

## 顔動画のインタラクティブ知的符号化における ローカル動き検出と眼鏡画像処理

大 関 和 夫<sup>†1</sup> 齊 藤 隆 弘<sup>†2</sup>  
金 子 正 秀<sup>†3</sup> 原 島 博<sup>†4</sup>

画像信号の統計的性質を利用した従来の波形符号化とは異なる符号化方式として知的符号化がある。筆者らはこの知的符号化における符号化部分、すなわち入力顔動画の分析部分にインタラクティブ操作を導入することにより実用化を図る方式について研究を行ってきた。本論文では、インタラクティブ操作の曖昧性について考察を行い、表情などのローカル動き検出と眼鏡処理に関し、その手順の明確化、操作量の評価と改良を行った。ローカル動きとグローバル動きの記述順序の違いをインタラクティブ操作量の比較により行い、顔部品に応じて異なる記述順序を用いることが有利なことを明確化した。次に、ワイヤフレームモデル上の1頂点に対する移動の操作が周辺の頂点に階層的に波及する移動規則を検討し、初期形状とその変形パターンとして0次、0.5次、1次の周期波形を用いた連動規則を開発した。これにより形状合わせの操作量の削減を符号量を増加させることなく実現した。開発したローカル動き検出ツールを拡張し、眼鏡画像処理を含んだ符号化を行った。新たな問題点として、眼鏡除去処理の操作量が定量化しにくいことを指摘した。眼鏡フレームを肌色で書き加える操作をどの程度まで行うかの定量的な判断基準として、テレビジョン受信におけるゴースト障害の消え残りレベルとの対応から求まる、DU比を導入した。

### Detection of Local Motion Parameters and Processing of Eye-glasses in Interactive Model-based Coding of Facial Image Sequence

KAZUO OHZEKI,<sup>†1</sup> TAKAHIRO SAITO,<sup>†2</sup> MASAHIDE KANEKO<sup>†3</sup>  
and HIROSHI HARASHIMA<sup>†4</sup>

For interactive model-based coding of facial image sequence, the number of interactive operations and difficulty of procedures should be clarified. This paper develops a local motion detection tool which realizes encoding in a limited number of operations with practically easy operations. A comparison of description order between local motion and global motion is examined from the viewpoint of amount of operations. Hierarchical and cooperative moving rule of vertices on wireframe model around facial part is proposed. By introducing zero-th and the 1-st order periodic waveform pattern for both horizontal edges of mouth, the number of operations for local motion is reduced. Extending the developed tool to the processing of eye-glass frame, new problems are found to be analyzed. Eye-glasses which overwrites glass frame by skin color is compared with a method using separated input of eye-glasses. A new deciding criterion for terminating human operation of eye-glasses processing is developed. The total encoding tool is fully practical and many facial image sequences are well encoded by the authors and others engineers.

#### 1. ま え が き

画像を含むマルチメディア通信においては、H.261、MPEGなどの波形符号化方式が実用化されているが、より低ビットレートの符号化や画像の構造化に対する検討が求められている。顔画像の知的符号化では、送・受信側双方で人物頭部の3次元形状モデル(ワイヤフレームモデル)を共有し、送信側では入力画像から形状変化や動きの情報を抽出する。受信側では、これら

†1 株式会社東芝小向工場放送映像機器開発部  
Broadcast Sys. & Equip. Develop. Dept., Komukai  
Works, Toshiba Corporation  
†2 神奈川大学工学部  
Faculty of Engineering, Kanagawa University  
†3 電気通信大学大学院電気通信学研究科  
Graduate School of Electro-Communications, The Uni-  
versity of Electro-Communications  
†4 東京大学工学部  
School of Engineering, The University of Tokyo

の情報に基づいて3次元形状モデルを変形し、さらに、あらかじめ1枚の顔写真として用意された顔部分のテクスチャ情報をこのモデルにマッピングすることによって、リアルな顔動画像を再生する。知的符号化を実現するうえで、現状では、送信側で入力画像から形状変化や動きの情報を安定して自動的に抽出することの難しさが問題となっている。これに対し、筆者らは、インターネットなどの非実時間通信環境の下での利用を想定し、符号化プロセスにおいて人間によるインタラクティブ操作を積極的に導入し、復号画像の画質の向上、処理の安定化を図る方法として、インタラクティブ知的符号化を提案した<sup>1),2)</sup>。

このインタラクティブ知的符号化は、知的符号化が有していた特徴を具体的に実用化していく作用を持つ。すなわち、従来のMPEGなどの波形符号化では達成できないような、超低ビットレートでの圧縮を可能とすることが第一の意義である。さらに、単に情報の圧縮を行うだけでなく、符号化の過程でなされる画像の構造化記述、動きパラメータ類が受信側で活用できる点が第二の意義である。特にインタラクティブ操作を介して得られる構造化情報や動きのパラメータは、抽象的な信号処理データとは異なり、顔へのワイヤフレームモデル配置、特徴点、3次元動きパラメータ、表情などに関する動きパラメータなどヒューマンインタフェース上有用な情報である。これらを用いると、受信側で、たとえば他人の顔のテクスチャとの入替えや動きないしは表情の入替えなど符号化されたデータの段階での編集、加工などの新たな処理の実現が可能となってくる。

そこでは、どのようなインタラクティブ操作を導入すれば、符号化に必要な動きパラメータなどを確実に検出できるかについての検討を行った。カメラのおよその配置と初期画像にフィッティングした3次元ワイヤフレームモデルの奥行き情報を仮定することにより、顔画像上に設定した12個の特徴点の2次元座標情報を用いて、奥行きを含む回転と平行移動からなる3次元動きパラメータを検出するアルゴリズムを導出した<sup>3)</sup>。このアルゴリズムを基に送信側における符号化ツールを構成し、また受信側における合成プログラムと合わせインタラクティブ知的符号化システムとしてまとめあげた<sup>1)</sup>。

ところで、インタラクティブ操作に関してはインタラクティブな画像認識ツールの開発が過去にも試みられたことがある。一例として、Cuginiら<sup>4)</sup>は画像理解のための手続きの定義を行うための研究用ツールを開発し、その後CADシステムなどのインタフェース知

識獲得の研究に発展させている。これに対し、筆者らの提案してきたインタラクティブ知的符号化は単なる補助や知識獲得とは異なり、符号化に必要なパラメータの解析を促進するような方式を採求することを特徴としている。ここで導入したインタラクティブ操作とは、計算機による全自動での認識ができない部分を人間が代行するというのではなく、どのような操作をすると、どのようなパラメータ検出ができるのかという関係を明確化し、より少ない操作量でより精度の良い検出を行うことを目標としている。

本論文では、顔動画像のインタラクティブ知的符号化において、ローカル動きパラメータ検出と眼鏡画像処理についてインタラクティブ操作を明確化する観点から検討する。まず、頭部全体の動きに関するグローバル動きパラメータ検出後、目、口などの表情に関する動きをローカル動きとして検出している。目に関しては瞬きにおける目の開閉度をマウスとGUIを用いて入力していく。口に関してはフィッティングされたワイヤフレームモデル上にある頂点をマウスによって移動していく。このようなインタラクティブ操作において、1つの頂点に連動して唇の初期形状を保存したまま、関連ある頂点集合を移動させることによる操作量の削減方式に加え、唇の形状として平坦な直線とうねりのある曲線をモデルに加える拡張を行っている。これらは周期波形のパターンでモデル化できるものであり、インタラクティブ操作の過程で人間の指示に対応して切り替える方式を構成できる。

次に、眼鏡のある顔動画像処理についての検討を行った。知的符号化においては従来、眼鏡のある画像は対象外とされていた。すなわち知的符号化では眼鏡を外した顔画像を初期画像として入力することによって、問題の回避を図っていたともいえる。しかしながら、たとえば講演者の顔動画像記録し、符号化することが考えられるが、符号化を行う都合上わざわざ眼鏡を外して初期画像を入力したり、講演中も眼鏡を外したり、眼鏡のみを借り受けて別途その画像を入力することなどを、いつも行えるとは限らない。このような想定では眼鏡をかけたままの顔動画像を処理する必要がある。このような応用のためにインタラクティブ知的符号化のツール開発を行い、画像から眼鏡除去の処理を行うときの操作が曖昧になることが問題であることを述べる。この問題に対応するために、眼鏡消去操作後の残留誤差に定量的評価を定めることを試みる。

以下、2章では、インタラクティブ知的符号化の概要並びにその操作回数に関する諸問題を検討する。3章では、ローカル動きパラメータ検出ツールにおける

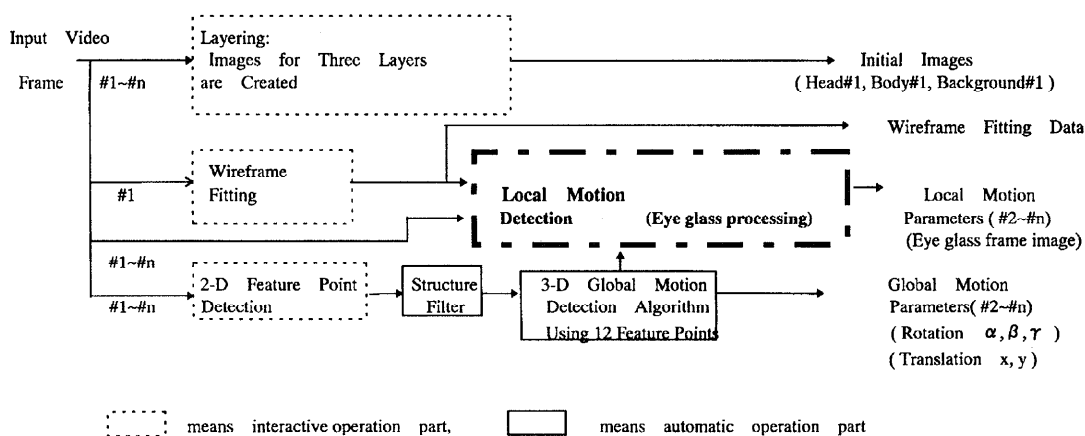


図1 インタラクティブ知的符号化システムの符号化部

Fig. 1 Encoder of interactive model-based coding.

ワイヤフレームモデルの頂点間の関係を調べ、インタラクティブ操作の負担軽減の観点から顔部品にパターンとしての運動規則を取り入れた検討について述べる。4章では構成されたローカル動きパラメータ検出ツールを用いて眼鏡のある顔画像のインタラクティブ知的符号化を行ううえでの操作量の問題を検討し、眼鏡消去処理の作業の程度に対する定量化を与える。

## 2. インタラクティブ知的符号化と問題点

図1にインタラクティブ知的符号化システムの送信部のブロック構成を示す。顔動画像系列に対し、初期画像1枚と動きなどのパラメータにより表現している。初期画像に対しては頭部の3次元ワイヤフレームモデルをフィッティングさせる。このワイヤフレームモデルに対して回転、平行移動などのグローバルな動きと表情などのローカルな動きによる変形を加え、初期画像のテクスチャマッピングを行い、各フレームの画像合成を行う。符号化されるパラメータは初期画像、ワイヤフレームモデルのフィッティング情報、グローバル動きパラメータ、ローカル動きパラメータである。ローカル動きパラメータのうち、目の形状は瞬きの度合いという一種の意味的なコマンドで表現されている。太い点線で囲まれた部分が本論文で検討している処理である。

インタラクティブ知的符号化の符号化部のうち、初期画像に対する頭部の3次元ワイヤフレームモデルのフィッティング、特徴点入力、ノイズ除去フィルタ、グローバル動きパラメータ検出アルゴリズムについては文献1), 3), 5)に述べられている。初期画像に対するワイヤフレームモデルのフィッティングは80の点についてなされるが、顔の輪郭などの約20点は省略で

きることが多い。グローバル動き検出は顔画像のワイヤフレームモデルを剛体と仮定し、頭部に対する回転と平行移動パラメータを求めている。頭部が剛体に近い場合は初期画像のワイヤフレームモデルを正確なグローバル動きパラメータにより回転・平行移動させれば移動後の顔部品にフィッティングしたワイヤフレームが得られる。しかし実際には発話や表情による変形、また顔自体が非剛体であることによる変形があり、これらを合わせてローカル動きとして再度変形することが必要になる。グローバル動きが線形な動きであるのに対し、ローカル動きパラメータは発話や表情などの非線形な動きも含んでおり、特徴点によってその動きを表現する場合、点数は1つの顔部品に対し10~数十個になってくる。これに対し顔の表情をアクションユニット(以下、AUと略す)と呼ばれるパラメータに分解し、表現する手法が研究されてきた<sup>6), 7)</sup>。Ekmanら<sup>8)</sup>によれば44種のAUにより顔の表情が表現できる。顔画像の合成においてはAUは有効だが、実画像からのAUの検出は正規化や大きい回転などに対する検出に問題が残っている<sup>9)</sup>。

従来ワイヤフレームモデルを用いた顔動画像の合成においては、初期ワイヤフレームモデルにAUを適用した後、グローバル動きを適用する手法がとられていた。これに対し、筆者らは、グローバル動き検出後に表情変形やその他非線形歪を合わせてローカル動きとして記述することにより、入力実画像に忠実な符号化方式を実現してきた。本論文では、その差異を操作量の観点から評価していく。

次に、知的符号化では従来より分析合成の都合上、眼鏡や髭のある画像の処理は対象外とされていた。これはテレビ電話など実時間応用を前提としていたため、

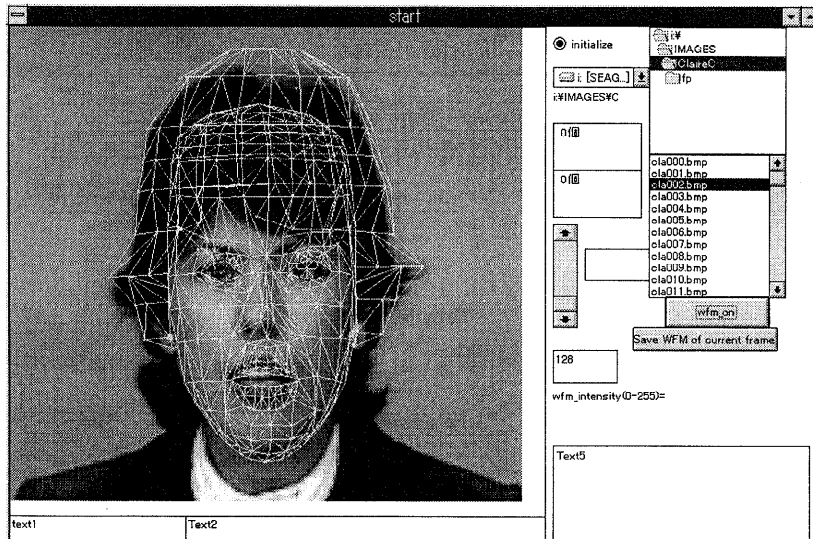


図 2 ローカル動きパラメータ検出ツールの動作画面

Fig. 2 Operational display of local motion parameter detection tool.

顔画像の特徴点の検出、追跡、動きパラメータ検出などにはまだ困難な部分があるため、その処理の負荷を増加させる眼鏡画像処理の検討がなされなかったとも考えられる。本論文のインタラクティブ知的符号化では非実時間応用であるため、インタラクティブ操作を介して眼鏡のある画像についても動き検出が可能となる。眼鏡画像処理においては眼鏡部分を肌色で上書きする処理と、移動後の眼鏡画像を作り上書きする処理が必要になる。眼鏡画像処理を含む符号化を行ってることにより、新たな問題点を提起する。肌色での上書き処理は入念に行えば行うほど、品質が向上することになるが、眼鏡フレームの境界部などの完全な消去は操作量の増大を招く。実用的なレベルとして、半自動的に、たとえば数分以内に行えることが望ましい。そこで肌色の上書きにおいてエッジ部分の拡散フィルタの適用や、操作量の上限を設定するため、消え残りに対する定量的評価を与えることなどが必要となる。

### 3. ローカル動きパラメータ検出ツールと操作回数

#### 3.1 ローカル動きパラメータ記述の操作量

図 2 は符号化ツールのうち、ローカル動きパラメータ検出に関する部分の動作画面例である。符号化する顔動画像ファイル (cla??? .bmp など) をリストから選択する。第 1 フレームに対するワイヤフレームモデルのフィッティングとグローバル動きパラメータの検出は終了しているものとする。wfm\_on により、第 1 フレームにフィッティングしたワイヤフレームモデル

を現在のフレームで検出したグローバル動きパラメータで変換し、顔画像上にはほぼ剛体変形のためのワイヤフレームモデルとして上書きしたうえで、表示する。符号化パラメータとして各フレームで入力した点数、累積点数が表示される。図 2 は、操作開始時点で、両方も 0 個と表示されている。顔部品付近のワイヤフレームモデルのポリゴンの頂点に関しては、発話や表情により上記グローバル動きパラメータを用いて変形されたワイヤフレームモデルと現在のフレームの顔部品との間にずれが生じる。ワイヤフレームモデルの各顔部品付近の頂点位置を移動させることにより、ワイヤフレームの局所的変形を行い、顔部品画像との位置を合わせる。

AU を用いた知的符号化においては、ローカル動き、グローバル動きのパラメータ記述は合成の順序のうえで、まず AU で指示されるローカル動き (lm) の変形を行い、次にグローバル動き (gm) を行うという順序 (以下、lm→gm と略記) で処理が行われてきた。これに対し、実画像入力の lm 記述に関しては、記述順序 lm→gm より gm→lm の方<sup>10)</sup> が操作量、計算量ともに有利であることを表 1 を用いて説明する。まず、gm は線形な変換であり、ローカル動きがなければ  $gm^{-1} * gm = I$  (単位変換) である。変換 gm は剛体仮定の基に行われているが、実際はたとえば顎の伸縮など非剛体の変形があり、ローカル動き記述でこれらを加味していくと線形の関係がくずれる。lm→gm 方式は初期画像を基準に第 n フレームのワイヤフレームモデルのローカル動きを求めることが必要であるた

表 1 ローカル動き記述順序の比較  
Table 1 Comparison of local motion description order.

Description order	lm→gm	gm→lm
To get lm	$Tx * gm^{-1}(wfm\_n)$	$gm(wfm\_0)$
Wireframe Model Fitting	$wfm\_n$ ( $n = 0, 1, \dots, N$ )	$wfm\_0$
$gm^{-1}$ Process	$gm^{-1}(wfm\_n)$ ( $n = 1, \dots, N$ )	0
Synthesis	$Tx * gm * lm(wfm\_0)$	$Tx * lm * gm(wfm\_0)$

lm: local motion, gm: global motion, Tx: Texture mapping, wfm: wireframe

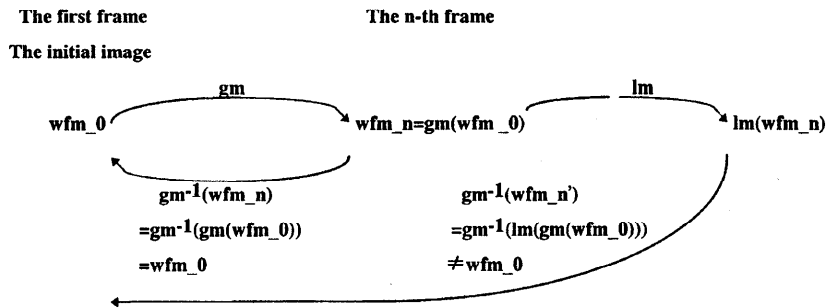


図 3 lm→gm 記述のために必要な処理  
Fig. 3 Process needed for lm→gm description.

め、第  $n$  フレームを基準とした初期画像への対応が必要となる。したがって、第  $n$  フレームのワイヤフレームモデルフィッティングが新たに必要になる。これはローカル動きを求めるためのものであるから、初期画像のワイヤフレームモデルをグローバル動きパラメータで変換したものより精度の高いフィッティングが必要となる。初期画像に対するワイヤフレームモデルフィッティングの方は、初期画像として処理するシーン全体にわたり最も正面向きで表情の強くないフレームを選択できるため、容易に行える。

図 3 に示すように、第  $n$  フレームの  $wfm\_n$  と初期画像の  $wfm\_0$  との関係は、 $lm$  がなければ、 $gm^{-1}(wfm\_n) = gm^{-1}(gm(wfm\_0)) = wfm\_0$  となる。しかし非線形なローカル動きが微少でなくなると、 $wfm\_n$  は  $wfm\_n' = lm(wfm\_n)$  と変形するため、 $gm$  による逆変換は  $gm^{-1}(wfm\_n') = gm^{-1}(lm(gm(wfm\_0))) \neq wfm\_0$  となり、顔輪郭や部品などの位置が初期画像のワイヤフレームモデル  $wfm\_0$  と違ったものとなる。これは初期画像のワイヤフレームモデルやテクスチャを複数用意しなければならないことを意味し、知的符号化の効率化に逆行する。表 1 の  $lm \rightarrow gm$  の方式を実現するためにはローカル動きの線形性を確保すべく、ワイヤフレームモデルのフィッティングを行い、その  $wfm\_0$  に対するグローバル動きの逆変換  $gm^{-1}$  を実行することなどが必要である。あるいは、別の手段により AU を検出

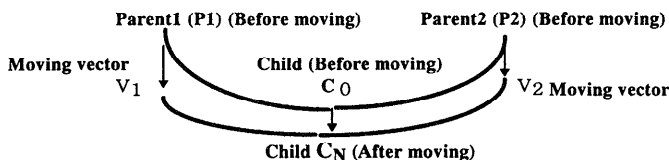
し、その AU で記述するなどの方式も考えられるが、AU 検出は自動処理でもインタラクティブ操作でも困難な処理の 1 つになっている。

表 1 から各フレームに対するフィッティング作業と  $gm^{-1}$  の演算量とが増加した処理量であることが分かる。具体的には CCITT の標準画像 Claire の場合、95～115 フレームの間は大きい回転によりフィッティングの難易度が高くなると予想される。ほかにも開口時には顎の変形によりフィッティング操作が増加すると予想される。ローカル動きの増加にともない、図 3 に示した  $gm^{-1}$  の対応のずれも増大してくる。さらに  $lm$  は検出時は実数ないしは画素単位の値（整数値）であるが、量子化によって、さらに精度が削減される。このとき  $lm$  という非線形なデータに対する  $gm^{-1}$  の非線形演算としての誤差が加わるため、 $lm$  を直接量子化した場合に比べ、精度が異なってくるという問題もある。

以上から  $gm \rightarrow lm$  の記述順序は顔部品でいうと、口の動きやその他表情・変形に適している。一方、目の動きについては、主に短い時間（3～4 フレームの間）の開閉が表現上重要であることから、顔の向きや表情にかかわらず開閉度のみを抽出すればよい。この開閉率はそのままロバスタな AU となるため、目の開閉に関しては、 $lm \rightarrow gm$  の記述が上で述べた処理量の増加なしに可能となる。もし目の開閉を  $gm \rightarrow lm$  で記述すると、検出された開閉度から AU を定めて、 $gm(AU)$

表 2 ハイブリッドローカル動きパラメータ記述方式  
Table 2 Hybred method of local motion parameter description.

Facial Parts	Eyes	Other than eyes (mouth)
lm discription	AU detection as blinking ratio	lm detection as coordinate values
Description order	lm1→gm	gm→lm2
Synthesis	Tx*gm*lm1(wfm_0)	Tx*lm2*gm(wfm_0)
Total synthesis	Tx*lm2*gm*lm1(wfm_0)	



$$C_N = C_0 + \alpha V_1 + \beta V_2$$

$$\alpha + \beta = 1, \alpha > 0, \beta > 0$$

図 4 親子の点の移動関係

Fig. 4 Relation of parents and child moving.

により開閉後のまぶた周辺のワイヤフレームモデルのポリゴンを変形させるか、各フレームでまぶた周辺のワイヤフレームモデルのポリゴンの多数の頂点を指示入力する必要がある。このいずれの場合も lm→gm 記述に比べ、処理量が増加することになり、AU 検出が容易な顔部品については lm→gm の記述方式の方が有利であることが分かる。以上を表 2 に示すように統合し、本論文では目のまばたきに関しては lm→gm 記述とし、口などの AU 検出の複雑性が高いものに対しては gm→lm 記述とするハイブリッド方式として構成した。

### 3.2 周期パターンを用いた連動処理

口形状を構成している頂点数は、本論文で採用しているワイヤフレームモデルにおいては約 50 点になっている。口の形状・表情の変化にともない 50 個の頂点をすべてインタラクティブ操作により移動させると操作回数が多く、操作者への負担が多い。発話・表情にともなう動きの多くは、ポリゴンの頂点が連動して移動している部分が多い。そこで、口形状を形成するポリゴンの頂点を階層的に定義し、上位階層の頂点が移動した場合はそれ以下の階層の頂点も連動して動くようにし、一様な動きに対しては指示操作の手間を省略できるようにする。ここでは連動に際して、復号器側の合成処理における演算量が増加しないように、内挿比率を乗算する程度の処理で階層処理方式を構築した。

口の付近のワイヤフレームモデルの連動規則を作成するために、口形状の動きのパターンをモデル化することを考える。まず、口形状のモデル化の例として、唇の輪郭上の 8 点の座標を制御点とし、残りの点をスプライン曲線で近似し、さらに丸い形状に対する特殊

処理を行う手法がある<sup>11)</sup>。これはモデル化の段階で 1 回行うものであるため、指示点数が多く、本方式には不適である。この場合の指示点数の削減は操作量の削減とともに、符号量の削減にもなる。そこで本論文では、口の両端 2 点を結ぶ線分に対し、周波数成分に対応したパターンによる表現を行う 0~1 次の周期波形をモデル化し、操作量の削減を試みる。0 次は平坦な線分、0.5 次は弓なりの半周期の円弧、1 次は 1 周期のうねりに対応するものを考える。

階層の順序に従って上位から階層ごとに移動量を求め、その階層の全頂点の処理が終了した後、次の階層の処理を行う。以下、繰り返し処理を行い、下位階層がなくなるまで処理を行う。上位階層の頂点を親とし、下位階層の所定の頂点が子として移動されるが、この親子関係を頂点ごとに後述の手順で定義しておく。親の移動にともない、まずその子がサーチされ、同時にその子に対する 2 親の移動後の位置がメモリから呼び出される。次に、図 4 に示すように、2 親の移動前の点の位置からの移動後の 2 個の移動ベクトルを内分比率に従って合成することにより子の移動量ベクトルが計算される。この移動量ベクトルが子の移動前の位置に加えられ子が移動される。この処理は巡回的になされるが親子とも移動前の位置はそのフレームの処理を開始する前の位置を用いている。この処理により上位階層の移動に起因する下位階層の形状には、そのフレームの処理を開始する前の始めの状態が継承される。したがって、口形状の場合、上下方向の開閉や平行移動の場合は、上位階層の少数点を指示するだけで他の頂点は連動するため、指示する必要がなくなる。さらに口部分に独特な形状・表情や歪みなどがある場合

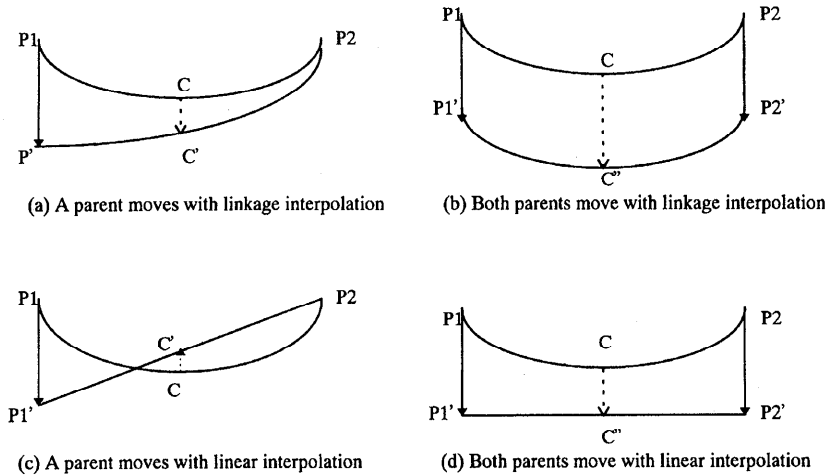


図5 運動的内挿と線形内挿の比較

Fig. 5 Comparison between linkage interpolation and linear interpolation.

は、下位階層にまで及んで指示を行うこともできる。

上のような定義に従って、運動的な動きをする場合と単純な線形内挿との比較を図5に示す。(a)は片方の親  $P_1$  の  $P_1'$  への移動にともない、子  $C$  が  $C'$  へ移動した場合で、(b)はさらにもう一方の親  $P_2$  が  $P_2'$  へ移動し、全体として子  $C''$  へ移動する様子を示している。(c)、(d)は同様の親の移動にともない、子が線形内挿される様子を示す。(a)、(b)のように親子で連携する関係は式では

$$C_{\text{new}} = C_{\text{old}} + \alpha \overrightarrow{P_1 P_1'} + \beta \overrightarrow{P_2 P_2'},$$

$$\alpha + \beta = 1, \alpha, \beta \geq 0 \quad (1)$$

となる。ここで、 $\alpha$ 、 $\beta$  は2親の影響力を与える内挿比率である。

### 3.3 運動規則の登録方式

前節の運動処理を行うためには親子関係の規則をあらかじめ登録しておかなければならない。以下、その手順を示す。まず、顔画像上にワイヤフレームモデルを配置し、次に運動させる頂点を選択する。選択された頂点に対し、その頂点に影響を与えるのに適している親となる頂点を2点選択し、登録を行う。口形状の場合は口の両端を最上位の階層とし、その中点を次の階層の点とし、さらにその中間の点を下位の階層の子として登録していく。また口形状が開口により広がるため、口から顔輪郭の間のワイヤフレームモデルの頂点に対しても、同じく親子関係の定義を行い、口周辺の画像のバランスを保つようにしている。また親から子を内挿する際の比率  $\alpha$ 、 $\beta$  は  $x$ 、 $y$  座標ごとに2親と子の距離の比を用いた。

### 3.4 運動と操作量

以上の規則を定めておくことにより、通常の口の動きに対し、操作の指示回数は4~6位で、ほぼ原画に重ねることができた。Claireの105フレームのローカル動きパラメータの形成過程を示したのが図6である。この階層的な移動規則は口の輪郭内の頂点のみではなく、顔輪郭にまで波及し、周辺のテクスチャに皺ができなくなっている。口付近で指示回数の多かった頂点は、口の両端と上唇下唇の中点の上下端であった。図7は口が閉じた場合で、初期形状が弓型になっているので、直線状にするため指示点数は増加する。この場合、上下の唇を互いに重ねるように配置することにより、指示点数を節約して、閉口を表すこともできる。

図8はCCITT標準画像の別の系列Miss Americaの第101フレームの画像である。唇の形が発話によりうねりを持った形になっており、特に唇の内側の輪郭の形状の入力を行う必要がある。

次に、具体的に運動の動作について説明する。0次である閉口の形状のとき、唇の両端点(図9(a)の  $q_1$ 、 $q_2$ )の修正を必要に応じて行う。グローバル動き変換後、両端点にずれがないときは修正はしなくてよい。次に閉口においては唇の中点  $q_3$  (または  $q_4$ ) の位置を指定するが、これが  $q_1$ 、 $q_2$  を結ぶ線分の中点  $m_3$  との中間より  $m_3$  側に近いとき、 $q_3$  を  $q_3'$  に移動させ、 $q_1$ 、 $q_2$  を親とする運動関係を移動後の親の線形内挿

$$C_{\text{new}} = C_{\text{old}} + \alpha q_1' + \beta q_2' \quad (2)$$

で行う。これにより  $q_3'$  の1回の入力で、上唇の下側の輪郭線が閉口に合わせて平坦に並ぶことになる。下

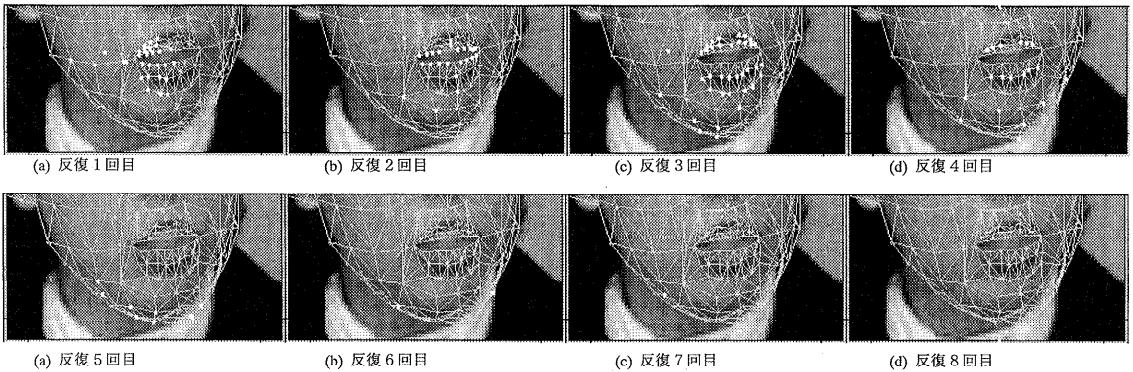


図 6 ローカル動きパラメータの座標入力と周辺の頂点の運動

Fig. 6 Local motion parameter input and its influence to surrounding points.



図 7 口を閉じた画像 (Claire 46 フレーム)

Fig. 7 Closed mouth (CCITT Claire, 46-th Frame).



図 8 口形状にうねりのある画像 (Miss America 101 フレーム)

Fig. 8 Mouth image with waving shape (CCITT Miss America, 101-th Frame).

唇も同様である。受信側での合成処理も同じ判断基準で行えるため、これによるオーバーヘッド情報は不要である。ただし、閉口に類似してさらに微妙な変形を表現しようとして、 $q_1$  または  $q_2$  と  $q_3'$  の中点がさらに入力された場合は、その点は更新され、その点を親とする子の点は以下のうねりのモードの場合を除いて式 (2) で内挿されるものとする。

次に図 9 (b) に、1 次波形のうねりのモードを示す。点  $q_5$  の指示点  $q_5'$  の位置が  $q_2, q_3'$  から式 (1) と (2) で決定される 2 つの内挿点  $q_5'(1), q_5'(2)$  の中点  $m_5$  に対し、 $q_5'(2)$  側に近い場合は通常モードとして、そうでない場合はうねりモードとして処理され

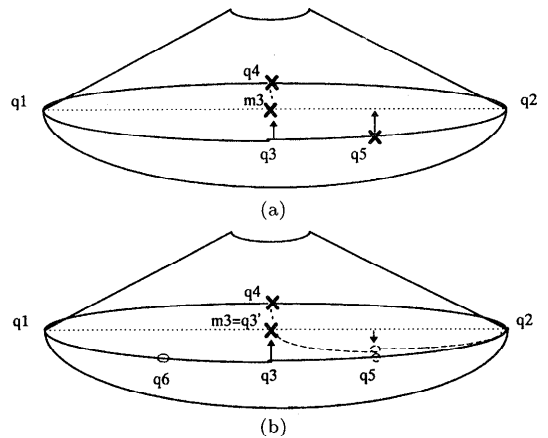


図 9 (a) 平坦モード時の頂点位置, (b) うねりモード時の頂点位置

Fig. 9 (a) Vertices arrangement for horizontal mode, (b) Vertices arrangement for wave mode.

る。 $q_3', q_5', q_2'$  で曲線を形成し、途中の子の点の内挿がなされる。親  $q_1', q_3'$  の内挿は式 (3) により  $q_6'$  の位置に設定し、 $q_1', q_6', q_3'$  で曲線が形成される。

$$p_6' = (q_1' + q_3')/2 - (q_5' - (q_3' + q_2))/2 \quad (3)$$

以上のうねりモードの処理における上記判定基準は受信側の合成過程でも同一の処理ができるため、オーバーヘッド情報は不要となる。

512×480 画素の顔画面上で、原画に重ねたワイヤフレームモデルが目視ではば一致した時点、画素にして 1 画素程度のずれに収まった時点进行操作終了の基準としている。ただし、再生画像のサイズが小さい場合は縮小比率の逆数の画素分のずれまで許容できるものとする。図 10 に 0 次、1 次のパターンを導入した場合のワイヤフレームモデルの形状を示す。図 10 (a) は図 7 と同じ Claire の 46 フレームを式 (1) の手法で処理を行った場合で、ローカル動きパラメータを 11



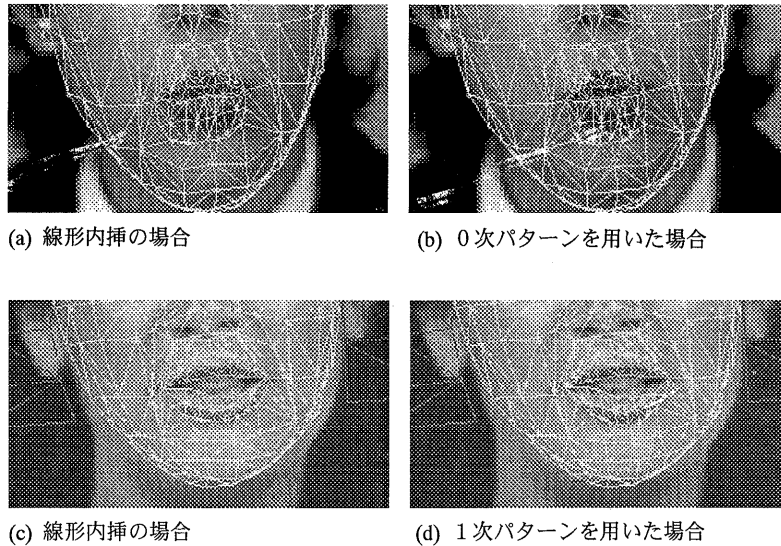


図 10 口のローカル動き (CCITT Claire 46 フレーム, Miss America 101 フレーム)  
 Fig. 10 Local motion of mouth (CCITT Claire 46th Frame, Miss America 101-th Frame).

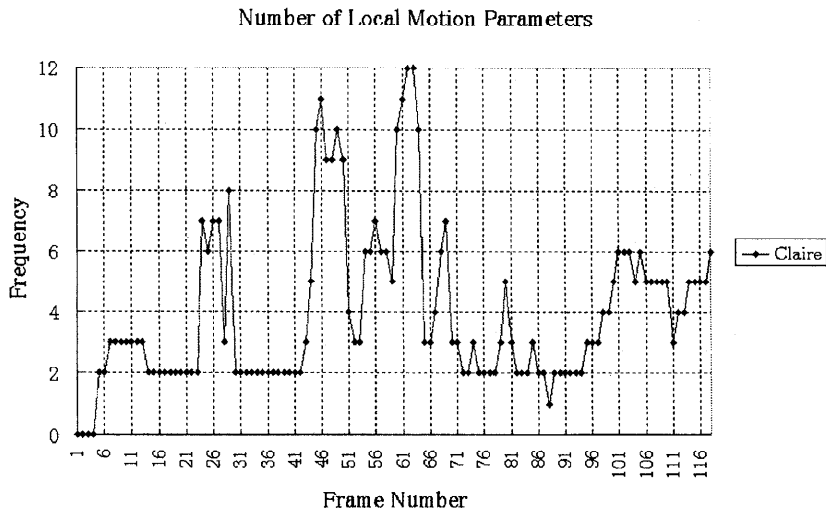


図 11 ローカル動きパラメータの符号化数 (CCITT Claire)  
 Fig. 11 Number of local motion parameters (CCITT Claire).

個使用している。(b)は式(2)の手法の場合で、ローカル動きパラメータは4個である。(c)は図8と同じMiss Americaの101フレームを式(1)の手法で処理した場合で、うねりを再現するためローカル動きパラメータを18個使用している。(d)は式(3)の手法で、ローカル動きパラメータは6個で表現されている。

図11にClaireについてのローカル動きパラメータのフレームごとの個数を示す。表3に画像Claire, Miss America, Naomiにつき5種の符号化時のローカル動きパラメータ符号数を示す。口のワイヤフレー

ムの頂点数50に比べ操作量、符号量がともに大きく削減されたのが分かる。操作量としては、平均2.6~8.2個/フレームに削減されている。出現頻度の高いフレームのうち44~48フレーム付近は唇の中心部が閉口により水平になっているためと考えられ、0次周期波形のパターン化によりこの部分は1/3にできた。

再合成画像の主観的画質評価は文献1)に述べられているが、レートに対し、従来波形符号化方式H.263よりも優れた特性を示している。また、このローカル動きパラメータを操作する図2に示すインタラクティ

表 3 ローカル動きパラメータの平均個数 (S = 再生画像サイズ 128 × 120, M = 256 × 240)

Table 3 Average number of local motion parameters (S = Reconstruction image size if 128 × 120, M = 256 × 240).

Sequence	Claire1 (S)	Claire2 (M)	Miss America (S)	Naomil (S)	Naomi2 (S)
Number of Frames	118	118	149	60	69
Average Number of Local Motion Parameters	3.983	8.237	2.624	4.183	3.348

ブなツールは筆者ら以外の技術者にも使用されている。通信・放送機構、映像分析プロジェクトにおいて 15 種以上の顔動画像系列の符号化実験がなされている。

#### 4. 操作量を考慮した眼鏡画像処理

##### 4.1 眼鏡処理を行わない場合

眼鏡画像処理は従来の実時間処理を前提とした知的符号化においては行われることはなかった。システム全体の中で特徴点の検出、追跡や 3 次元動きパラメータの検出などが困難であり、眼鏡のない人物顔画像であっても厳密には無条件での動作は難しかった。汎用性確保にあたり、ほぼ正面向きであること、大きな回転はない、背景に紛らわしい物はないなど、いくつかの前提条件を設定する必要がある<sup>12),13)</sup>。したがって、たとえ初期画像に対して、眼鏡部分の検出処理などを施すことができたとしても、それ以後実時間で入力される眼鏡付き画像から正しい動き検出を行うことは、処理の負荷の増加となるため、眼鏡付きの画像処理は行われていなかったと考えられる。これに対し、筆者らの提案してきたインタラクティブ知的符号化は、いくつかのインタラクティブ操作事項とその回数、精度などを規定しておき、カメラと被写体のおよその配置という拘束条件を加えることにより、汎用的に動きパラメータを検出するアルゴリズムを開発してきたため、従来行われていなかった眼鏡処理などの拡張は、動き検出とは独立に可能となった。ここでは実際に眼鏡画像処理を含む知的符号化を実行し、そこに生じる新たな問題を調べる。ポリゴンの頂点を移動させることにより、非剛体としての変形もするワイヤフレームモデル上に剛体として変形しない眼鏡フレームモデルを付加した形での処理を行う。

まず、眼鏡付き顔画像を眼鏡処理なしに直接知的符号化した場合について述べる。図 12 (a) は眼鏡付きの初期画像、図 12 (b) は AU による表情変形を与えた場合の結果で、目の周辺に表情による変形があるため、剛体であるはずの眼鏡が変形している様子が分かる。

##### 4.2 眼鏡画像処理

眼鏡なしの顔画像を入力するためには入力画像系列を撮影する際に被写体となる人物に眼鏡を外してもら

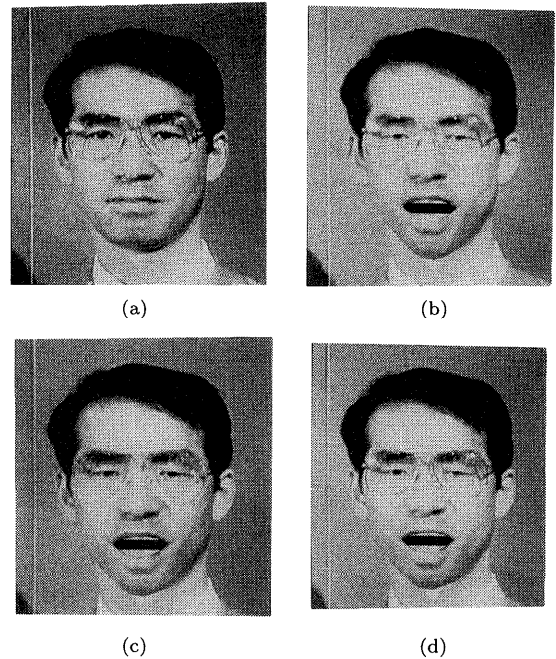


図 12 眼鏡画像処理。(a) 原画、(b) 眼鏡画像処理なしの合成画像、(c) 眼鏡除去処理結果、(d) 眼鏡画像処理付きの合成画像  
Fig. 12 Eye glass processing. (a) Original image, (b) Synthesized image without eye glass processing, (c) Eye glass removal result, (d) Synthesized image with eye glass processing.

い、初期画像の入力を行い、また眼鏡を借りてモデル化に必要な入力を行うという作業が必要である。さらに眼鏡のみ別系統で入力する場合、照明条件を人物撮影時とほぼ同じ状態にしておくことが望ましい。合成後の上書きにおいて眼鏡フレームと顔画像の境界が不自然に強調されたものになるという、照明条件が異なる眼鏡画像を別系統で入力する処理例として、眼鏡フレームの選択のために仮想的に試着するメガネオーダーメイドシステムがある<sup>14)</sup>。実際にメガネを装着できない状況で、そのイメージを視察するのに役立つと思われるが、やはり若干眼鏡エッジが強調されている。

図 13 に眼鏡処理方式の構成を示す。眼鏡処理では、眼鏡フレームやレンズの反射によるハイライト部分に対応する肌色で上書きして眼鏡を除くとともに、眼鏡フレーム部分を切り出し眼鏡フレーム画像として保存

表 4 眼鏡画像処理の方式比較

Table 4 Comparison of eye-glasses processing methods.

	Method A Proposed method	Method B Separated input of eye-glass frame
Input	Facial image with eye-glasses	Facial image and eye-glasses separately transferred
Eye-glasses modeling	From multi-frame images	From multi-view images
Motion Detection	From synthesized vertices of wireframe	
Texture adjustment	LPF for edge	Color adjustment for eye-glass frame Naturalization of edge
Synthesized image	Natural texture	Tolerable for large motion

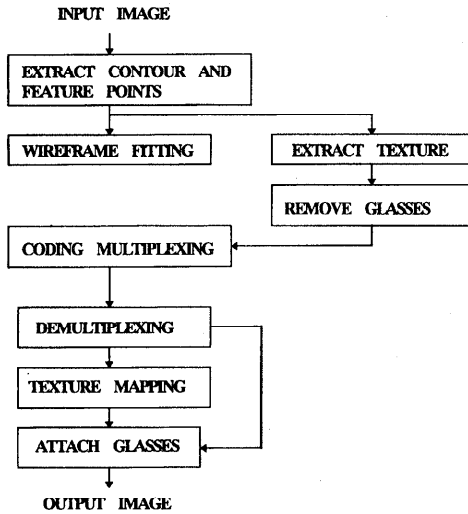


図 13 眼鏡画像処理の流れ図

Fig. 13 Flowchart of eye glass processing.

しておく。眼鏡除去画像を初期画として動きパラメータによりワイヤフレームモデルに変形を与え、初期画像をテクスチャマッピングした後、眼鏡フレーム画像を上書きする。このように眼鏡付き画像を入力として処理する方式に対し、眼鏡のない顔画像を入力し、かつ眼鏡は別系統で入力し、テクスチャマッピングによる顔画像合成後、眼鏡部分を上書きすることも可能である。この結果を図 12(c), (d) に示す<sup>15)~17)</sup>。

表 4 に眼鏡画像処理における上記 2 種の方式の比較をまとめている。今後眼鏡をかけた人物の講演などをビデオテープに収録したものを符号化するような応用も多くなると考えられる。符号化対象である人物が眼鏡を外した顔画像を初期画像としてわざわざ入力したり、眼鏡のみを取り外して撮影したりすることに協力的な場合は B 方式でも符号化が可能だが、実際はそうでない場合も想定されるため、A 方式による符号化が必要となる。

眼鏡処理における上書きする領域の(端)点、領域の接続、閉曲線内の塗りつぶしにより、眼鏡フレーム部、ハイライト部、かげ部など眼鏡の存在に起因する

画像の変化部を肌色または眉などの色で置換えを行う。この置換えはその部分の肌色に近いと思われる部分を数画素サイズのブロックとしてコピーし、上書きする。これは肌色部分は同一色が連続しているのではなく、肌の細かい模様テクスチャとして存在するため、単一色や類似の 1 色を選択し置換すると、肌目が失われて光沢が出たりするためである。さらに置換え画像は類似した肌色のブロックで上書きした場合でも実際の肌色そのものは未知であるため、眼鏡フレームのエッジ状の痕跡を完全に消すには相当な回数の試行作業を要することがある。ここではまず上書き後のエッジ部分に拡散処理を行い、残留するエッジ成分の軽減を図った。また、後述するように、眼鏡を上書きした時点での消え残りや眼鏡の輝度振幅の比較から、操作を終了する基準の設定を行った。

眼鏡の動きに関しては

- (1) 両目の中間点からの  $z$  方向に関する眼鏡の距離を  $d$  とし、眼鏡フレームを 3 次元モデルとして扱う。
- (2) グローバル動きパラメータから眼鏡フレーム部分の 3 次元スケーリング処理を行う。
- (3) グローバル動きパラメータから眼鏡フレーム部分の 2 次元スケーリング処理を行う。
- (4) 実画像の眼鏡両端部の位置座標から、2 次元スケーリング処理を行う。

などの方式が考えられる。(1) は表 4 の B 方式に対応し、眼鏡フレームの 3 次元モデルが必要となる。(2)~(4) はスケーリングのみを行う方式であり、その計算手法が異なるものである。ここでは各画像フレームにおけるワイヤフレームモデルの頂点座標が得られていることから(4)の手法で実験を行った。 $(X_1', Y_1'), (X_2', Y_2')$  を図 14 の 2 点の座標とすると、傾きは  $(Y_2' - Y_1') / (X_2' - X_1')$  となる。ただし、傾きは通常最大でも 30~40 度であると見なせるので、90 度になることはないものとする。抽出された眼鏡フレーム画像をこの傾き値により、眼鏡フレームに向かって左端を中心に回転させる。眼鏡フレームの左端

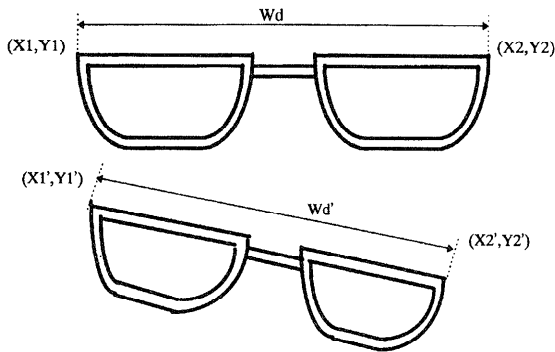


図 14 眼鏡フレームの左右端点の位置. 上: 初期画像の配置, 下: 回転のあるフレームの配置  
Fig. 14 Position of left and right edges of eye glass frame. Top: The case for the initial frame, Bottom: The case with rotating motion.

は、顔画像ワイヤフレームモデル上に設定された眼鏡フレーム左端の頂点位置に固定されている。また  $(X_1', Y_1')$ ,  $(X_2', Y_2')$  の距離により眼鏡フレームのサイズのスケージングがなされる。これにより右端も顔ワイヤフレームモデルの眼鏡フレームの右端の位置に配置されていることになる。図 15 (a) に動きパラメータとして 3 次元回転を含んでいる画像 Claire を使用した場合の処理結果を示す。

#### 4.3 消え残り評価

図 15 (b) に、表 4 の B 方式に相当する別系統で入力された眼鏡を上書きした例を示す。ただし、これは B 方式の様子を調べるためのものであるため、3 次元モデル化や 3 次元処理は省略してある。照明条件が異なるため、眼鏡フレームに違和感がある。またフレーム境界に対して特には処理を行っていないため、境界部分で顔画像との親和性が悪くなっている。一方、図 12 (c) は初期段階で肌色で上書きした眼鏡フレームの痕跡が残っているが、眼鏡フレームを上書きした図 12 (d) では、視覚上の感度は低下していると考えられる。図 16 に眼鏡部分を横切る線上の輝度レベルの変化を示す。眼鏡フレームによるレベル変化は最大で 54 (13 dB, ただし最大振幅を 255 としている), 消え残りと思われる部分の変化は 6~14 で平均 10.75 (28 dB) である。信号としては存在するが、上書きした眼鏡に着目した場合、消え残りは視覚感的には十分小さくなくなると考えられる。ある正弦波の場合、コントラストが 5 倍近いと目立たなくなり、10 倍近いあればほとんど気がつかなくなるという研究がある<sup>18)</sup>。視覚的には、このように消え残り成分があっても主たる信号との差が一定以上あれば主信号がより強く着目され、消え残りは実際より目立たなくなるという性質がある。そこ



(a)



(b)

図 15 大きい回転のある場合の眼鏡画像処理 (Claire 第 90 フレーム). (a) 提案する方式 A 眼鏡画像処理による合成画像, (b) 別系統入力した眼鏡画像を用いた方式 B の合成画像

Fig. 15 Eye glass processing for large rotation image (Claire, the 90th frame). (a) Synthesized image by the proposed method A, (b) Synthesized image by the method B with separated input eye glass image.

でテレビゴースト消去実験における主観評価データを用いて評価することにする。

眼鏡処理画像の消え残りは、眼鏡フレーム部の周辺にやや変形したフレーム画像として重畳しているため、テレビ受信機におけるゴーストのような画像となる。テレビゴーストのレベルとその消え残りの主観評価実験が行われている<sup>19), 20)</sup> が、それによると、事実上気にならないという画質 4 と気になるレベルである画質 3 の境界は DU 比\*で 24 dB となっている。これから、本眼鏡処理においても消え残りレベルの平均がこのレベルに達した時点で眼鏡の消去処理を停止することが妥当と考えられる。ここで、参考として主な D/U 値に対する眼鏡付き画像のゴースト画像を合成し、確認を行った。図 17 (a) は原画に対し右と下の方向にそれぞれ 25 画素の平行移動を行い、ゴーストとして重ねた画像である。(b) は D/U = 18 dB, (c) は同 24 dB, (d) は同 30 dB に対応する。眼鏡部分に関しては、24 dB で存在はするがほとんど気にならない

\* DU 比: 希望波 (Desired) とゴースト波 (Undesired) のキャリアレベルでの比で  $20 \log_{10}(D/U)$  で定義されている。

## Eye-glass Frame

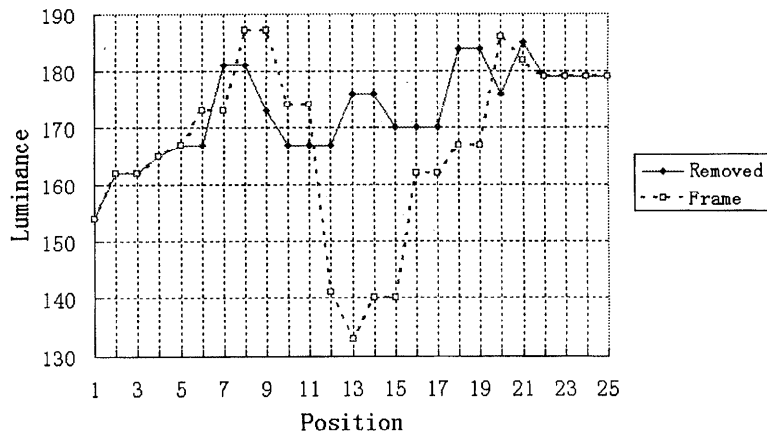


図 16 眼鏡フレーム部分を横切る線分上の輝度とその消え残り  
 Fig. 16 Luminance on the line across eye-glass frame and removed frame.

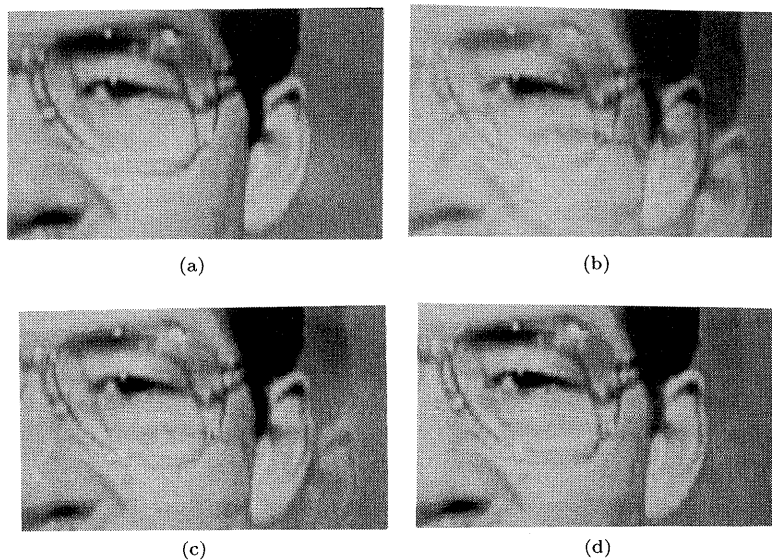


図 17 眼鏡部の消え残り評価のためのゴースト画像による比較. (a) 原画像, (b) ゴースト画像 (18 dB), (c) ゴースト画像 (24 dB), (d) ゴースト画像 (30 dB)  
 Fig. 17 Comparison of ghosted images for evaluating residual eye glass frame.  
 (a) Original image, (b) With 18 dB ghost, (c) With 24 dB ghost, (d) With 30 dB ghost.

レベルであるという基準に合致している。

## 5. ま と め

筆者らは知的符号化にインタラクティブ操作を導入し、三次元動きパラメータ検出などの問題を解決してきた。

本論文では、インタラクティブ知的符号化におけるインタラクティブ操作の曖昧性についての考察に基づ

き、符号化システムにおけるローカル動き検出以降の部分について、手順の明確化と操作量を評価する検討を行った。口などに対しては、gm→lm方式が、目の開閉に関してはlm→gm方式が有利であることを示し、ローカル動き記述順序の操作回数と難易度の比較によりその裏付けを行った。

次に、ワイヤフレームモデル上の顔部品の頂点に階層的な親子関係を定義し、通常の多くの発話シーンで

存在する唇の上下動に対し、全点指定の場合に比べ、操作回数と符号量の削減を図る頂点移動方式を開発した。低ビットレート伝送の特徴を生かしたまま、唇両端点に対する0次、0.5次、1次の周期波形をパターン化することにより、閉口やうねるように曲った唇の形状表現におけるインタラクティブ操作量の削減を図った。

開発したインタラクティブ知的符号化の符号化ツールを拡張し、眼鏡処理を用いた符号化を行い、眼鏡フレームに対する肌色による上書き消去処理の操作量の限界が曖昧になることを新たな問題点として指摘した。定量的な判定基準として、テレビゴーストレベルの主観評価実験データを用いて、有限の操作回数で行うための上書き消去処理操作の終了の指標を作成した。具体的には上書き消去処理の不完全性による消え残りが、D/Uで24dB以上になったところで、処理を終了するという操作に対する基準が作成できた。

以上、操作量を考慮して、実用的な時間内に符号化操作を行うことのできるローカル動き検出ツールの開発を行った。本方式は別に発表するグローバル動きパラメータ検出ツールと合わせてシステム化し、著者らのほかにも技術者が使用し、15種以上の動画系列の符号化が行われている。

今後は、顔動画像から人物全体のモデルとの統合的処理の研究を行っていく。

### 参 考 文 献

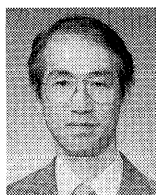
- 1) 大関和夫, 斎藤隆弘, 金子正秀, 原島 博: インタラクティブ知的符号化システム, 電子情報通信学会論文誌 (A), Vol.J80-A, No.8, pp.1293-1303 (1997).
- 2) 金子正秀: 知的画像情報処理, テレビジョン学会誌, Vol.49, No.6, pp.750-759 (1995).
- 3) Ohzeki, K., Saito, T., Kaneko, M. and Harashimam, H.: Interactive Model-based Coding of Facial Image Sequence with a New Motion Detection Algorithm, *IEICE Trans. Communi.*, Vol.E79-B, No.10, pp.1474-1483 (1996).
- 4) Cugini, U., Dell'Oca, M., Merelli, D. and Mussio, P.: An Interactive Approach to the Definition of Digital Image Interpretation, *IEEE Conference on CVPR83*, pp.168-169 (1983).
- 5) 大関和夫, 斎藤隆弘, 金子正秀, 原島 博: インタラクティブ知的符号化における動き検出, 電子情報通信学会技術報告, MVE95-59 (Feb, 1996).
- 6) Aizawa, K. and Harashima, H.: Model-based Analysis Synthesis Image Coding (MBASIC) System for a Person's Face, *Signal Processing: Image Communication*, Vol.1, pp.139-152, Elsevier Science Publishers (1989).
- 7) 森島繁生, 岡田信一, 原島 博: 知的インターフェースのための顔の表情合成法の一検討, 電子情報通信学会論文誌 (D-II), Vol.J73-D-II, No.3, pp.351-359 (1990).
- 8) Ekman, P. and Friesen, W.: *Facial Action Coding System*, Consulting Psychologists, Palo Alto, CA (1977).
- 9) 崔 昌石, 原島 博, 武部 幹: 顔の3次元にモデルに基づく表情の記述と合成, 電子情報通信学会論文誌 (A), Vol.J73-A, No.7, pp.1270-1280 (1990).
- 10) Ohzeki, K., Saito, T., Kaneko, M. and Harashima, H.: Interactive Model-based Coding for Face Metaphor User interface in Network Communications, *Proc. International Conference on Intelligent User Interfaces (ACM IUI97)*, pp.245-248 (Jan. 1997).
- 11) 金子正秀, 小池 淳, 羽鳥好律: テキスト情報に対応した口形状変化を有する顔動画像の合成, 電子情報通信学会技術報告, Vol.IE89-4 (1989).
- 12) Lavagetto, F. and Curinga, S.: Object-oriented Scene Modeling for Interpersonal Video Communication at Very Low Bitrate, *Signal Processing: Image Communication*, Vol.6, No.5, pp.379-395, Elsevier Science Publishers (1994).
- 13) 金子正秀, 羽鳥好律, 小池 淳: 形状変化の検出と3次元形状モデルに基づく顔動画像の符号化, 電子情報通信学会論文誌 (B), Vol.J71-B, No.12, pp.1554-1563 (1988).
- 14) 萩原和浩, 加藤誠巳: ベジエ曲線によるメガネの形状表現とそのメガネオーガメイドシステムへの応用, 情報処理学会論文誌, Vol.36, No.11, pp.2642-2652 (1995).
- 15) Ohzeki, K., Kaneko, M., Saito, T. and Harashima, H.: Interactive Model-based Coding for Multimedia E-mail Environment, *Proc. International Workshop on Coding Techniques for Very Low Bitrate Video (VLBV'95)*, Vol.A-2 (1995).
- 16) 大関和夫, 斎藤隆弘, 金子正秀, 原島 博: 顔動画像のインタラクティブ知的符号化方式, PCSJ95 予稿, 2-3 (Oct. 1995).
- 17) Ohzeki, K., Saito, T., Kaneko, M. and Harashima, H.: Interactive Model-based Coding for Non-realtime Environment, *Proc. 1996 Picture Coding Symposium*, pp.663-668 (Mar. 1996).
- 18) 松井利一, 平原修三: マルチチャネル構造視覚モデルを用いた複合正弦波知覚特性解析, 電子情報通信学会論文誌 (D-II), Vol.J75-D-II, No.9,

pp.1597-1605 (1992).

- 19) 山崎 滋, 宮沢 寛: 複数ゴースト障害の主観評価定量化における基本評価 DU 比, テレビジョン学会誌, Vol.34, No.5, pp.414-419 (1980).
- 20) 小原正晴, 宮沢 寛, 村上純造, 牧野進一, 大関和夫: アダプター型ゴーストキャンセラーのシステム検討, テレビジョン学会, 技術報告, IT52-1, pp.51-56 (Feb. 1982).

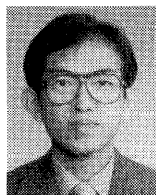
(平成 10 年 1 月 5 日受付)

(平成 10 年 5 月 8 日採録)



大関 和夫

1974 年早稲田大学理工学部数学科卒業。同年東芝総合研究所入所。以来, ファクシミリ符号化, 静止画像符号化, 動画像符号化の研究・開発に従事。1994 年, 通信・放送機構, 映像分析プロジェクトにて, インタラクティブ知的符号化の研究を行う。1997 年, (株) 東芝研究開発センターに戻る。1998 年より (株) 東芝小向工場放送映像機器開発部, 参事。IEEE, 電子情報通信学会, 映像情報メディア学会, 日本顔学会各会員。



齊藤 隆弘 (正会員)

1976 年東京大学工学部電気工学科卒業。1981 年同大学院博士課程修了。工学博士。同年神奈川大学工学部電気工学科専任講師, 現在, 同大学教授。1980 年度電子情報通信学会学術奨励賞, 1982 年度丹羽記念賞, 1989 年度同学会業績賞, 1990 年度同学会論文賞等を受賞。画像入力; 画像通信, 信号・画像処理, 医用応用等の研究に従事。IEEE, 電子情報通信学会, 映像情報メディア学会, 電気学会各会員。



金子 正秀 (正会員)

1976 年東京大学工学部電子工学科卒業。1981 年同大学院博士課程修了。工学博士。同年, 国際電信電話(株)に入社。研究所に勤務。1994 年東京大学工学部助教授。1997 年国際電信電話(株)に戻り研究所に勤務。1998 年電気通信大学大学院電気通信学研究科助教授, 現在に至る。1985 年度日本 ME 学会論文賞等を受賞。画像符号化, 3 次元統合画像通信, マルチメディア情報処理, ヒューマンコミュニケーション, 顔画像処理等の研究に従事。電子情報通信学会, 日本 ME 学会, 映像情報メディア学会, 日本顔学会各会員。



原島 博 (正会員)

1968 年東京大学工学部電子工学科卒業。1973 年同大学院博士課程修了。工学博士。同年同大学工学部専任講師。1975 年同大学助教授, 1991 年同大学教授。通信理論, デジタル信号処理, 画像符号化, マルチメディアと知的通信の融合, 顔画像処理等の感性コミュニケーション, 3 次元空間共有コミュニケーションに関連する研究に従事。電子情報通信学会業績賞(2 回)他受賞。IEEE, 電子情報通信学会, 電気学会, 映像情報メディア学会, 日本顔学会各会員。