

ディスプレイ光反射を利用した 顔領域検出の研究

坂本 真吾 前田 佐嘉志 鶴田 直之
福岡大学大学院 工学研究科

顔認識や表情認識を用いた高度なヒューマンインターフェースを構築するには、顔の検出ができることが前提となる。この顔検出が安定に行うことができるかどうか、顔認識や表情認識の精度を大きく左右する。従来の肌色情報を用いた顔検出法には背景の肌色物体の誤検出をする、環境光の色変化に弱いという問題があった。本論文では、ディスプレイの照射光を用いた顔領域の検出法を述べる。映り込みの変化量から物体本来の色を見つけ、背景の肌色物体の除去することで、背景や環境光に左右されない顔領域の検出をする。

Face Detection using Reflection of Display Light

Shingo Sakamoto Sakashi maeda Naoyuki Tsuruta
Graduate School of Engineering, Fukuoka University

In order to build the advanced human interface using face recognition or expression recognition, it will be the requisite that detection of a face can be performed. Whether this face detection can carry out stably influences the accuracy of face recognition or expression recognition greatly. There was a problem that detects the skin color object of a background that it was weak to color change of environmental light in the face detecting method using the conventional beige information. In this paper, a method of face detection using reflection of display light is described. The intrinsic color of object is found from the amount of change of the reflection by this method, the face detection can carry out stably under various light condition and complex background.

1 序論

1.1 背景

人間のコミュニケーションにおいて顔は重要な役割を果たしている。それは、人間が視覚によって顔から多様な情報を読みとることで滑らかに意志の疎通をはかることができるからにほかならない。視覚的に顔から得られる情報は、感情、興味、意志などの心理的情報、体調などの身体的情報や口の動きに示される発話情報など多岐にわたっている。このような顔から得られる情報をコンピュータが読みとることができるようになれば、様々な可能性が期待できる。その一つに在宅医療への応用がある。

現在、我々は高齢化社会で生きている。高齢化が進むにつれ、高齢者介護施設の欠乏により、いずれ施設に入り切れない高齢者が自宅での療養を余儀なくされることは明らかである。今後の在宅医療では、患者となる高齢者は健康状態を定期的にコンピュータで検診することとなる。そこで、コンピュータの操作に不慣れな高齢者にでも、機械を意識せず自然に人間と対話ができるような擬人化エージェントの実現が必要となってくる。

擬人化エージェントの実現というヒューマンインターフェースの分野において、顔認識や表情認識を行うにあたり、画像から人物の顔の位置を判断することが最低条件となる。

在宅医療に応用するにあたり、背景や環境光に左右されることなく、患者だけを対象とした実時間での正確な顔領域検出が求められる。

顔検出は、移動→接近→対話という3モード解析のうち、接近モードと対話モードとの切替えの際に必要であり、瞬時に顔検出ができれば滑らかに対話モードに移ることができ、表情認識などの精度にも繋がることとなる。

1.2 目的

本研究は対象人物がディスプレイ前に座っている状況を想定した上で、ディスプレイ光反射を利用した顔領域検出システムを提案し、背景や環境光に影響を受けない、対象人物のみを対象とした顔領域検出を行う。また、実時間処理に近付ける。

2 肌色情報を用いた顔検出

顔検出法は様々であり、その一つに肌色情報がある。肌色情報は顔を見つけるのにより手がかりとなる。

肌色には固定の色がなく、人それぞれに個人差があるのは明らかである。それは輝度、彩度の違いによるものである。これに対し、色相には個人差は少ない。肌色領域の色相のヒストグラムから色相の閾値を決定する。従来法は色相を用いることで環境光の明るさに左右されない肌色領域の検出が可能となるとしているが、次のような問題が残る。

- 環境光の色の変化に弱い
- 背景の肌色領域を検出してしまう

そこで、本研究ではディスプレイ光の映り込みを利用して、環境光に左右されない顔本来の固有色を検出し、同時に背景の肌色領域を分離する方法を提案する。

3 ディスプレイ光の利用

物体は光を反射、吸収する性質を持つ。我々は反射した光を視覚で感じ、色を認識している。すべての物体が全く同じ反射をするわけではなく、反射は物体の反射率、鏡面反射や拡散反射といった反射の種類、輝度や光源との距離などによって異なる。反射の度合は反射率、反射の種類などにも影響されるが、特に輝度と光源との距離に影響される。

図1のようにディスプレイの前に対象人物がいる状況で光を照射し、照射前と後の画像を撮る。その二枚の画像の各画素ごとのRGBの各成分の差をだす。その差のRGB値を新たな画像、差分画像として扱う。差分画像には雑音が目立つのでスムージングを行う必要がある。

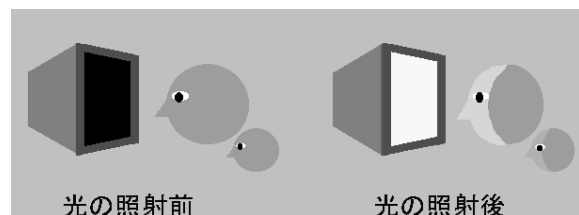


図 1: ディスプレイ光の映り込み

3.1 映り込み変化を用いた固有色の抽出

差分画像を用いて固有色を求める。

はじめに

対象人物の固有色 : (R_O, G_O, B_O)

環境光 : (R_I, G_I, B_I)

ディスプレイの照射光 : (R_L, G_L, B_L)

とする。光を照射する前の対象人物の色は

$$(R_O R_I, G_O G_I, B_O B_I)$$

で表され、光を照射した後の対象人物の色は

$$(R_O(R_I + R_L), G_O(G_I + G_L), B_O(B_I + B_L))$$

で表される。色の変化量は光の照射後と照射前の差であるから

$$(R_O R_L, G_O G_L, B_O B_L)$$

となる。またディスプレイの照射光は事前にディスプレイ光とカメラ色特性のキャリブレーションを行うことによって求めておくことができる。この二つの値から固有色を求めることができる。

$$\begin{aligned} R_O &= \frac{R_O R_L}{R_L} \\ G_O &= \frac{G_O G_L}{G_L} \\ B_O &= \frac{B_O B_L}{B_L} \end{aligned} \quad (1)$$

3.2 映り込み強度を用いた背景除去

顔領域を検出する際、背景に肌色の物体、つまり顔領域の候補が存在することを考慮しなければならない。対象人物以外の人がある場合、誤った検出をする恐れがある。その際、対象人物の顔であるかどうかを見分けるため、背景除去が必要となる。対象人物がディスプレイに最も近いと想定しているので、ディスプレイ光を照射すると、対象人物が最も強く光の影響を受けることになる。これは、光には照射地点から離れるにつれ弱まるという減衰の性質のためである。この光の減衰の性質を利用して、対象人物の顔領域のみを検出する。

この差がディスプレイ光による色の変化量であり、ディスプレイに近い物体と遠い物体では、この色の変化量に差が生じる。対象人物と背景物体の間

にある程度の距離があれば、色の変化量の大きなものは顔領域候補とし、変化量の小さなものを背景物体と見なして除去することができる。

4 顔領域検出アルゴリズム

1. 照射前の画像を撮影し、ノイズを除去する
2. ディスプレイ光を照射する
3. 照射後の画像を撮影し、ノイズを除去する
4. 二枚の画像の RGB 各成分の差を計算する
5. 色の変化量により背景を除去する
6. 環境光と照射光から固有色を求める
7. 固有色の色相を求める
8. 色相の閾値処理により顔を検出する

5 実験

5.1 ディスプレイ光とカメラ色特性のキャリブレーション

照射光のデータを得るために RGB 各原色の赤、緑、青と白について環境光の無い状況下でディスプレイ光の照射部分のみを撮影した。その際、RGB 値を 50、100、150、200、250 の五段階にした。50 の場合では、赤 (50, 0, 0)、緑 (0, 50, 0)、青 (0, 0, 50)、白 (50, 50, 50) というように設定した。撮影した画像の RGB 各成分の平均値をグラフ化すると、図 2 のようになった。

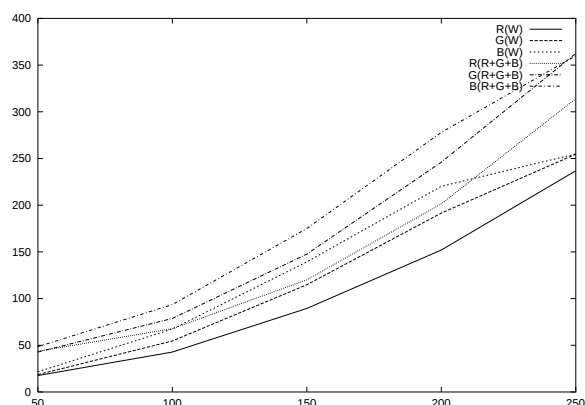


図 2: 白の平均値と三原色の平均値の和

ここで、ディスプレイ光の色が赤、緑、青のとき

光の色 (display)	レンズの絞り (iris)	成分 (RGB)	設定上での RGB 値に対する実際の RGB 値				
			50	100	150	200	250
白	暗	B	15.4852	39.905	84.5628	138.715	183.412
		G	14.1538	32.6849	69.3891	119.803	175.663
		R	14.2002	26.6349	53.6444	94.2293	147.245
	中	B	21.6818	67.392	139.316	220.171	254.862
		G	18.6716	54.5793	114.803	191.535	254.603
		R	17.6293	42.8026	89.4636	151.994	236.813
	明	B	32.7581	104.462	210.278	254.759	255
		G	27.1072	85.2642	177.931	250.133	255
		R	23.3807	66.4635	139.631	226.675	255

表 1: 白の場合の RGB 平均値

光の色 (display)	レンズの絞り (iris)	成分 (RGB)	実際の RGB 値 (0, 0, 0)
黒	暗	B	12.1616
		G	11.1778
		R	12.1651
	中	B	13.6845
		G	12.2803
		R	12.8345
	明	B	17.0137
		G	14.124
		R	14.2065

表 2: 黒の場合の RGB 平均値

の各成分の和と白の各成分が等しくなるはずが、グラフによると三色の和の方が値が大きくなっている。このことから設定上、純粋な原色にしてもビデオカメラで撮影すると他の成分を含むことが判った(表 2 参照)。また、設定上で黒 (0, 0, 0) にしていても、実際には RGB 各成分が微妙に含まれていることも判った。

ディスプレイ光の照射前とはディスプレイが黒のときで、照射後とはディスプレイを白にしたときである。黒のときを基準 (0, 0, 0) としたので、照射光は白の RGB 値から黒の RGB 値を引いたものである。肌色領域について、昼間、晴天、室内、白色蛍光灯の自然環境下で撮影した顔画像の色相のヒストグラムと式 (1) から求めた固有色での色相のヒストグラムを図 3、図 4 に示す。これにより顔の肌色領域の色相を $-330^{\circ} \sim 30^{\circ}$ までと求めた。

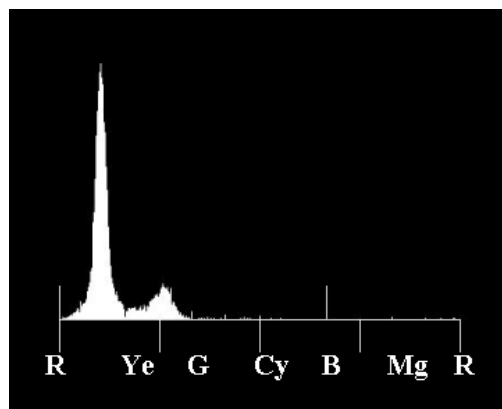


図 3: 自然環境下での色相のヒストグラム

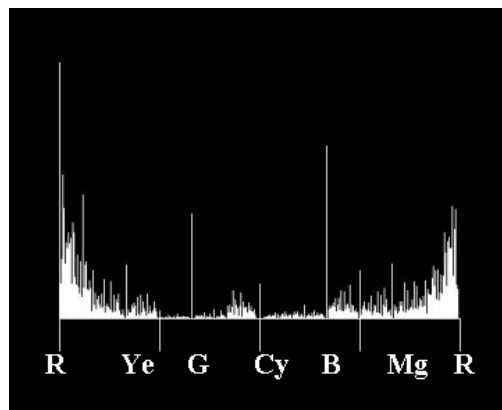


図 4: 固有色での色相のヒストグラム

5.2 実験の概要

入力画像はフルカラーで、画像サイズは 320×240 ピクセル、ディスプレイ光の照射は黒から白に変化させた。

使用器具について、ディスプレイは ViewSonic 社の ViewPanel VE150、ビデオカメラは SONY 社の XC-003、キャプチャボードは ITI 社の IC4RGB061 を使用した。

5.3 実験結果

環境光の明るさ、色、そして背景に影響されない検出ができた。また、CPU は PentiumIII 550MHz、メモリは 128Mb で、処理速度は 0.10 秒であった。

照射前と照射後の画像とその結果得られた画像を最後に示す。

6 結論

本研究では実時間顔領域検出システムを構築した。このシステムでは環境光の明るさ・色、背景の肌色物体の有無に関係なく正確な顔領域の検出が可能であるといえる。ゆえに、昼間にしろ夜間にしろ関係なく、またどのような照明にたいしても安定に顔領域の検出ができるということである。更に、背景に人や肌色物体が存在しても検出せずに、対象人物のみの顔領域の検出ができた。

しかし、ディスプレイ光の照射前に一枚目の画像を撮ってから、照射後の二枚目の画像を撮るまでの間に背景物体が大きく動いた部分を、肌色領域だと誤った検出をすることがあった。それは、背景物体の移動があった部分の差分値が肌色領域とする差分値の範囲に含まれていたためである。

6.1 ■ 今後の課題

- 背景物体の移動による誤検出のない検出法を検討する

その問題を解決する一例として、背景画像の移動によって生じる可能性のある色を予め調べておき、そのとき生じる可能性のない色の光を照射する。白に近い色は、光の色をそのまま返しやすいため、その部分を顔領域にあ

る白目や歯だと考え、その顔特徴に接している肌色領域の連結成分を顔領域とする方法も考えられる。

- 処理時間を向上させ実時間処理を目指す

処理速度は

1. プログラムの改良
2. MMX アーキテクチャによる並列処理化などの方法が考えられる。

参考文献

- [1] 高木 幹雄, 下田 陽久: “画像解析ハンドブック”, 東京大学出版会, 1991.
- [2] 杉本 岳夫, 徐剛: “カラー画像による実時間顔検出”, 情報処理学会研究報告, vol. 97-CVIM-107, No. 6, 1997.
- [3] 井上 誠喜, 八木 伸行, 林 正樹, 中須 英輔, 三谷 公二, 奥井 誠人: “C 言語で学ぶ実践画像処理”, オーム社, 1999.
- [4] 安居院 猛, 長尾 智晴: “C 言語による画像処理入門”, 昭晃堂, 2000.

照射前



照射後



結果 1



図 5.1: 実験結果 1

照射前



照射後



結果 2

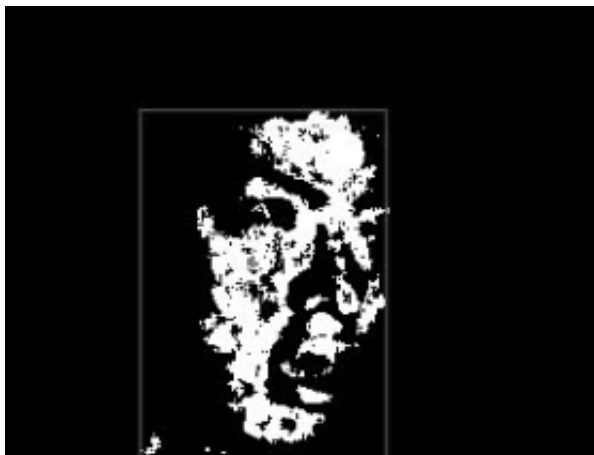


図 5.2: 実験結果 2