Webカメラの映像からの目検出ソフトの作成

４年Ａ組　稲益　秀成
指導教員　米田　隆恒

１．要約

私は、コンピュータに搭載されているWebカメラを用いて利用者を撮影し、その映像から利用者の顔、そして目線を検出する方法を研究した。

今回作成したソフトは、Webカメラで利用者を撮影し、その映像から顔の位置を検出し、そこから映像中の利用者の目の位置を検出することができるものである。

キーワード　目、目線、目線検出、Webカメラ、OpenCV、Processing

２．研究の背景と目的

人間は他人の目を見ると、その目線と目線の先にあるものから、その人が何を見ているのかをおおまかに知ることができる。そこで私は、人間の目は視覚情報を得るという感覚器官としての機能だけでなく、自分の視覚情報を他人に知らせるという情報伝達の役割も持っているのではないかと考えた。今日幅広く普及しているコンピュータやスマートフォンなどが人間の目線を情報として取得できれば、これらの電子機器の使い方が広がるのではないだろうか。例えばコンピュータが利用者の見ているものを検出できれば、そこから利用者の興味のあるものを予測する、といったことが考えられる。

そこで、コンピュータが利用者の目線を得る方法について考えていくことにした。

現時点では目線検出までは完成していないため、顔検出、目検出について報告する。

３．研究内容

コンピュータが利用者の目線を取得するために、Webカメラを用いた。

Webカメラとは、SkypeやFaceTimeなどといった、インターネットでテレビ電話を楽しむときに相手にリアルタイムで送信するための自分の顔の映像を撮影するカメラのことである。

図１　Webカメラ

図２　スマートフォン等のインカメラ

また、図２のように最近のスマートフォンや、タブレット端末などには同様な用途のカメラが標準で内蔵されていることが多い。たとえ、標準で搭載されていない場合や、搭載されているカメラが低性能であったとしても、Web カメラは複数のメーカーから多数販売されており、手軽に高性能なカメラを用意することができる。よって、他に目線検出専用の機器を追加するよりも Web カメラを利用したほうが、手軽に目線検出を行うことができると考え、Web カメラで目線検出に向けた目検出を行った。

Web カメラの映像から利用者の目線の動きを検出する流れとしては、

（A）Web カメラの映像から利用者の顔部分のみを切り出す。

（B）切り出した顔の範囲から目の位置を検出する。

（C）目の位置から、目の中心を計算し、その動きから目線を計算する。

といったものである。

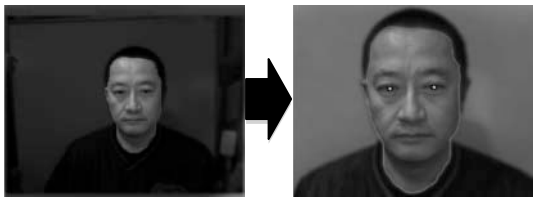
図３　目検出までの流れ

これまで、目の認識までは成功していたが、今年度は精度を高めるために、顔検出方法についてさらに深く研究した。

現在、目線計算の部分はまだ完成していないため、（B）までを報告する。

３．１　ノイズ除去

昨年度の時点で、目検出自体は安定したものが完成したが、その処理を行う顔検出が安定していないため、最終的な目検出の精度は十分なものではなかった。そこで、顔検出が安定しないのはカメラのノイズが原因だと考え、映像のノイズを除去する方法を調べた。そして、「メディアンフィルター」「平均値フィルター」という方法を検討した。

まず、「メディアンフィルター」とは、映像から、例えば３ピクセル×３ピクセルの大きさを切り出し、その範囲の９つのピクセルの濃度の中央値をその範囲の中心のピクセルの濃度にする、という方法である。「平均値フィルター」は、同様に９つのピクセルについて考え、その平均値をその範囲の中心のピクセルにする、というものであるが、今回のプログラムではこの２つの方法は使えないと考えた。なぜなら、もし色相の中央値をとってしまうと、肌の色相が大きく変化してしまう可能性が高いからである。そうすると、逆に肌の色の検出が安定しなくなってしまう。また、平均値をとることでは大きく色相の値が変化しないが、それでも色相が変化してしまい、肌の色だったものが肌の色と認識されなくなる可能性があるので、他の方法を考えた。

そして考えたノイズ除去方法は、まず肌の色が検出されたピクセルから左方向と上方向に、５ピクセルの範囲で、近いところから順番に肌の色のピクセルがあるかを探す。肌の色のピクセルが５ピクセル以内にあれば、最初の肌の色の点とその５ピクセル以内の点の間も肌の色にする。という方法である。

顔の映像にあるノイズを除去できればいいので、肌の色だけを見ていくことにした。この方法を使うことで、元々肌の色だったピクセルは変化せず、肌の色だと予想されるものは肌の色だと認識できるので、はじめの２つの方法に比べて良いと考えた。

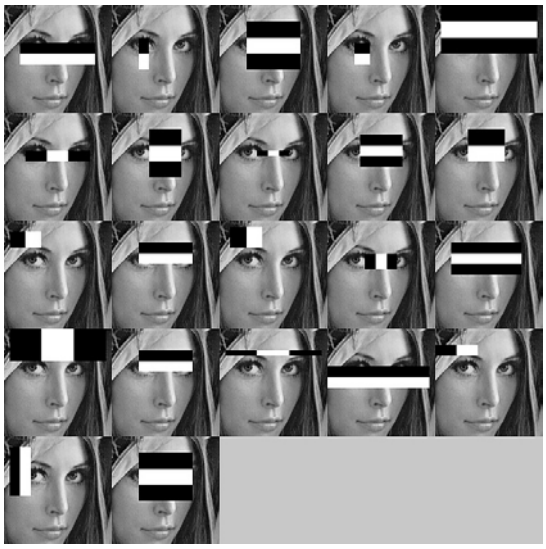

図４　顔検出におけるパターン認識の例
（参考文献[６]より引用）

結果

この方法では、ノイズが強い状態では顔以外の部分も肌の色に変更してしまったという問題点があった。また、リアルタイムで計算を行うには処理が重いものであった。

３．２　OpenCV を用いた顔検出

肌の色を探していく方法では、なかなか精度を出すのが難しかったため、今度は肌の色ではなく、パターン認識を使う方法を用いた。たとえば、鼻筋の方がその隣の目のある部分の方より手前にあり、光が当たりやすいため明るくなる、といったように顔の一部分と別の部分との色の違いの特徴をデータベースに持っておき、このデータをパターンとして映像の左上から１ピクセルごとにずらしながら、今調べている範囲の映像の特徴が、データベース上の顔の特徴と一致するかを調べていく方法である。

また、この顔の特徴のデータは１つひとつでの認識率は低いものだが、２つの部分の色の違いであるため、シンプルで計算量が少なくて済み、それらをたくさん用意して調べていくことで認識率を高めていくものである。

そこで、この顔検出を手軽に行うことができる、インテル株式会社が発表したオープンソースライブラリーである OpenCV を利用することにした。

この方法は点ではなく、部分と部分の色の差を用いるので、ノイズの影響を受けないという利点がある。

結果

OpenCV による顔検出の精度は高いものだったが、リアルタイムで目検出を行うには処理が重すぎるものであった。

しかしながら、この OpenCV による高精度かつ処理の重い方法と昨年度研究した、低精度かつ処理の軽い方法とを組み合わせれば今回の研究目標を達成するのに十分な性能の顔検出が行えると考えている。

４．考察

今回、コンピュータで利用者の目線を取得するために、Web カメラで利用者を撮影し、その映像から目の位置を検出する方法を研究した。

顔の範囲の検出は、まだまだ改善の余地があるが、目の位置の検出の処理は軽く、安定しており、顔検出さえ改善すれば目認識をリアルタイムで動かすことができると考えている。

今後は、肌の色を使う方法と OpenCV によるパターン認識を用いる方法をどのように組み合わせて行くか、考えて行きたい。

５．今後の展望

今回検出した目の中心から目線を推定し、それをコンピュータに活用させるようにしていくことが次の課題である。

そのためには、顔の向きや、顔と画面との位置関係を取得し、それらの情報から利用者の見ているものを計算していく必要がある。

これが完成すると、コンピュータは、利用者が何を見ているのかという情報を得ることができるようになる。これにより、例えばインターネットブラウザ利用者の目線から、その利用者の興味がどこにあるかがわかることになる。あるいはどのような内容が読んでもらえるかを今まで以上に調べることができるようになり、うまく使えば、Web サイトの質の改善につながると考えられる。

また、体の不自由な人でも目線の動きだけでコンピュータを操作できるようになる。例えば、画面に「あいうえお…」といったような文字の表を表示しておき、それを一定時間見ることでその文字を選択したことにすれば、文章入力はもちろんだが、その文字を音声で出力することで、体が不自由な人でも会話することができるようになる。このような社会貢献につながると考えられる。

６．参考文献

［１］ メディアンフィルター
「http://www.gifu-nct.ac.jp/elec/yamada/iwata/median/index.html」

［２］ Processing Reference
「http://processing.org/reference/」

［３］ OpenCV ¥ Library
「http://ubaa.net/shared/processing/opencv/」

［４］ コンピュータビジョンのセカイ - 今そこにあるミライ
「http://news.mynavi.jp/series/computer_vision/008/index.html」

［５］ Adam Harvey Explains Viola-Jones Face Detection
「http://makematics.com/research/viola-jones/」

［６］ OpenCV.jp
「http://opencv.jp/」

７．謝辞

今回の研究にあたり、サイエンス研究会物理班顧問の米田隆恒先生、藤野智美先生には、研究に関する助言や指導をしていただきました。また、同研究会のメンバーには研究のサポートをしていただきました。この場を借りてお礼申し上げます。