肌色領域抽出手法の検討

正員 足達義則 (中部大学)

非会員 今井昭宏 (中部大学)

正員 尾崎正弘 (名古屋女子大学)

非会員 石井直弘 (名古屋工業大学)

Study on Extraction of Flesh-Colored Region

Yoshinori ADACHI, Member, Akihiro IMAI, Non-member (Chubu University), Masahiro OZAKI, Member (Nagoya Women's University), Naohiro ISHII, Non-member (Nagoya Institute of Technology)

Five color spaces, RGB, XYZ, YIQ, HSV and LUV color spaces, were examined to know which was the most suitable to extract a face region. From the observations of flesh-colored pixel distributions in those color spaces, a face region was most successfully extracted by an ellipse from the UV plane in the LUV color space. A proposed method using LUV color space and ellipse extraction was an extension of that proposed by Hongo et al. that used a circle instead of an ellipse, and it improved extracted results by 6 points in our agreement index. A real time flesh-colored region extraction system was constructed on a personal computer with a 400MHz celeron CPU. It could extract flesh-colored regions with the rate of 12 frames, 160×120 pixels each, per second.

キーワード : flesh color, LUV color space, ellipse, extraction, real time system

1. はじめに

我々が、日常難く行なっている個人識別を、コンピュータで自動に行う研究が盛んに行われている⁽¹⁾⁽²⁾。特に、銀行や犯罪捜査などのセキュリティシステムでの使用を目的とした顔画像による個人識別に関する研究⁽³⁾が多くなされてきた。また最近では顔画像から表情、向き、動きを解析し、それらの情報をもとにコンピュータと人間の間のインターフェイス (マンマシンインターフェイス) をより自然な操作で行なえるようにしようという研究が行われている⁽⁴⁾⁽⁵⁾。さらにテレビ電話などの画像処理を伴う技術の普及なども合わせ、様々なアプローチで研究が行われており、ますます「顔に関する情報処理」は重要になっていくと考えられる。

これらの研究の多くは、顔の各部品 (目・鼻・口など) の形状や、相対的な配置を重要な特性としている⁽⁶⁾。そして、顔の各部品を抽出する前段階として顔領域を特定することが多く、そのために画像中での肌色の領域 (以下、肌色領域) を適切に抽出する処理を行うことが重要となってくる。

しかし、画像中での肌色の画素は照明条件や個人差などによって、色や濃度値が動的に変化するために安定した抽出を行うことは難しい。

本研究では、色を扱う上で最も一般的であると考えられる RGB 表色系ではなく、明度情報と色情報が分離されており、2つの色の色空間上での距離が、人間の目を感じる色の差と比例するように均等に定められた均等知覚色空間の一つである LUV 表色系を選択する。その UV 軸で構成される画素数の2次元ヒストグラムにおける肌色領域の分布形状を考慮することにより、照明条件や個人差などにより肌色領域が画像ごとに異なる場合でも適切に肌色領域が抽出できるようにした。また、実際に本手法の処理をリアルタイムで行うシステムを開発し、顔領域の追跡を行った。

以下では、まず5種類の色空間上における肌色領域の分布状態を示す。次に、LUV 表色系を使った本手法について述べる。そして、本手法の有効性を検証するために行った実験とその結果について述べる。最後に、本研究で開発したリアルタイムで肌色領域を抽出し顔領域の追跡を行うシステムについて述べ、まとめと今後の課題について述べる。

2. 色空間と肌色領域の分布

通常、我々が扱う画像信号は R(赤)、G(緑)、B(青) の3原色で表される RGB 表色系で表現されているが、色相と明度が区分されておらず、照明の違いによって RGB 値が大きく変動する。このため、肌色領域を分離する際、広い

図 1 肌色領域の分布を検討する画像
Fig.1. Sample image to examine flesh-colored distribution.

範囲を検討しなければならないなどの特徴がある。表色系にはこの他にも様々な表色系があるが、ここでは混色系である RGB 表色系、XYZ 表色系、YIQ 表色系、顕色系の HSV 表色系と LUV 表色系における肌色領域の分布について検討する。

〈2・1〉各色空間での肌色領域の分布 肌色領域の抽出に適した色空間を検討するために、図 1 に示す日中室内で撮影した背景を含む人物画像を用いた。これは、様々な条件下で撮影した画像に対し色空間上での分布を調べた結果、相対的な分布形状が余り変化しないことから、日中室内（明るく影が少ない）での画像を検討することにした。図 1(a) はオリジナル画像であり、(b) は肌色画像、(c) は背景画像である。これらの画像を各色空間に投影し、各画素を色空間内の散布図として表示し、肌色空間の分布形状と背景領域との分離の程度から、肌色領域抽出に最も適した色空間を探る。

図 2 の RGB 表色系と図 3 の XYZ 表色系では、R、G、B の線形変換で X、Y、Z を生成しているためか分布の形状が非常によく似ており、どちらにも同じことが言える。これらの空間では肌色領域と背景とのオーバーラップは少ない (XY 平面を除いて) が、肌色領域そのものの分布形状を単純な式で表現することは難しい。

図 4 に示す YIQ 表色系では、(a) と (c) の分布がよく似ており、肌色領域が背景の中に溶け込んだ形になっている。したがって、RGB 表色系などよりも抽出はさらに難しくなっている。

図 5 の HSV 表色系では、背景は H 軸に関して幅広く分布しているが、肌色領域はコンパクトな分布になっており、混色系に比べて抽出は容易になっているが、肌色領域の形状が単純でなく、背景とも混ざっているため切り出しが難しい。

図 6 の LUV 表色系では、肌色領域は分布の明確な楕円形状 (特に UV 平面) を示しており、容易に式で表現できる。しかし、背景と混じった部分も多く見られ、肌色領域

図 2 RGB 表色系における画素の散布図
Fig.2. Pixel distribution in RGB color space

図 3 XYZ 表色系における画素の散布図
Fig.3. Pixel distribution in XYZ color space

のみの抽出が難しいことが伺える。

これらの検討の結果、一長一短はあるものの、肌色領域が容易に楕円の式で近似でき、抽出が容易な LUV 表色系と楕円切り出しを提案する。

〈2・2〉楕円体による肌色領域抽出手法の提案 LUV 表色系を用いて肌色領域を抽出する試みは、本郷⁽⁶⁾が行っている。彼らは、UV 平面における基準肌色値を予め求めておき、ここからの距離ヒストグラムを判別分析で 2 値化してダイナミックに閾値を求め、そこまでの距離を半径と

図 4 YIQ 表色系における画素の散布図
Fig. 4. Pixel distribution in YIQ color space

図 6 LUV 表色系における画素の散布図
Fig. 6. Pixel distribution in LUV color space

図 5 HSV 表色系における画素の散布図
Fig. 5. Pixel distribution in HSV color space

図 7 肌色領域抽出のための楕円の決定
Fig. 7. Obtaining an ellipse for flesh-colored region extraction.

して円形の切り出しを行っている。

これに対して、我々は画素ヒストグラムの分布形状を検討した結果、より適切な楕円による切り出しを提案する。ここで楕円の長径、短径は以下のように決定する。

まず、図 7 に示すように、本郷らと同様にして求めた基準肌色値 (U_{max}, V_{max}) を通り、傾き 1 の直線を

$$V = (U - U_{max}) + V_{max} \dots\dots\dots (1)$$

とし、この直線上での画素ヒストグラムを作成する。ヒストグラムの高さが基準肌色値での値の 10% 以下、または判

別分析で分離されるところまでを肌色領域の長軸とする。長軸の中心 (P, Q) を通り、式 1 に直交する直線

$$V = -(U - P) + Q \dots\dots\dots (2)$$

に対しても同様に、画素ヒストグラムから両端を決定し、両端までの平均距離で短軸を決定する。

長軸、短軸の長さをそれぞれ A, B としたとき、抽出領

域を規定する楕円は

$$\left[\frac{(U-P)+(V-Q)}{A/2}\right]^2 + \left[\frac{(U-P)-(V-Q)}{B/2}\right]^2 = 1 \dots\dots\dots (3)$$

となる。

(U_{max}, V_{max})の値は、画像毎の明るさなどによって多少変化するが、日中の明るい室内から薄暮の中、および夜間蛍光灯下で撮影した画像まで幅広く調べた結果、

$$U_{max} = 6 \sim 25 \quad \text{全体の値域：} -134 \sim 220$$

$$V_{max} = 14 \sim 32 \quad \text{全体の値域：} -140 \sim 122$$

となっていた。

3. 実験結果

つぎに、本手法による抽出結果を、本郷らの手法と比較する。一致率は、長谷ら⁽⁷⁾が提案した式と同形の

$$\eta = \frac{s - s_1}{s - s_2} \times 100 \dots\dots\dots (4)$$

で評価する。ここで、 s は理想肌色領域、 s_1 は欠損部分、 s_2 は過剰抽出部分の画素数である。また、抽出形状の適切さは欠損率を評価する式

$$\zeta = \frac{s_1}{s} \times 100 \dots\dots\dots (5)$$

で、抽出範囲の適切さは過剰率を評価する式

$$\xi = \frac{s_2}{s} \times 100 \dots\dots\dots (6)$$

で調べる。

人物、照明条件の異なった30枚の画像に対する上記3つの指標の平均値を表1に示すと共に、抽出結果の例を図8と図9に示した。図8は薄暮の中外に顔を向けた状態で撮影した画像であり、全体に暗いが影が少ない。一方、図9は日中明るい室内で側面から光が差している状態で撮影しており、顔半分に影ができています。

本研究では、一致率90ポイント以上、欠損率10ポイント以下、および過剰率10ポイント以下を目標としたが、表1に示すように一致率は平均87.4ポイントと若干目標に達しなかった。しかし、欠損率や過剰率は目標値を下回り、目標に達することができた。また、本郷らの方法に比べ過剰率が1ポイント程度しか変化していないにもかかわらず、一致率や欠損率は6～7ポイントも向上しており、本手法によってより適切な抽出が行われたことを示している。このことは図8や図9にも現れており、首や額の部分の欠損部分が改善している様子が分かる。具体的には、図8では、一致率88.6ポイント、欠損率7.85ポイント、および過剰率4.01ポイント（本郷らの方法ではそれぞれ、86.1、11.0、および3.36ポイント）となっており、図9では、一致率73.6

表 1 抽出結果の比較

Table 1. Comparison of extracted flesh-colored regions

	本郷らの手法	提案手法
一致率 η	81.7	87.4
欠損率 ζ	14.1	7.50
過剰率 ξ	7.21	8.15

図 8 肌色領域抽出結果の比較 (1)

Fig. 8. Comparison of extracted flesh-colored regions(1).

図 9 肌色領域抽出結果の比較 (2)

Fig. 9. Comparison of extracted flesh-colored regions(2).

ポイント、欠損率16.9ポイント、および過剰率12.8ポイント（本郷らの方法ではそれぞれ、62.8、31.6、および8.82ポイント）となっていた。図8のように陰影の差が小さい場合には満足のいく結果が得られるが、図9のように陰影の差が大きい場合には影の部分の抽出が難しく目標値を下回る結果となっているが、本郷らの方法による結果よりも大きく改善している。

4. 実時間肌色領域抽出システムの開発

ここでは、本研究で提案した肌色領域の抽出手法の処理をリアルタイムで行うシステムの開発について述べる。

図 10 実時間肌色領域抽出システムの概要
Fig. 10. Schematic diagram of the flesh-colored region extraction system.

表 2 システムの開発環境
Table 2. Environment of the system development

項目	内容
計算機	PC/AT 互換機 (Celeron 400MHz)
ビデオカメラ	MASTACS MOVIE S63(Hitachi)
ビデオキャプチャカード	GV-VCP/PCI(I/O Data)
OS	Microsoft Windows95, 98
開発言語	Visual C++ 5.0

〈4・1〉 システムの概要 本システムは、ビデオカメラからの画像のストリームデータをパーソナルコンピュータに搭載したビデオキャプチャカードで取得し、取得画像に対し本研究で提案した肌色領域の抽出手法の処理をリアルタイムで行うものである。これにより、画像中の肌色領域がリアルタイムで抽出されるために顔領域や、他の肌色領域の位置特定および追跡が可能となる。システムの大まかな流れを図 10 に示す。

〈4・2〉 開発環境 開発したシステムは、表 2 の環境で動作する。

〈4・3〉 動作過程 動作の具体的な手順を次に示す。

- (1) 使用するビデオカメラの特性と照明条件が異なるので、予め基準肌色値 (U_{max}, V_{max}) を特定しておく。しかし、LUV 表色系の特徴で値に大きな変化はない。
- (2) 高速に肌色領域を特定するため、縦横とも 2 画素に 1 画素の割合で間引いて調べる。しかし、この間引き間隔は処理系のスピードに依存する。
- (3) ビデオカメラから取得しているのは RGB のバイトストリームデータであるので、LUV 表色系に変換する。
- (4) UV 平面に射影されたデータ画素から式 1, 2 上でのヒストグラムを調べ、楕円 (式 3) を決定する。
- (5) この楕円内に存在する画素の縦横方向の最大最小を調べ、これらの値よりも 1 画素分 (間引き間隔に

図 11 実時間肌色領域抽出システムの動作例
Fig. 11. Example of the extracted flesh-colored region by the proposed system.

依存する) 外側までの領域で、最初の探索で間引いた画素について調べる。

- (6) 抽出した肌色部分のみを残して、他の部分は黒くした状態で画面に表示する。

〈4・4〉 システムの動作状況 本システムの動作状況を確認するため、日中室内と夜間蛍光灯下の 2 種類の照明条件下での肌色領域の抽出状況を調べた。

その結果、日中室内、夜間蛍光灯下ともに安定した抽出をリアルタイムで行うことができることを確認した。実行速度は 160×120 の解像度で 1 フレームあたり約 0.08 秒 (毎秒 12 フレーム相当) であった。図 11(a) は日中室内での動作例、(b) は夜間蛍光灯下での動作例である。また、今回は顔領域の特定を目的として肌色領域の抽出について研究を行ったのであるが、他の肌色領域抽出を前処理として必要とする様々な研究 (例: 手話認識を目的とした手領域抽出など) にも応用できると考えられる。図 11(c) は本システムにより手領域をリアルタイムに抽出している様子である。図中、抽出結果の画面の中に画面表面での反射光が所々に写っているが、CRT 画面を撮影した為であり、抽出結果とは無関係である。

5. ま と め

本研究の結果、以下の事柄が明らかとなった。

5 つの色空間における肌色領域の分布を調べ、LUV 表色系に射影することによって最も容易に抽出できることを確認した。

肌色領域を LUV 空間上の楕円領域で近似することによって、これまででない精度で抽出が可能になった。

ほぼ実時間 (12 フレーム/秒) での肌色領域抽出処理が可能となった。

しかし、今回目標とした一致率 90 ポイント以上は実現できず (本研究の結果は平均 87.4 ポイント)、今後の課題で

ある。また、顔の認識等の処理を含めた場合、現在の処理速度で十分であるとは言えず、今後、高速化が必要であると考えられる。これには、前フレームの抽出結果の利用による追跡や MMX 命令の利用、一部アセンブラの実装などが考えられる。また、本研究では肌色領域が点在している場合には、無駄な領域も調べてしまっている。高速化のためには、部分領域で調べるように変更する必要があると考えられる。

(平成 12 年 04 月 07 日受付, 同 12 年 08 月 02 日再受付)

文 献

(1) 平山泰崇, 中村 納: "表情変化にロバストな新しい個人識別方法", 信学技報, PRMU97-132(1997-11).

(2) Rama Chellappa, Charles L. Wilson, and Saddi Derohey: "Human and machine recognition of face", A survey. Proceeding of The IEEE, Vol.83, No 5, pages 704-740,1995.

(3) 吉野峰生: "犯罪捜査における顔画像の個人識別", 信学技報, PRMU99-28(1999-06).

(4) 中村修二, 山下久志, 佐藤公則, 高橋行俊: "動画像における部位画像の追跡に関する一検討", 信学技報, PRMU98-69(1998-07).

(5) 赤松 茂: "コンピュータによる顔の認識 -サーベイ-", 電子情報通信学会論文誌, Vol.J80-D-II No.8, pp.2031-2046 (1997. 8).

(6) 本郷仁志, 羽飼直記, 山本和彦: "4 方向面特徴と色面特徴の統合による顔および顔部品の抽出", MIRU98, II (307-312).

(7) 長谷博行, 丸山博, 松下靖征, 米田政明, 酒井充: "カラー文書画像中の文字領域抽出のための領域分割方式", 信学技法, PRMU97-217(1998-02).

足 達 義 則 (正員) 1953 年生。1976 年 3 月名古屋大学理学部数学科卒業。同年名古屋工業大学助手。1988 年中部大学講師。1990 年同助教授。1996 年同教授。工学博士。主としてパターン認識に関する研究に従事。

今 井 昭 宏 (非会員) 1975 年生。2000 年中部大学経営情報学研究科博士前期課程修了。同年 日立システムアンドサービス入社。経営情報学修士。

尾 崎 正 弘 (正員) 1946 年生。1984 年慶應義塾大学文学部哲学科卒業。1989 年岡崎女子短期大学講師。1992 年同助教授。1997 年名古屋女子大学助教授。1999 年同教授。主としてパターン認識に関する研究に従事。

石 井 直 宏 (非会員) 1940 年生。1963 年東北大学工学部電気工学科卒業。1968 年同大学助手。1975 年名古屋工業大学助教授。1982 年同教授。工学博士。主として人工知能に関する研究に従事。