末をネットワークに接続し、端末間で共有する3次元仮想空間内に表示される操作者の分身を通してコミュニケーションを行うシステムにおいて、我々はより現実のコミュニケーションに近づくように利用者の分身を表現する方法について検討を重ねてきた。^[1]

3次元仮想空間内にコミュニケーション相手の分身を表示する方法として、本人の顔画像を動画で表示する方法^{[2][3]}や複数の方向からビデオ撮影して作成したアバタで表示する方法^[4]などが検討されている。顔画像の動画の場合、リアルタイムに相手の表情を把握することができるという利点があるが、表示が平

面的であるために広角度・遠距離からの本人識別が困難である。また、複数の方向からビデオ撮影して作成したアバタの場合は、さまざまな方向から相手を認識することが可能である。しかしこのアバタを作成するために特別な装置を必要とするとともに、リアルタイムに作成することは困難であると思われる。

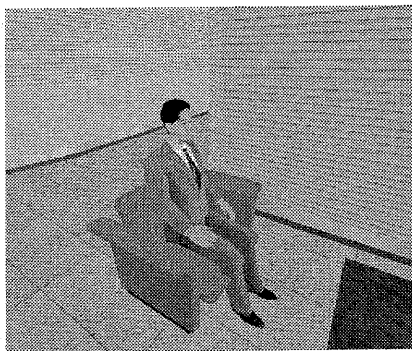
以上の背景を踏まえて、本研究ではブレ・コミュニケーションの情報として本人識別が重要なフォーマルな用途、かつ端末の操作者が頻繁に交代して利用するような用途を想定して、利用者本人の仮想人物モデルをリアルタイムに生成して操作者が自分の3次元顔モデルを使って参加できる仮想空間コミュニケーションシステムを開発した。

2 システム構成

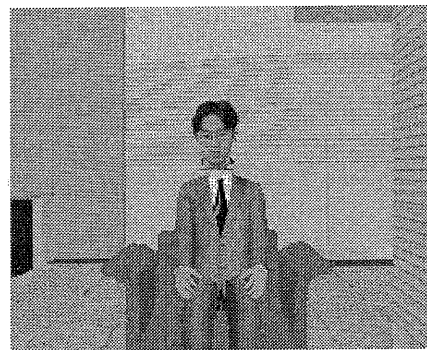
2.1 3次元仮想空間コミュニケーションシステム

我々は、CGで描かれた3次元空間を複数の利用者間で共有してコミュニケーションを行うためのシステムについて研究開発を進めている。利用者がこの空間に参加すると、自分の分身が相手の画面に表示される。利用者が仮想空間内をウォークスルーすると、その分身もウォークスルーしている視点の位置にあわせて移動する。利用者はコミュニケーション相手の分身の姿と、その分身の顔部分に表示される相手の顔のビデオを見ながら、音声と文字チャットを利用してコミュニケーションを行うことができる。本システムはインターネット上で動作させることを目的としているため、単位時間に伝送可能な情報量は制限される。したがって顔のビデオも相手の表情がリアルタイムに把握できるという特徴がある反面、仮想空間内での相手の分身から少し離れるとそのビデオのサイズおよび画質は誰であるか容易に判別するために十分とは言えなかった。またこのビデオは平面上に表示されているため、相手の前に回らないと顔から誰であるか判別できないという課題があった。

そこで今回我々は、操作者がリアルタイムに自分の3次元顔モデルを作成し、自分の顔で仮想空間に参加するコミュニケーションシステムを開発した。自分の顔モデルの表示では表情をリアルタイムに把握できなくなってしまうため、接近したコミュニケーションが必要な状況になると顔モデルから従来のように顔のビデオに切り替わる手法を採用した。本システムの利用画面例を図1に示す。



(a)CGモデル表示時



(b)ビデオ表示時

図1 本システムの利用画面例。

2.2 リアルタイム3次元人物顔作成システム

本システムにおけるコミュニケーション端末では、参加者が自分の分身として仮想空間に表示する3次

元アバタの顔モデルをリアルタイムに作成することができる。このモデル作成は図2に示すようにPCとそれに接続されるCCDカメラを組み合わせた汎用的なシステム構成で実現されている。端末には複数の種類の仮想人物の骨格モデルをあらかじめ用意する。利用者はこれらの骨格モデルから適当なモデルを選択する。この選択されたモデルに正面から撮影した顔画像に画像処理を施して作成される顔画像を貼り付ける。さらに利用者は骨格モデルと同様に用意された頭髪モデル群から適当なモデルを選択し、これらを組み合わせて3次元モデルとして自分の分身が端末画面に表示される。



図2 利用者端末のシステム構成.

3 人物顔作成システムにおける顔モデル生成処理

3.1 顔領域の抽出

入力画像は人物の正面から撮影した静止画像が1枚あれば十分である。この入力画像として撮影した顔画像の中で利用者に両目と口の位置を指定させる。これらの位置の近傍の肌色を参照して抽出すべき肌色の範囲を動的に決定し、顔領域を背景から抽出する。この処理の様子を図3に示す。



(a)入力画像



(b)抽出領域

図3 顔領域の抽出.

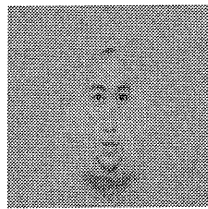
3.2 合成処理

抽出した顔領域の中の両目および口の位置とテンプレートとして用意された顔テクスチャ中の予め調べられた両目と口の位置を比較し、テンプレートのそれらの位置に合うように抽出顔領域に縦横方向の伸縮処理と回転処理を施す。この処理画像からマスクを使って余計な領域を削除した後、テンプレート顔画像上に添付することで合成顔画像を作成する。

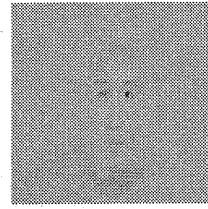
3.3 顔色の補正処理

撮影時の照明条件の影響を受けないように、また仮想空間に表示した時に適切な顔色になるように抽出顔領域の色を調節する。抽出した顔の領域の色の平均値がテンプレートの顔色の平均値と等しくなる

ように全体的な調節を施す。さらに頬や顎などの影による局所的な色の違いを補正するためのグラデーション処理を行い、骨格モデルに添付する。この処理の様子を図4に示す。



(a)単純合成処理



(b)顔色補正後

図4 顔色の補正処理。

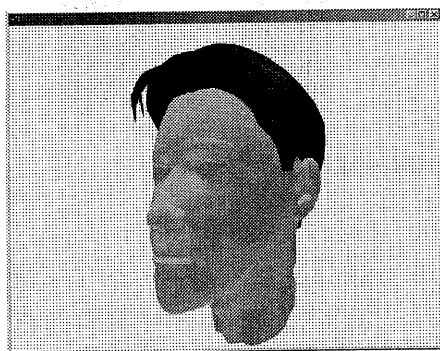
4 コミュニケーションシステムにおけるアバタの表示方法

4.1 顔画像の配信

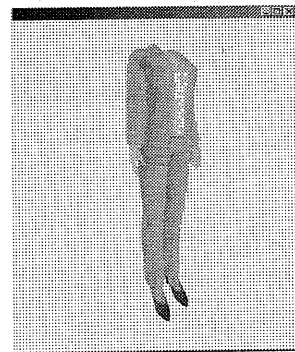
画像処理を施した顔画像は、コミュニケーション相手の端末画面に表示するために各端末に配信される。現在はあらかじめ指定した相手端末のみに配信されるように設定されている。この顔画像と組み合わせて表示される顔の骨格モデル、頭髪モデルおよび胴体モデルはあらかじめ各端末に保持されている。そのため、これらのモデルは配信される必要がなく、利用者が選択したモデルを示す識別子及び胴体モデルを決定するための性別識別子が顔画像とともに配信される。

4.2 モデルの結合

データの受信側の端末では、配信されたモデルの識別子から参加者が選択した顔骨格モデル及び頭髪モデルを決定する。また性別の識別子から男性用胴体モデルか女性用胴体モデルを決定する。決定されたこれらのモデルを組み合わせて人物モデルを形成し、これに配信された顔画像を添付することで相手の利用者の分身として表示する。この人物モデルを構成する各モデルの例を図5に示す。



(a)頭髪モデル及び顔骨格モデル



(b)胴体モデル(女性)

図5 人物構成モデル。

4.3 顔モデルとビデオ画像の切り替え

前述のように、顔を3次元モデルで表示したことによって、相手と接近したコミュニケーションで表情がリ

リアルタイムに把握できなくなってしまう。そこで接近したコミュニケーションではビデオによる表示と切り替える方法を採用した。従来のビデオ表示では、利用者の顔を撮影する際に背景も含めて四角いビデオとして表示するか、あるいはクロマキーバックを利用して背景を切り抜いて人物の頭部のみを表示していた。そのため表示された人物は人物としてのリアリティがなかったり、撮影側の端末の汎用性が低下するという問題が生じていた。そこで全てソフトウェアによる人物の背景除去処理を施した。この処理では背景除去のために、ビデオのフレーム間差分をとり、移動物体のある領域内のみを表示している。背景除去の処理が施された顔のビデオは各端末にマルチキャストされ、CG 胴体モデルの首の上に設けられたスクリーンに表示される。

5 利用例

本システムの利用例を示す。図6に入力画像(左列)及び生成された3次元人物モデル(中, 右列)を, 図7に顔のビデオ表示例を示す。利用者は両目と口の位置を指定した以外にもかわらず、鼻の頂点や唇など、顔骨格モデルの起伏にテクスチャが適切に貼られている様子が観察される。各モデルの作成時の

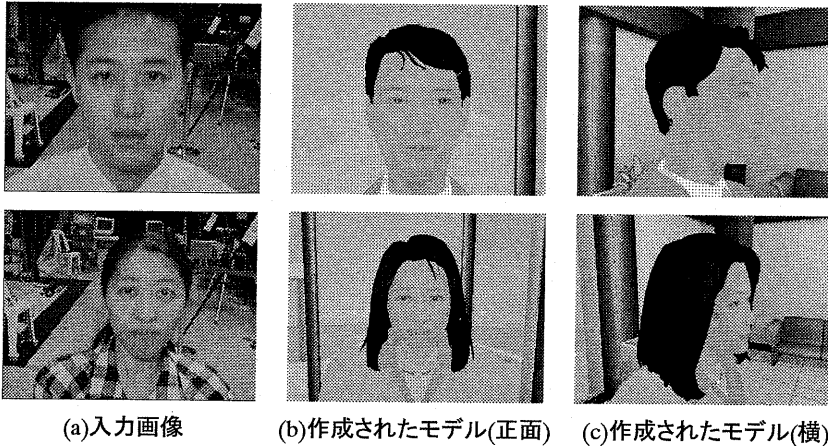


図6 本システムにおける3次元人物モデルの作成例。



図7 本システムにおける顔のビデオ表示例(拡大).

画像処理時間は Pentium700MHz の PC を用いて平均 10 秒であった。

顔モデルのポリゴン数は平均4700程度である。各クライアントが送信する画像は圧縮なしで平均1MBのデータサイズである。ビデオは毎秒3フレーム程度の速度で表示されることを確認した。

6 おわりに

リアルタイムに3次元人物顔モデルを作成し、自分の顔を持つ分身を介したコミュニケーションシステムの開発について報告した。本システムは、3次元仮想空間でのコミュニケーションシステムにおいて広角度・遠距離から相手がわかること、かつ汎用的な装置を用いて誰でもその場で簡単に顔モデルを作成できることを目的としている。そのためモデルの正確さは相手を識別可能な程度にとどめ、リアルタイムに顔モデルを作成できることが最重要であると考えた。

実際に複数の利用者に適用した結果、良好な処理時間を得たと同時に、両目と口の位置の指定だけで十分自然に見える顔モデルを作成できることがわかった。

今後は仮想空間への参加時に自端末に表示すべき参加者の顔画像を自動的に相手の端末から取得できるように顔画像の配信処理を改良するとともに、3次元仮想空間コミュニケーションにおける有効性を確認していく予定である。

参考文献

- [1] 箕浦大祐, 松浦宣彦, 松本敏宏, 前田泰宏, 菅原昌平, 正樹茂樹, “3次元仮想空間の参加者表示方法の検討,” 第59回情処全大, Vol.4, pp.119-120, 1999.
- [2] S. Sugawara, G. Suzuki, Y. Nagashima, M. Matsuura, H. Tanigawa, and M. Moriuchi, “InterSpace: Networked virtual world for visual communication,” IEICE Transactions on Information and Systems, Vol.E77-D, No.12, pp.1344-1349, 1994.
- [3] 中西英之, 吉田力, 西村俊和, 石田亨, “次世代ヒューマンインタフェース・インタラクション Free Walk 3次元仮想空間を用いた非形式的なコミュニケーションの支援,” 情報処理学会論文誌, Vol.39, No.5, pp.1356-1364, 1998.
- [4] 広瀬通孝, 小木哲朗, 山田俊郎, 玉川憲, KIM S., “CABIN間通信におけるビデオアバタの生成,” ヒューマンインタフェース学会研究報告, Vol.1, No.2, pp.43-46, 1999.