

テレビ電話のための仮想衣服着用システム

島田 英之・塩野 充

岡山理科大学工学部情報工学科

(1993年9月30日 受理)

1. まえがき

音声による手軽な情報伝達の手段として、電話は今日では社会に広く受け入れられるに至った。また、電話回線は広範囲に及ぶネットワークでもあることに注目し、最近では音声以外にファックスなどによる画像情報の伝送もごく普通に行われている。

このような背景のもとで、次世代の電話としてかなり以前からテレビ電話の研究開発が行われているが、近年の各種ハードウェアの性能向上の度合に比して、テレビ電話の普及は遅れているように思われる。この主な原因として、

- (1) 音声のみに比べ、画像の情報量（特に動画像）は膨大であるため、伝送にコストがかかる
- (2) 画像の圧縮や伝送の規格が統一されていない

などの問題がまず考えられる。これらの問題に対しては、顔画像から顔の形状や表情を検出し、高度に符号化して伝送する知的画像符号化の研究¹⁾などが行われている。しかし、誤った顔の表情の伝達はテレビ電話によるコミュニケーションにおいて大変危険な事態であるにもかかわらず、現時点で提案されている知的符号化方式は、当然考えられるべき眼鏡や頭髮などの影響を吸収し得るほどロバストなものとはいえない。

また、テレビ電話は、リアルタイムに画像を伝送できるという利点と同時に、撮影された画像（服装、部屋の様子など）がもし送信者にとって不都合なものであっても、リアルタイムに相手側に送信してしまうという欠点を持つ。そのため、上述の技術的な問題が解決され、テレビ電話が一般に普及した場合でも、心理的な抵抗感という新たな問題が生ずることが十分予想される。

本論文では、テレビ電話に存在する以上のような問題点をふまえ、人物画像中の衣服や背景を選択的に加工したのちに相手側に伝送することを目的として行った、仮想衣服着用システムのための基礎的実験について述べる^{2,3)}。

2. 仮想衣服着用システムについて

2.1 処理手順

本システムでの処理手順を図1に示す。実際に動画を処理する場合にはこの手順を動画

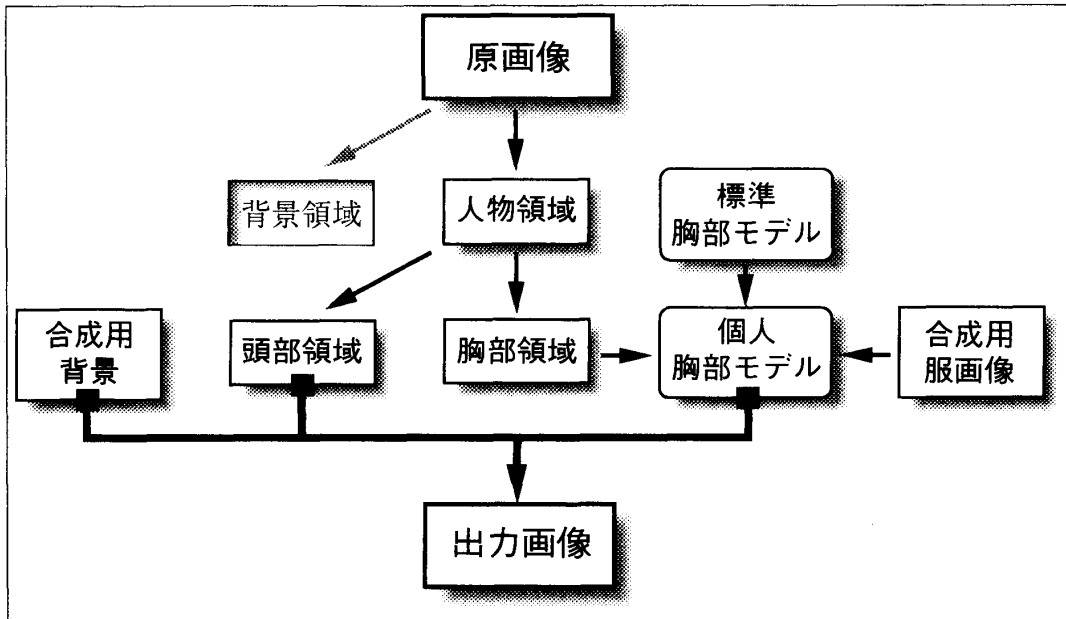
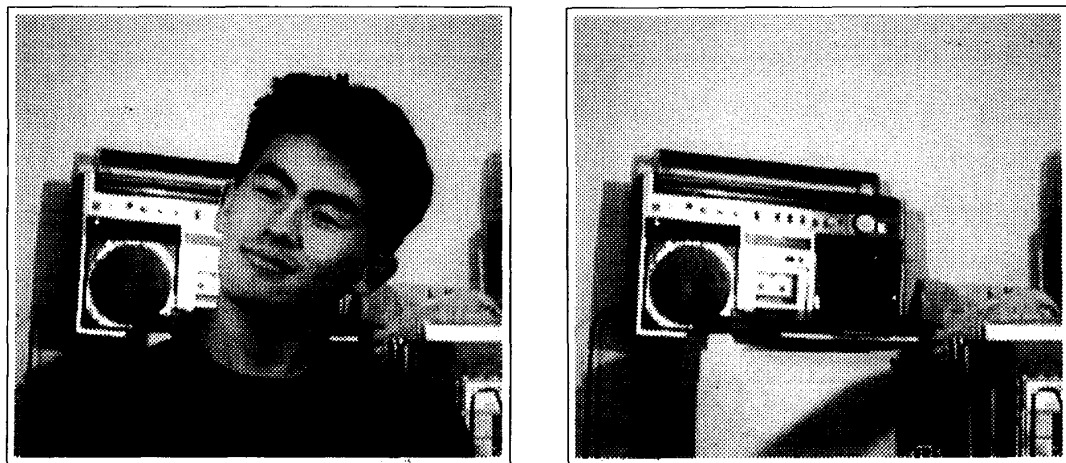


図1 仮想衣服着用システムの処理の流れ



(a) 原画像

(b) 背景画像

図2 原画像と背景画像

の各フレームに対して適用する。

2.2 画像の撮影

カラー TV カメラを固定し、人物と背景を含む画像、背景のみの画像をそれぞれ撮影する。以後、前者を原画像、後者を背景画像と呼ぶ。原画像は、人物の頭頂から胸の上部まで完全に画像内に収まり、かつ、ほぼ人物の正面から撮影した画像とする。

撮影された両者の画像を図2に示す。

2.3 人物領域の抽出

2.3.1 背景領域の除去

まず、撮影された原画像と背景画像より、対応する画素ごとに濃度の差分値を計算し、濃度の差分値のヒストグラムを作成する。理想的な撮影環境では、原画像中の背景領域は背景画像中の対応する部分と等しい濃度値を持つため、ヒストグラム上で0の位置にピークが存在すると考えられる。しかし、実際には照明条件の変動、TVカメラ側の自動濃度補正機構などの影響により、ピークは原点上には存在せず、原点から若干ずれた位置に出現する⁴⁾。

一方、人物領域に対応する領域では、原画像と背景画像は一般には異なる濃度値を持つため、濃度の差分値のヒストグラム上では原点から大きく外れた位置にピークが出現する。ゆえに、背景領域は以下の手順で除去できる。

- (1) ヒストグラムのピーク位置を検出する。
- (2) ピーク位置からある程度の範囲の領域は、原画像と背景画像がほぼ一致した領域なので、背景画像とみなす。原画像中から、この領域に対応する画素をすべて除去する。

安定して背景画像を除去するためには、ピークからどの範囲までを背景とみなすかを適応的に求める必要があるが、現時点ではピーク値から -10 、 $+30$ の範囲の領域を背景領域とみなしている。原画像には、背景画像に存在しなかった影が生じることがあるため、しきい値はピーク値を中心として非対称に設定した。

上述の処理により原画像から背景領域を除去した画像を図3に示す。

2.3.2 雑音除去

前述の処理で得られた人物領域には、人物領域以外の雑音はかなり残留しているため、雑音除去処理が必要である。ここでは、原画像を、人物領域（画素値1）と背景領域（画

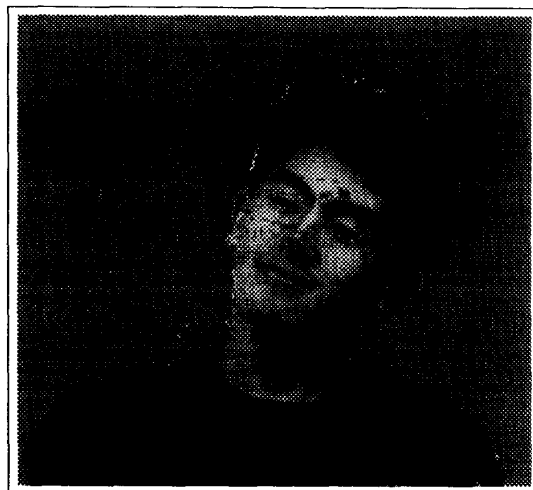


図3 背景領域除去後の画像

素値 0) からなる 2 値画像とみなし、以下の手順で処理を行う。

- (1) 大きさ 3×3 の論理マスクをラスタ走査することにより、1 画素の大きさの孤立点を除去する。
- (2) 画面内に存在する全ての閉領域（画素値 1 からなる領域）の面積を計測し、一定面積以下の領域を除去する。
- (3) 膨張処理を 3 回、縮小処理を 3 回行うことにより、人物領域に存在する穴や欠けを除去する。

2.3.3 輪郭追跡

次に、抽出された人物領域の輪郭線を抽出する。処理手順を以下に示す。ここで対象としている画像は、上記の雑音除去処理を経て、人物領域として 1 つの閉領域のみが存在する画像である。

- (1) 画像の中央最上端から下方へ向けて、人物領域を探索する。
- (2) 人物領域を発見した座標から、時計回りに人物領域の輪郭を追跡する。追跡を開始した座標に再び戻ってきた時点で、追跡を終了する。
- (3) 人物領域の最も左下に対応する場所が輪郭の開始点となるように、輪郭線のデータを再編成する。輪郭線の開始点を明確に定義することは、以降の処理の簡略化にとって有用である。

次に、全ての場所において等しく $1/7$ の重みを持つ長さ 7 の平滑化マスクを輪郭線に重畳することにより、輪郭線の平滑化を行う。

以上の処理を用いて人物領域を輪郭線のデータで表現することにより、以降の処理で必要となる曲率計算などの処理がより容易になる。

2.4 胸部領域の抽出

原画像中の人物に仮想衣服を着用させるためには、得られた人物領域を頭部と胸部の領域に分割する必要がある。頭部と胸部の明確な境界線を一義に定めることは容易ではないが、ここでは曲率に注目し、以下の方法により境界線を定義する。

2.4.1 頭部と胸部の分離点の決定

まず、人物領域の頭部と胸部の境界線は、人物領域のうち最もくびれた箇所の 2 点を両端に持つ線分であると定義する。

人物領域の左下端を始点、右下端を終点とする、時計回りの曲線 $Q(s)(s = 1, 2, \dots, L)$ に着目する。

一般に、パラメトリックな曲線 $Q(t)$ 上の点 t_0 における曲率 $r(t_0)$ は次式で表される。ただし、プライム記号 ($'$) は、パラメータ t についての微分を表す。

$$r(t_0) = |\mathbf{Q}''(t_0)| = \left| \lim_{\Delta t \rightarrow 0} \frac{\mathbf{Q}'(t_0 + \Delta t) - \mathbf{Q}'(t_0)}{\Delta t} \right| \quad (1)$$

式(1)を曲線 $\mathbf{Q}(s)$ に直接適用した場合、首のくびれの部分特有の特徴が曲率に反映されると考えられる。しかし、首の部分のような水平方向のくびれを検出するためには、曲線の垂直方向の移動は考慮する必要がないために、曲線の水平方向 (x 方向) の移動のみに着目する。最終的に、曲率 $\tilde{r}(t_0)$ は次式により求められる。ここで、 $Q_x(t)$ は点 $\mathbf{Q}(t)$ の x 成分を示す。また、 d は実験的に定める定数である。

$$\tilde{r}(t_0) = Q_x(t_0 + d) - 2Q_x(t_0) + Q_x(t_0 - d) \quad (2)$$

横軸に $t(t = 1, \dots, L)$ 、縦軸に $\tilde{r}(t)$ をとった場合、このグラフには明確な2つのピークが出現する。そこで、人物領域の輪郭線上でこのピークに相当する2つの点を頭部と胸部の分離点として採用する。

2.4.2 頭部と胸部の境界線の決定

次に、上述の処理により決定された2点を接続して頭部と胸部の境界線を作成する。このとき、単純に2点間を結ぶ直線を境界線とした場合、仮想衣服と頭部の接続が不自然に見えることが予想される。そこで、スプライン曲線を用いて、分離点間を滑らかに補間する。

スプライン曲線を発生するためには、3点以上の節点が必要である。両端の2点は決定されているので、この2点の位置関係から新たな節点をもう1点設け、これら3点を補間する。スプライン曲線は、自由端を持つ3次のパラメトリックスプライン曲線を用いている。

画面下端、左肩、右肩、首の合計4本の線分で囲まれる閉領域を、最終的な胸部領域とみなす。

抽出された人物領域の輪郭について、境界線を決定した結果を図4に示す。

2.4.3 輪郭線の分割による特徴点抽出

胸部領域の輪郭形状は、後述の標準胸部モデルの変形のために用いる重要な情報である。そこで、領域全体の性質を反映し、領域が変形しても互いの位置関係がなるべく変化しない座標点列を選定するために、胸部領域の輪郭線をセグメントごとに等分割し、特徴点を設ける。

胸部領域の輪郭線を構成するセグメントについて、曲線長のパラメータを用いて、左右の肩については14、首については6、下端については10の部分にセグメントごとに等分割する。その結果44個の特徴点が得られる。得られた特徴点を、図5に示す。

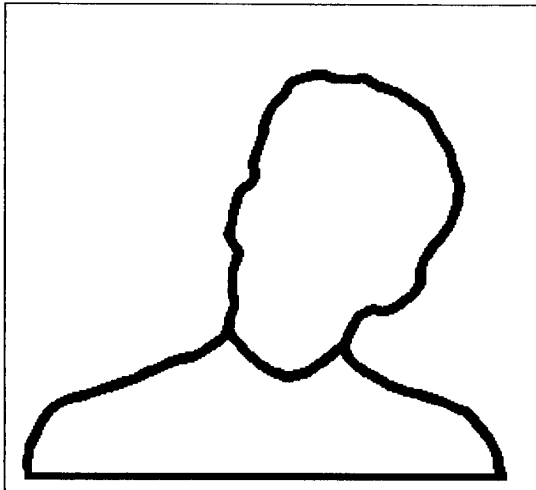


図4 人物領域の輪郭線と頭部・胸部の境界線

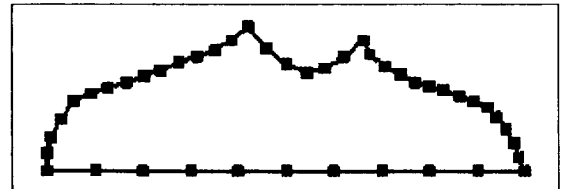


図5 胸部領域輪郭上に定められた特徴点

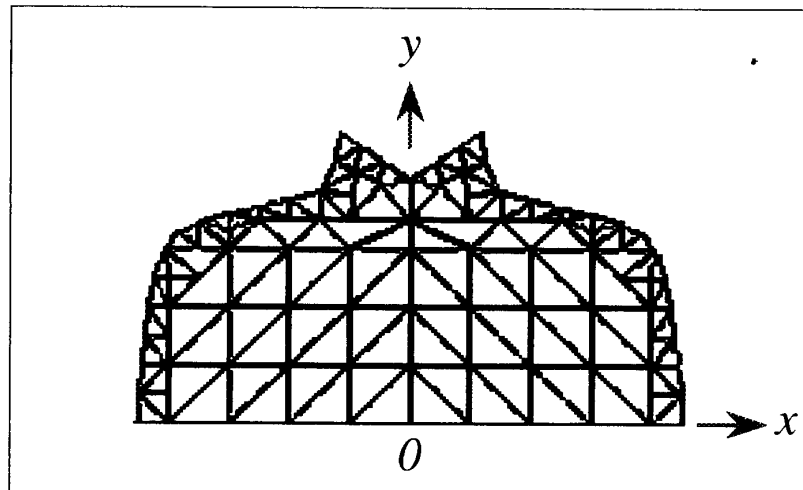


図6 標準胸部モデル

2.5 標準胸部モデルの変形

仮想衣服着用に際しては、あらかじめ用意された衣服の形状を個人の胸部領域に合わせて変形させ、両者の形状を一致させる必要がある。また、形状が一致するだけでなく、胸部領域の動きに合わせて衣服全体が柔軟に変形しなければ、不自然さが生じる。そこで、以下の手順により、モデルに基づいた処理を行う。

- (1) 多数の多角形からなる標準的な胸部のモデルを用意する。これを標準胸部モデルと呼ぶ。標準胸部モデルを図6に示す。標準胸部モデルは、合計96点の2次元座標情報と、それらを頂点とする合計132個の3角形ポリゴンを得るための連結情報から構成されている。
- (2) 前節で得られた特徴点列を用い、個人の胸部領域の形状に合わせて標準胸部モデルを変形させる。この結果得られた胸部モデルを個人胸部モデルと呼ぶ。ポリゴンの連結情

報は、標準胸部モデルと共通である。結果的に、特徴点は個人胸部モデルの座標点列の部分集合となる。

- (3) 同様に、あらかじめ用意された衣服画像の胸部領域の形状に合わせて標準胸部モデルを変形させる。得られたモデルを衣服モデルと呼ぶ。
- (4) 衣服モデルから個人胸部モデルへ、対応する多角形ごとにテクスチャマッピングを行う。

本節では、標準胸部モデルを対象とする胸部の形状に合わせて変形させるための手法について述べる。なお、以下では、標準胸部モデルと個人胸部モデルについて、特徴点では常に陽に移動ベクトルが得られるものとし、それ以外の点も含めた座標点は単に頂点と呼ぶこととする。

2.5.1 標準胸部モデルの大まかな変形

標準胸部モデルは、得られた胸部領域とは座標系が異なるうえに、位置や大きさにかなり差があるため、以下のような簡単な座標変換により大まかな変形を行う。

- (1) 標準胸部モデルの y 座標の符号を反転する。
- (2) 標準胸部モデルを囲む外接矩形枠が、全特徴点を囲む外接矩形枠に一致するように、標準胸部モデル全体のスケールリングと平行移動を行う。

2.5.2 標準胸部モデルの細部のフィッティング

一般に、変形前のモデルと変形後のモデルについて、対応する座標点の移動量（移動ベクトルと呼ぶ）の全てが陽に得られることは少ない。そのため、少数の移動ベクトルのみを用いてモデル全体を変形させる手法が種々提案されている。ここでは、秋本らによる手法⁵⁾を用いた。この手法では、

- A. 観測された移動ベクトルは、各特徴点に内在する、陽に得られた移動ベクトルとは異なる移動ベクトル（内部移動ベクトル）の合成の結果生じたものである。
 - B. ある頂点の移動ベクトルは、全ての重み付けされた内部移動ベクトルの和である。
- という仮定のもとで、陽に得られた移動ベクトルの位置関係に関する連立一次方程式を解くことによってモデル全体の移動ベクトルを算出する。

以下では、次の記号を用いる。

- $S = \{S_i \mid i \in 1, 2, \dots, M\}$: 標準胸部モデル頂点列
- $\hat{S} = \{\hat{S}_j \mid j \in 1, 2, \dots, N\}$: 標準胸部モデル特徴点列
- $P = \{P_i \mid i \in 1, 2, \dots, M\}$: 個人胸部モデル頂点列
- $\hat{P} = \{\hat{P}_j \mid j \in 1, 2, \dots, N\}$: 個人胸部モデル特徴点列
- $V = \{V_i \mid i \in 1, 2, \dots, M\}$: 頂点の移動ベクトル
- $\hat{V} = \{\hat{V}_j \mid j \in 1, 2, \dots, N\}$: 特徴点の移動ベクトル

$\tilde{V} = \{\tilde{V}_j | j \in 1, 2, \dots, N\}$: 特徴点の内部移動ベクトル

実際には, $M = 96, N = 44$ である。

この段階で既知のものは, S, \hat{S}, \hat{P} のみである。未知のものについては, 前述の仮定より次の関係が得られる。

$$P = S + V \quad (3)$$

$$\left. \begin{aligned} V_1 &= \sum_{j=1}^N \tilde{V}_j \cdot f(S_1, \hat{S}_j) \\ V_2 &= \sum_{j=1}^N \tilde{V}_j \cdot f(S_2, \hat{S}_j) \\ &\vdots \\ V_M &= \sum_{j=1}^N \tilde{V}_j \cdot f(S_M, \hat{S}_j) \end{aligned} \right\} \quad (4)$$

$$\left. \begin{aligned} \hat{V}_1 &= \sum_{j=1}^N \tilde{V}_j \cdot f(\hat{S}_1, \hat{S}_j) \\ \hat{V}_2 &= \sum_{j=1}^N \tilde{V}_j \cdot f(\hat{S}_2, \hat{S}_j) \\ &\vdots \\ \hat{V}_N &= \sum_{j=1}^N \tilde{V}_j \cdot f(\hat{S}_N, \hat{S}_j) \end{aligned} \right\} \quad (5)$$

$$\hat{V} = \hat{P} - \hat{S} \quad (6)$$

$$\left[f(X, Y) = \frac{C}{|X - Y| + C} \quad (C : \text{Const}, 0 < C) \right]$$

上式より, 式(5)の連立一次方程式を解けば \tilde{V} が得られ, 結果的に個人胸部モデル P が得られる。なお, 用いているベクトルは全て2次元ベクトルのため, 連立方程式を2回解く必要がある。定数 C は, 移動ベクトルから受ける影響の度合いを調整するためのもので

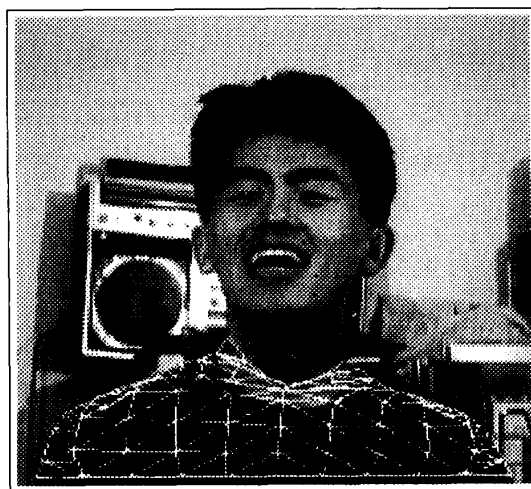


図7 フィッティング後の胸部モデル

あり、実験的に定める。

上述の方法によって標準胸部モデルを変形し、人物領域に合わせてフィッティングを行った結果を図7に示す。

2.6 テクスチャマッピング

まず、テクスチャマッピングに先立って、人物に着用させたい衣服の画像を準備する必要がある。この処理は、任意の人物に実際に任意の衣服を着用させ、これまで述べた手順をそのまま適用する。その結果、衣服領域の画像と衣服モデルが得られる。

次に、個人胸部モデルの各3角形ごとに、スキャンコンバージョン⁶⁾により3角形ポリゴン内に含まれる全ての画素を順に走査し、各画素の濃度値を衣服モデルの対応する3角形内から得られた濃度値に入れ換える。

衣服モデル内のある3角形の各頂点座標値を (p_x, p_y) , (q_x, q_y) , (r_x, r_y) 、個人胸部モデルでの対応する3角形を (p'_x, p'_y) , (q'_x, q'_y) , (r'_x, r'_y) とし、前者の3角形内の1点の座標値を (x, y) 、後者の3角形内の1点の座標値を (x', y') とすると、2点間の変換式は以下のようなアフィン変換式で表される。

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} \quad (7)$$

但し、

$$\begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} p_x & q_x & r_x \\ p_y & q_y & r_y \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} p'_x & q'_x & r'_x \\ p'_y & q'_y & r'_y \\ 1 & 1 & 1 \end{pmatrix}^{-1} \quad (8)$$

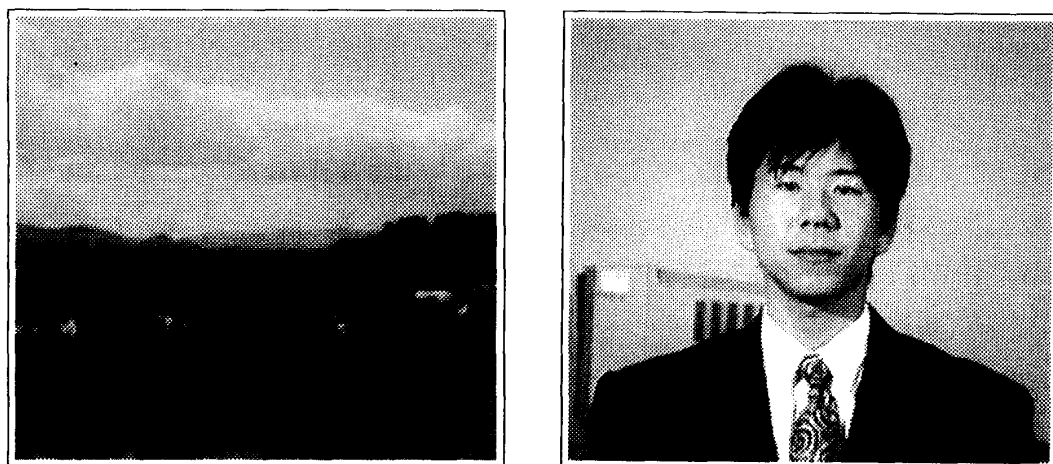
胸部モデル以外の領域については、2.4での処理により人物領域がすでに抽出されているので、人物領域以外の領域が背景であり、人物領域のうち胸部モデル以外の部分が頭部であることが自明である。そのため、頭部領域、背景領域はそれぞれ単純な処理で抽出できる。このとき、背景領域は必要に応じて任意の画像で代用することが可能である。

最後に、テクスチャマッピングされた胸部モデル、頭部領域、背景領域の3者を合成することによって、最終的な出力画像が得られる。

3. 実験結果と検討

3.1 実験結果

一連の処理による仮想衣服着用実験結果を以下に示す。図8は前もって用意された背景画像と服画像である。図9は、実際の人間の動作を撮影した一連の原画像であり、各画像に対する処理結果が図10である。



(a) 背景画像

(b) 服画像

図8 背景画像と服画像



図9 仮想衣服着用前の原画像

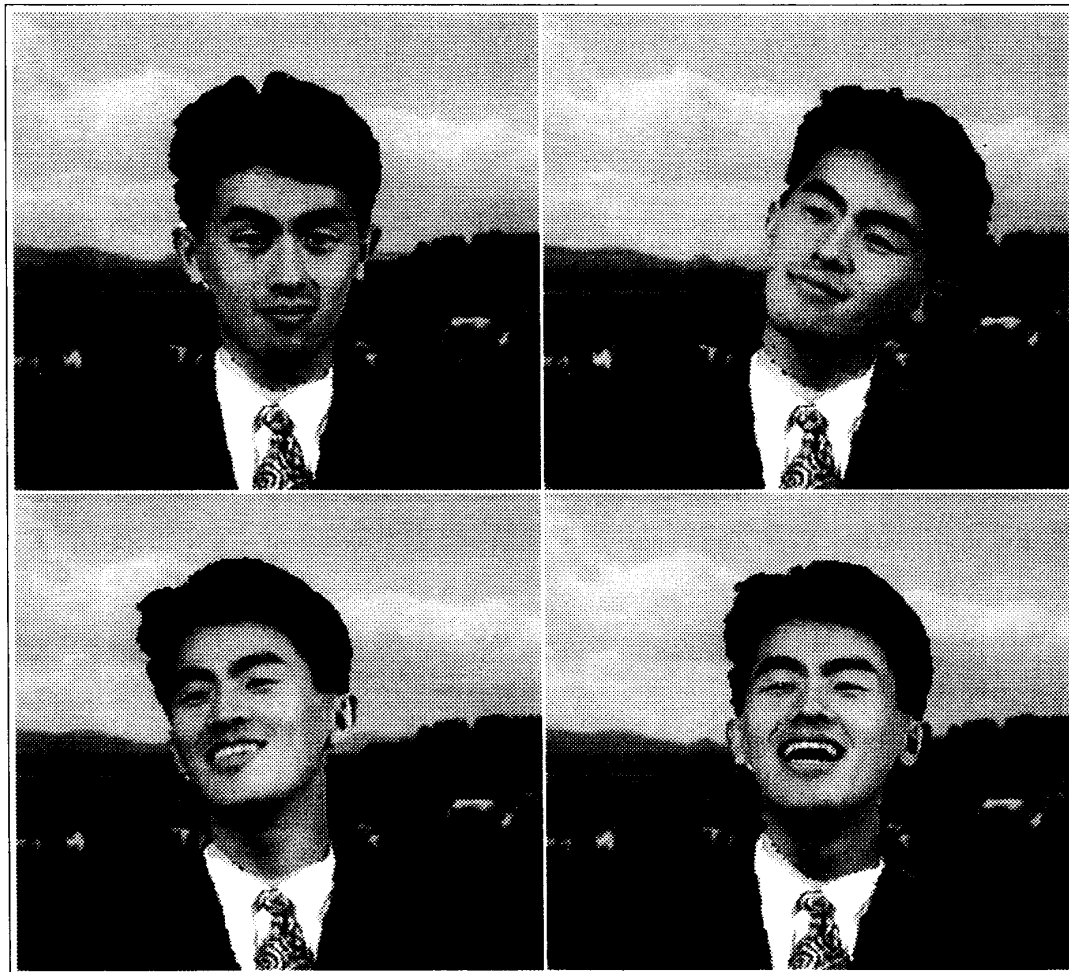


図10 仮想衣服着用実験結果

図10では、服画像がよく動きに追従しており、原画像とはかなり受信者に与える印象が異なるにもかかわらず、全体的に大きな違和感は生じていない。

3.2 人物領域の抽出について

今回用いている人物領域抽出法の最大の利点は、特別な背景を必要としないことである。テレビ放送などでは、青色のクロマキーバックを背景とし、人間の肌色と背景の青色がほぼ補色関係であることを利用して人物領域を抽出する手法が用いられるが、一般家庭ではこのような背景のためのスペース確保が困難であるために、望ましい手法とはいえない。

しかし、現在用いている手法においても、

- 背景領域と人物領域の間の濃度差が小さい
- 背景領域に動物体（例えば、放映中のテレビ画面など）が存在する

などの場合には正確な人物領域を得ることが困難になるため、ある程度の背景の制限はやむを得ないと考える。また、ヒストグラム上でのしきい値は環境によりかなり変動するた

め、適応的に決定するための対策も必要である。

また、頭部の領域は、その領域の画像に画像圧縮などの処理を施し、画像として伝送する必要がある。しかし、頭髪や顔の領域は一般に

- 濃度の変化があまり大きくない
- 色相をある程度の範囲に制限できる

などの性質を持つため、画像圧縮の際に、服や背景の領域に比べて安定した高圧縮率が期待できる。

3.3 胸部モデルの作成について

実験では、胸部モデルに基づいた処理を行うことにより、服画像が身体の動きに柔軟に追従することが示された。しかし現時点では、胸部モデルは人物領域の輪郭線の情報に完全に依存しているために、例えば画面中の人物が受話器を耳に当てる動作を行い、腕が画面内に現れた場合などには、胸部モデルにかなり大きな歪みが生ずることが予想される。また、人物領域が縦方向に平行移動した場合などには、不自然な服の伸縮が生じるなど、いくつかの問題が存在するため、胸部モデル作成に際してはさらに何らかの知識を導入する必要がある。

実際に動画像を処理する場合には、専用ハードウェアの搭載などによる高速処理が必要となる。本実験で用いた手順のうち、胸部モデルのフィッティングは、44元の連立方程式を2回解く必要があるために、最も処理時間を要し、かつハードウェア化が困難な部分であると思われる。しかし、通常の動画像では直前のフレームからの動きは大きくないため、前のフレームとの相関を利用して計算量を削減することも可能であると考えられる。

4. ま と め

本論文では、テレビ電話の使用時における服装などにともなう心理的抵抗感の軽減を主な目的とした、仮想衣服着用システムについて述べた。実験の結果、ある程度通話環境を制限する必要があるものの、仮想衣服が人物の動きに柔軟に追従するという良好な結果が得られた。さらに本手法では、初回のみ合成用背景画像と服画像を送信し、次の動画フレームからは頭部領域の画像と胸部モデルの特徴点のみを送信すれば受信側で画像を合成できるため、通信路の負担をかなり軽減できる。

実際に完成されたシステムを構築するためには、

- 人物領域のロバストな抽出
- 専用ハードウェアによる高速処理
- 動画像指向の処理

などの項目について改善の必要があるが、本論文で述べた手法は、心理的抵抗感の軽減にとどまらず、画像圧縮技術と知的符号化技術のそれぞれの欠点を補うためのハイブリッド

な符号化の一方法として、今後大いに発展の余地があるものとする。

最後になったが、本実験を行うにあたって御助力を頂いた本研究室の平成3年度卒業生金田真也君（現在、(株)中国日本電気ソフトウェア）ならびに平成4年度卒業生尾崎耕之君（現在、(株)岡山日本電気ソフトウェア）に厚く感謝する。

参考文献

- 1) 相澤清晴, 原島 博, 齊藤隆弘: 構造モデルを用いた画像の分析合成符号化方式, 信学論 (B-I), **J71-B-1**, 3, pp. 200-207 (1989)
- 2) 金田真也, 島田英之, 塩野 充: テレビ電話のための仮想衣服着用システムに関する研究, 平成3年度岡山理科大学工学部電子工学科卒業論文 (1992)
- 3) 尾崎耕之, 島田英之, 塩野 充: テレビ電話のための仮想衣服着用システムに関する研究 (II), 平成4年度岡山理科大学工学部電子工学科卒業論文 (1993)
- 4) 藤本 強, 正満峰夫, 長谷雅彦: 背景メモリを用いた背景除去法式の検討, 信学技報, **PRU90-146**, pp. 9-16 (1990)
- 5) 秋本高明, 末永康仁: 頭部の正面・側面像と基本3次元形状を用いた顔画像生成手法, 信学技報, **PRU88-47**, pp. 61-66 (1988)
- 6) 山口富士夫: 実践コンピュータグラフィックス, 日刊工業新聞社 (1988)

Virtual Clothes Wearing System for Videotelephones

Hideyuki SHIMADA and Mitsuru SHIONO

Department of Information Engineering

Faculty of Engineering

Okayama University of Science

Ridai-cho 1-1, Okayama 700, Japan

(Received September 30, 1993)

In the coming 21st century, the B-ISDN (Broad band Integrated Services Digital Network) system will become very popular in our daily life. The B-ISDN system can carry so great quantity of information that videotelephone system will become a practical equipment in our life. But a videotelephone has a basic problem, that is, not only user's face but also private scene which the user would not like to show is send to the receiver of the call. For example, an employee will be troubled if his boss call up him when he has just taken a bath. In this paper, we propose a virtual clothes wearing system for videotelephones, and show the results of fundamental experiments. Using this method, an user can send his own image above the chest to the receiver with virtual colthes that he likes best. For example, the receiver can see him wearing his good suit even if he is indeed naked then.