

## 3次元構造モデルに基づく自然顔画像の陰影制御

† 佐々木康      † 小野英太      † 森島繁生      ‡ 原島 博

†成蹊大学工学部

‡東京大学工学部

あらまし 実画像には照明の影響が含まれており、テクスチャマッピング手法により画像を合成する場合、原画像に含まれる照明成分は対象物体の表面の模様として合成されている。本論文では、対象を人物顔画像に限定し、人の顔の構造は一般的に滑らかで、構造によって生じる陰影成分も空間的に滑らかであるという仮定を用いて、人物顔画像からテクスチャと照明による影響を分離し、照明による影響のみをCGによる手法を用いて再合成する試みについて述べる。

和文キーワード テクスチャ・マッピング 陰影 分析合成符号化

### *Control Shade effect of Human Face Image by 3D Model*

†Yasushi SASAKI    †Eita ONO    †Shigeo MORISHIMA    ‡Hiroshi HARASHIMA

†Faculty of Engineering, Seikei University

‡Faculty of Engineering, University of Tokyo

Abstract A real facial image includes both a minute texture of human's surface and the illumination effect. A Texture mapping technique projects the texture that include not only minute texture but the shade information onto 3D surface model.

This paper presents a new method to synthesize the face image considering an illumination effect. This method is composed of a separation of the illumination effects from original image, and reproduction of the illumination effects to the synthesized image.

英文 key words Texture Mapping    Shade    Model Base Decoding

## 1.はじめに

テレビ電話やテレビ会議システム、臨場感通信等を目的とする分析合成画像符号化、あるいはコンピュータと人間とのコミュニケーションを円滑化する知的インタフェースの表示部として人物顔画像をリアルに合成する研究は近年盛んに行なわれている。

顔画像の合成方式は、大別すると2種類に分けられる。一方は人物頭部モデルから肌、目、唇のテクスチャまでパラメータ設定し、それらに任意の光源を当てた時の輝度値を計算して画像を合成する、コンピュータグラフィックスによる手法[1]。もう一方は、対象となる人物の正面顔画像を用い、そのテクスチャを3次元のモデルに貼り付けるテクスチャマッピングによる手法[2]である。

前者の手法は、3次元のモデルと合成したテクスチャを持ち、任意の照明環境を設定し、その環境下で任意の方向から見た画像を合成できる[1]。しかし、人物頭部の3次元構造から肌等のテクスチャに至るまでデータとして入力する必要があるため、実際の人物のように生きている顔を合成するのは今のところ困難である。そこで、実際の人物顔画像をテクスチャ解析し、そのテクスチャを元に画像を合成する手法も提案されている[3]。

これに対し、後者の手法では対象人物の2次元のテクスチャ（正面画像を扱ったものが多い）を実際に用いているため、特定人物の顔画像合成は容易で、かつ、かなり自然な画像を得ることが可能である。しかし、2次元画像に投影されたテクスチャを利用するため、正面顔画像が有している照明環境でのみ画像は合成され、任意の照明条件での画像を得るのは困難である。また、正面顔画像を用いる場合、頭部の前半分のテクスチャのみ取得可能のため、頭部の後半分は合成できない。これについては、正面テクスチャ以外に対象人物の全周囲テクスチャと3次元構造を同時に取得し、合成する手法が提案されている[4]。

テクスチャマッピングによる顔画像の合成は、簡単に質の高い画像を合成できる点にメリットがある。この利点を活かしてテクスチャマッピングを行なう場合に、その元となる原画像の撮影の際の照明環境を考慮し、照明によって生じる陰影成分を分離して制御する手法が望まれる。

また、ビジョンの研究として、画像に含まれる照明の影響を分析する試みも行なわれている。光源色と物体色をそれぞれ色空間内のベクトルとし、画像内の各画素の輝度値を2つのベクトルの和で表わし、それぞれのベクトルの強さを用いてハイライト領域を求める2色性反射モデルを用いる手法[5]や、距離画像と輝度値を同時に取得し、正確な3次元構造と輝度値から対象物体の反射係数を求める手法[6]などがある。

また、建物の内装シミュレーションで、壁の模様のようにテクスチャは変化可能であるが、画像に含まれる陰影は変化しない場合のテクスチャマッピング手法[7]や、正面顔画像と3次元モデルから陰影成分を取り除く手法[8]も提案されている。

本論文では、人の顔の構造は一般的に滑らかであり、構造によって生じる陰影成分も空間的に滑らかであるという仮定を用いて、人物顔画像からテクスチャと照明による影響を分離し、照明による影響のみをCGによる手法を用いて再合成する試みについて述べる。まず2.で、テクスチャマッピングによる表情合成手法について述べ、3.で任意照明環境での陰影制御のアルゴリズム、4.で顔画像の領域分割法について解説する。次に、5.で画像に含まれるテクスチャと照明成分の分離法、6.で照明成分の再合成手法、7.で任意照明環境での顔画像とCGで生成した頭髮画像との融合法について説明する。最後に、8.で結論を述べる。

## 2.テクスチャマッピングによる表情合成

人物の顔画像を3次元モデルで近似し、そのモデルを表情、パラメータおよび口形パラメータにしたがって変形、移動させ、その後、テクスチャマッピングにより正面像の自然画像から得られた輝度値をモデルに投影して画像を合成する手法が提案されている[2]。この原理を図1に示す。この手法ではモデルに投影する輝度値（テクスチャ）に、実画像の模様を用いるため、品感そのまま保存され高品質な画像を合成することが可能である。

しかし、テクスチャマッピングでは投影する輝度値に含まれる照明による影響も、結果として対象物体の模様として扱ってしまう。この場合、正面顔画像に含まれる陰影やハイライトといった照明による影響もモデルの動きにともない移動し、不自然な印象を与える。本来、照明による影響は、観察者の視点と光源の位置、対象物体の構造、位置によって決まるものであり、対象物の表面の模様としては扱うのは妥当ではない。したがって、照明による影響と対象物体の模様（テクスチャ）とを分離して別々に扱い、模様のみをテクスチャマッピングし、照明の影響は別にCGの手法で合成する必要がある。

## 3.顔画像のテクスチャ解析と陰影操作法の適用

図2に本手法の原理を示す。本手法では人物の顔画像を対象とし、原画像を同一色相領域に分割して、そのうちの肌領域と唇領域に対し周波数分析を適用してテクスチャの解析を行っている。また、三次元構造モデルを用いCGの技法を利用して陰影の生成をするこ

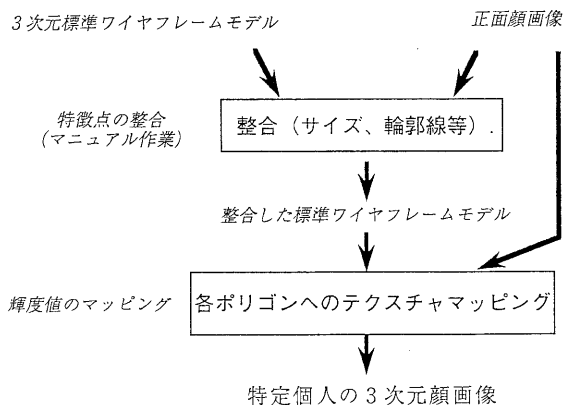


図1. 3次元顔画像の生成

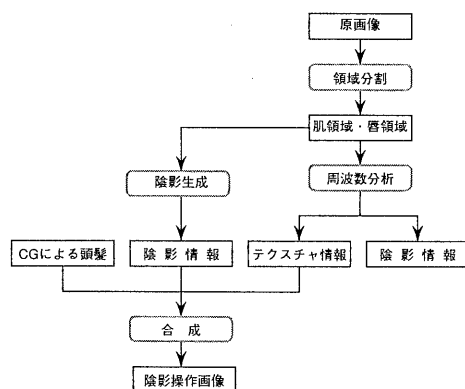


図2. 処理の流れ

により陰影の操作を行なっている。以下にアルゴリズムの流れを示し、次節以降でその詳細について述べる。

#### (1) 領域分割

人間の顔画像は頭髮、肌、目、眉、口領域から構成されている。各領域でその表面形状ならびに反射特性が異なるため、各領域毎に処理をする必要がでてくる。そこで、色相を基にして領域分割を行なう。

#### (2) FFTによるテクスチャ情報と陰影情報の分離

分割した同一色相領域にはテクスチャ情報と陰影情報が混在している。人物顔画像においてはテクスチャの変化に対し陰影の変化は緩やかであると考えられ、テクスチャは高周波成分、陰影は低周波成分に現われると見なすことが出来る。そこで高速フーリエ変換を施し、画像を周波数成分で表した後、高周波成分と低周波成分に分け、それぞれを逆変換してテクスチャと照明成分画像とを得る。

#### (3) テクスチャ情報と陰影情報の再合成

三次元構造モデルを用い、設定した照明環境下での陰影成分をCGの技法により合成し、テクスチャ成分と合成する。これにより任意の照明環境に対応する画像が得られる。

#### (4) 陰影を操作した自然顔画像と頭髮との合成

人物顔画像の各領域のうち、頭髮、眉毛領域については陰影制御は今のところ行っていない。本稿では比較的陰影が顕著にあらわれる人物の肌領域と唇領域

を処理対象としており、頭髮領域にはハイライト成分のみを扱うこととする。このため照明環境が肌、唇領域と等しくなるようにCGで頭髮を合成し、顔画像と融合している。これについては6.で改めて述べる。

#### 4. 領域分割

通常、人間はカラー画像を見る際、同一色相部分を一つの領域として認識している。これには照明による情報は影響を与えない。また、色相の異なる各々の領域によって光の反射特性も異なると考えられる。そこで陰影の処理の前に原画像を同一色相領域に分割する必要がある。

領域分割のアルゴリズムを図3に示す。まず、過分割の原因となる雑音を抑制するために原画像に対して平滑化フィルタを施す。これは3画素×3画素のマスクサイズ内での平均値を中心画素の値とするものである。次に、RGBデータをYIQ信号に変換する。これは、カラーテレビジョンのNTSC方式で用いられるYIQ信号は主に顔画像を最適表示するために考案されたものであるので、顔画像の領域分割には有利であると考えられるからである。

分割方法は、I信号(色差信号)はオレンジからシアンにかけての色を表すので肌領域で大きな値を持ち、Q信号(色差信号)は緑からマゼンタにかけての色を表すので唇領域で大きな値を持ち、Y信号(輝度信号)は毛髪領域で小さな値を持つ。この特徴を利用して、しきい値処理により同一色相領域に分割し、本手法の対象となる肌領域と唇領域を抽出する。また、領域分割の際にしきい値処理がからむが、しきい値を自動的に決定する手法についても検討を進めている[9]。ここではマニュアルでとりあえずしきい値を決定した。

図4に領域分割した結果を示す。

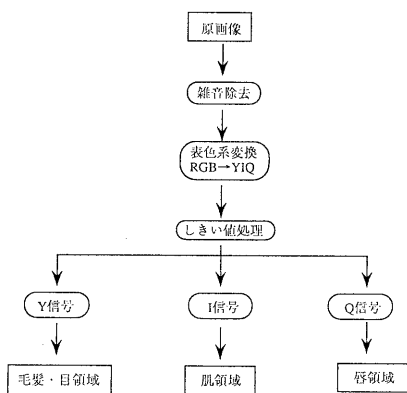
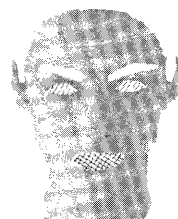


図3. 領域分割アルゴリズム



斜線部: 目領域  
 網かけ部: 唇領域  
 白色部: 毛髪・背景領域

図4. 領域分割結果

### 5. テクスチャ情報と陰影情報の分離法

本節では顔画像のテクスチャ情報は高周波成分、陰影情報は低周波成分であるという周波数特性の相違を利用した分離法について述べる。本手法ではこの分離に当たって対象物体の三次元構造や照明の配置といった幾何学的な情報を用いないことが特徴である。抽出されるテクスチャ情報は対象物体本来の色彩との差分情報として得られるので正負の値を持つ。

#### (1) RGB表色系による分離法

周波数分析には二次元FFTを用いる。入力画像は512画素×512画素とする。領域分割された画像の画素値の無い部分（領域外の部分）には対象物体のR, G, Bそれぞれの成分の平均値を入れる。これによりエッジ部分での高周波成分の発生を抑制することができる。この画像のR, G, B成分それぞれに対し二次元FFTを施し512×512個の周波数成分に分解する。周波数軸上で高周波成分と低周波成分とに分離し、それぞれに逆FFTを施し、テクスチャ成分と照明成分とを得る。図5に肌領域を画像とそれから求めたテクスチャならびに照明成分を示す。テクスチャについては正負の振幅となるため128のバイアスをかけている。

#### (2) 知覚的表色系による分離法

通常、カラー画像ではRGBデータが用いられる場合が多いが、これから色彩情報を得ることは困難なので、RGBデータを人間の色知覚の三属性である色相、明度、彩度に変換することによりデータを感覚的に扱いやすくする。今回はHVC表色系を用いた。RGBデータからHVCへの変換には数学的な処理によって変換を行なうMIM法を用いた[10]。

本手法での対象物体は同一色相としているので明度と彩度について考慮すればよいことになる。よって、明度と彩度に対し、RGB表色系の場合と同様に、それぞれFFTを施し、その後しきい値処理によりテクスチャ情報と陰影情報を分離する。

### 6. 任意の照明環境下における画像合成

分離した陰影情報を任意の照明環境に対して生成することができれば、テクスチャ情報と合成し、目的の任意照明環境下での顔画像を得ることができる。

陰影情報の生成には三次元構造モデルの一つであるワイヤフレームモデル（三角形パッチ数531）を用い物体表面の光の反射を計算することにより陰影を生成する。

顔の3次元モデルの表面は各領域毎にパラメータは異なるが、拡散反射と鏡面反射を起こすモデルで近似している。ここで、鏡面反射を起こす画素ではその輝度値は光源色となるが、本論文では光源は昼光であると仮定している。

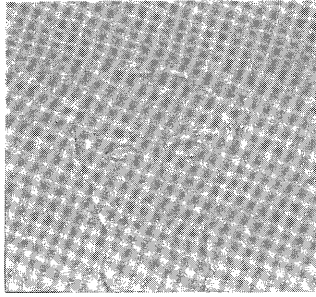
#### (1) RGBを用いた場合の陰影情報の生成法

対象物体の画素値を対象領域全体のRGB成分それぞれの平均値に置き換える。ここでは、この平均値を対象物体本来の色彩であると仮定する。この対象領域全ての画素(x, y)に対し以下の式を用いて陰影付けを行う。

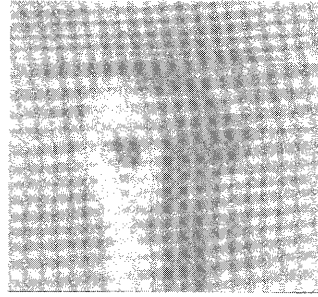
$$l_d(x,y) = k_a \vec{n}(x,y) \cdot \vec{L} + l_s \quad (1)$$

$$i'(x,y) = i(x,y) \cdot l_d(x,y) \quad (2)$$

ここで(1)式の $l_d$ は陰影の原因となる拡散反射成分、 $k_a$  ( $0 \leq k_a \leq 1$ ) は陰影の強度を表わす重み係



(a) テクスチャ成分



(b) 照明成分

図5. 肌領域のテクスチャ成分と照明成分

数、 $\vec{N}(x,y)$ は座標 $(x,y)$ における面法線単位ベクトルで三次元構造モデルにより求まる。 $\vec{L}$ は入射光単位ベクトルであり任意の値を入力することができる。 $l_s$ は環境光で、拡散反射成分 $l_d$ の取り得る最大値が1となるように設定する。(2)式の $i'$ は陰影付加後の画素値で $i$ は陰影付加前の画素値(対象領域の平均値)である。

鏡面反射の影響は拡散反射のみの場合より輝度値が上がり、色は白色に近づくと考えられるので次式のようにして表現する。

$$\vec{R}(x,y) = 2 \cdot (\vec{L} \cdot \vec{N}(x,y)) - \vec{L} \quad (3)$$

$$l_s(x,y) = k_s (\vec{R}(x,y) \cdot \vec{S})^n \quad (4)$$

$$i''(x,y) = l_s(x,y) \cdot i_{max} - i(x,y) + i'(x,y) \quad (5)$$

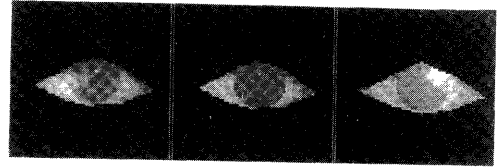
ここで(3)式の $\vec{R}$ は反射光単位ベクトル、(4)式の $l_s$ は鏡面反射係数、 $k_s$ ( $0 \leq k_s \leq 1$ )は鏡面反射の強度を表わす重み係数、 $\vec{S}$ は視線単位ベクトルで、 $n$ は鏡面反射光の広がる度合を示す指数である。(5)式の $i''$ は鏡面反射と拡散反射を考慮した画素値で $i_{max}$ はRGB各成分の最大値(255)である。

各画素のRGB成分それぞれに対して $i''(x,y)$ を求めることにより陰影情報を生成することができる。

#### (2) 知覚的表色系の場合の陰影情報の生成法

対象物体各々の画素の明度値と彩度値を対象領域全体の平均値に置き換える。これを対象物体本来の色彩であると仮定する。陰影付けは明度値に対して(1)、(2)式を用いる。鏡面反射の影響は彩度に表れ、その変化は明度の変化とは逆になり、鏡面反射成分が強いほど彩度値は小さくなる。よって、次式の様にして表現する。

$$c'(x,y) = c(x,y) \cdot (1 - l_s \cdot k_s) \quad (6)$$



(a) 原画像

(b) ハイライト除去画像

(c) ハイライト付加画像

図6. 目のハイライトの除去と付加

ここで、 $c'$ は鏡面反射を考慮した彩度の画素値、 $c$ は計算前の画素値(対象領域の彩度の平均値)である。

以上により陰影情報を生成することができる。

#### (3) 目・眉の照明効果の除去

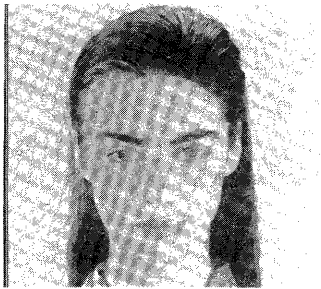
CGで陰影を付加する際には、拡散反射が最大の時の輝度値に対しどの程度輝度値が減衰しているかを計算している。したがって、元画像のテクスチャにハイライトが含まれているとその部分の本来の輝度値が未知なため照明が変化したときの輝度値を計算できない。そこで、目などのハイライト部分を本来の色と思われる色に置き換える。また、目、眉、共に左右が同じ輝度値になるように輝度補正をする。

#### (4) 目のハイライトの付加

目領域は肌や口領域に比べ表面がなめらかなため、眼領域で生じる鏡面反射は鋭くスパイク状になる。そこで、眼領域の3次元構造も滑らかにするために眼球を球モデルで近似し(4)、(5)式を用いてハイライトを付加する。図6にハイライトの除去と付加の結果を示す。



(a) 原画像



(b) RGBの場合



(c) 知覚的表色系 (HVC) の場合

図7. 陰影操作画像

#### (5) テクスチャ情報との合成

生成した陰影情報にテクスチャ情報を合成することにより本研究の目的である任意の照明環境に対応する自然顔画像を得ることができる。図7にRGBの場合と知覚的表色系の場合の結果を示す。当研究室でRGBの場合と知覚的表色系の場合ではどちらの方法が望ましいか評価を行なったところ各個人の評価の基準の違いにより意見が異なったため、本論ではその点については述べない。

#### 7.CGで生成した頭髮との融合

人間の頭髮は、髪の毛の構造が非常に細かいことと、それにより構成される形状が複雑なことからモデル化が難しい。また、髪の毛1本1本の鏡面反射が強く、照明の影響を多大に受けるため、モデルに頭髮の輝度値を投影するだけでは任意の照明環境での画像を合成するのは困難である。そこで、CGで頭髮を合成する研究が盛んに行なわれており、近年高品質な頭髮を合成できるようになっている[11]。本章では、このようなCGによって生成された頭髮と自然顔画像との融合について述べる。

##### 7-1 CGによる頭髮生成

本研究では、髪の毛の微細構造を直接的に幾何モデルで近似し、レンダリングする手法[11]を用いて頭髮を生成した。この手法は、頭部の3次元モデルの表面からモデル化した髪の毛を1本ずつ生成し、そのモデルに対して設定した照明環境に対する輝度値を計算している。髪の毛のモデル化には様々なモデルが提案されているが、ここでは円柱モデルを使い、3次Bezier曲線にしたがった滑らかな空間曲線となる様にモデル化したものを用いている。図8に髪の毛のモデルを示す。今回の生成した髪の毛の本数は約5万本で、対象人物の正面顔画像にたいし整合した3次元モデルの頭部から髪の毛を生成した。照明環境は肌および口領域と同一の環境としている。

本論文では、CGにより生成された頭髮と実際の人物顔画像に照明成分の処理を施した画像との融合について述べ、髪の毛のデザインの詳細については省略する。

##### 7-2 CGで生成された頭髮と自然顔画像の融合

実際に撮影された人物顔画像とCGによって合成された頭髮画像をそのまま融合しようとすると、髪の毛と顔が分離した様な不自然な画像が合成される。この原因としては、顔画像と頭髮画像の照明環境が一致していないことが挙げられる。実画像の照明環境が既知であれば、その照明環境に対する頭髮画像を合成できるが、一般に入力される顔画像の照明環境はほとんどが未知であり、また画像から照明環境を求めることも

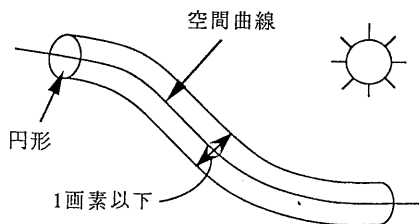


図8. 髪の毛のモデル

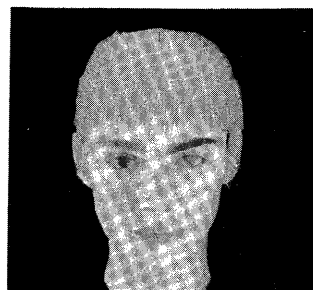


図9. CGによる頭髪との融合結果  
(頭髪の色は原画像と異なる)

現在の段階では不可能である。そこで、未知の照明環境の影響を受けている人物顔画像から照明成分を取り除き、頭髪画像を合成した時と同一の照明環境を顔画像に付加して頭髪画像と融合させ照明環境を統一した。融合した結果を図9に示す。

頭髪領域と肌領域で照明環境が一致しているため、比較的自然な画像が合成できている。これまでCGで生成された頭髪は実際の人物に髪の毛を生成したものは少なく、マネキンに髪の毛を生成したもののばかりであったが、本手法を用いることで任意の対象人物に任意の頭髪を合成することが可能となった。

#### 8.まとめ

本論文では、人物顔画像を対象に、画像に含まれる照明成分を除去し、任意の照明環境での照明の効果を付加した画像を合成する手法を提案し、その結果を評価した。また、自然画像に含まれる陰影成分を制御することでCGで生成した画像と違和感なく融合することが可能となった。

現在の段階では正面顔画像に対して処理を行なっているため、人物頭部が側面を向いたときにテクスチャの情報量が減少している。そこで、現在正面画像と側面画像から全周囲のテクスチャを融合する手法について検討を進めている。

#### 参考文献

- [1]N.M.Thalman:"Visualizing Humans by Computer",  
The 3rd International Forum on the Frontier of  
Telecommunications Technology,1991
- [2]S.Morishima et al:"A Proposal of A Knowledge Based  
Isolated Word Recognition", ICASSP'86,Vol1.1,  
pp713-716,1986

- [3]石橋、岸野："領域分割を用いた人物像の色彩  
テクスチャ分析・合成"、信学論(D-2), J74-D-2  
No.11, 1991
- [4]渡辺、末永："3D形状と輝度(色)を同時計  
測可能なスキャナとその応用"、第6回  
NICOGRAPH90,pp181-189,1990
- [5]Gudum.J.klinker, Steven.A.Shafer, Takeo Kanade:  
"Using A Color Reflection Model to Separate  
Highlights from Object Color",  
1st, ICCV, pp145-150, 1987
- [6]Katsushi Ikeuchi, Kosuke Sato:  
"Determining Reflectance Parameters using Range and  
Brightness Images",  
3rd, ICCV, pp12-20, 1990
- [7]森田、加藤、宮岡："陰影を反映したテクスチャ  
マッピング方法の検討"、電子情報通信学会春  
季全大 D-464, 1992
- [8]荒川、登、前原："知覚表色形と3次元構造モ  
デルを用いた陰影効果の除去"、電子情報通信学  
会秋季全大 A-155, 1992
- [9]五味、小野、森島、原島："自動閾値設定によ  
る顔画像からの口領域抽出"、電子情報通信学会  
春季全大D-435,1992
- [10]宮原、吉田："色データ(RGB)-(HVC)数学的変  
換方法"、テレビジョン学会誌Vol43, No.10, 1989
- [11]小林、森島、原島："糸状物体の運動モデルと  
CGによるシュミレーション"、電子情報通信学  
会 PRU90-127, 1991