

顔方向から気持ちを表す動作の判定

Detection of Actions which Express Feelings
from Facial Orientation

篠田 裕人*, 山田 昌尚*, 土江田織枝*
Hiroto SHINODA*, Masanao YAMADA*, Orie DOEDA*
*釧路工業高等専門学校
*National Institute of Technology, Kushiro College

Email: yoshida@kushiro-ct.ac.jp

あらまし: 近年, インターネットを使ったオンライン授業や会議が増えている. 遠い場所にいる相手とも気軽に機会を作れるので便利な反面, 対面のときと比べると相手の様子や反応を直接感じることができないという難点がある. そこで本研究では, ウェブカメラの映像から得た情報を用いて, 顔の向きを取得し, そこから, 肯定や共感を表す「うなづく」動作と, 否定や共感をしていないことを表す「首を傾げる」動作を判断するシステムの開発を行った. 本稿ではシステムの概要や評価実験の結果について述べる.

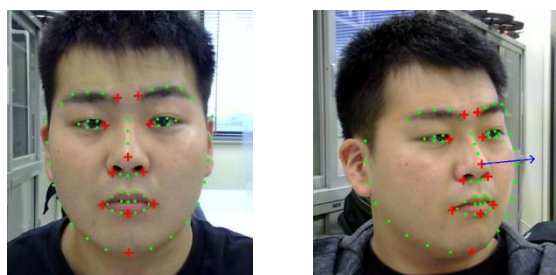
キーワード: 顔の特徴点, 顔の方向, 気持ちを表す動作

1. はじめに

近年, 感染症対策のためインターネットを通じたオンライン授業や会議が増えている. これらは, お互いが離れた場所においても気軽に機会を作れるので便利である. しかし, 対面で接するときのように相手の様子や反応を, 直接感じ取ることができないという難点がある. そこで表情や顔の向いている方向(以後, 顔方向と呼ぶ)などの情報を用いることで相手の反応を判断できないかと考えた. 本研究では, 顔の特徴点の情報を用いて取得した顔方向から, 肯定や共感を表す「うなづく」動作や, 否定や共感していないことを表す「首を傾げる」などの動作を判定するシステムの開発を目的とする.

2. システムの概要

システムは CPU i7 950, グラフィックボード GeForce 9800GT を搭載したパソコンと, 画素数 0.9MP 解像度 1280×720, リフレッシュレート 30fps 程度のウェブカメラ(以後, カメラと呼ぶ)で構成した. カメラはディスプレイの中央上部に設置する. 顔の特徴点を取りやすいように, カメラと顔の距離は大体 65cm から 100cm 以内とし, 顔全体がカメラの画面内に納まるようにする. 図 1 は 70cm 程度の距離に顔が位置した状態でシステムを使用したとき



(a) 特徴点を表示

(b) 顔方向を表示

図 1 特徴点と顔方向

の画像である. システムの使用中はマスクの着用はせず, 口はできるだけ開めた状態とする. カメラの映像から OpenCV のライブラリである Dlib により特徴点を検出^[1]し, その情報から顔方向を取得する. リアルタイムに処理を行うため, 特徴点の検出は回帰ツリー分析を用いている.

本システムで判断する動作は 2 種類とする. 「うなづく」動作をすることで, 理解している又は, 共感しているなどの肯定的な気持ちを表し, 「首を傾げる」動作をすることで, 理解していない又は, 共感していないなどの, 否定的な気持ちを表すことにした.

3. 顔方向の取得

顔方向は, 両眉の内側・両目の内側・鼻頂点・両小鼻の外側・口の両外側・唇の下・顎の先の 11 点の位置情報を用いる. Dlib により検出した 68 点の特徴点は図 1(a)の点と十字記号となっている. 十字記号は顔方向の取得に使う 11 点を示している. この 11 点の特徴点の値に対する 2 次元・3 次元の各座標上の値(Roll, Pitch, Yaw)と, カメラパラメータから求めた回転ベクトルと平行移動ベクトルで, 鼻頂点の向きを表すベクトルを取得する^[2]. ここで用いる 3 次元ベクトルの回転に使う Roll, Pitch, Yaw を図 2 に示す. 図 1(b)の鼻の先端の矢印記号が鼻頂点の向きを表しているが, 鼻頂点の向きと顔方向は一致すると

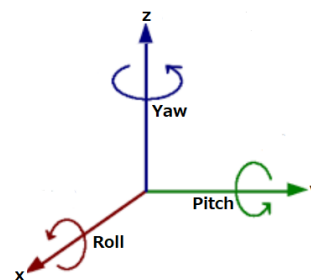


図 2 3次元のベクトルの回転

考えられるため、本システムでは、鼻頂点の向きを顔方向とする。

4. 動作の判定

4.1 動作の判定方法

「うなずく」と「首を傾げる」などの動作の判定は、3章で顔方向を取得する際に用いた Roll, Pitch, Yaw から Roll と Pitch の値を使って行う。顔の動きの判断は基準となる値である閾値を使うが、閾値に対して一定以上短い間隔で Pitch の値に変化があったときには「うなずく」動作をしたと判定する。また、「首を傾げる」の動作は、Roll の値の変化で判定する。

4.2 顔の位置の調整

図3の赤色の横軸は Pitch の現在の値を示し、縦軸は Roll の現在の値を示している。Dlib で取得した特徴点の鼻の頂点の位置が原点となり交差している。緑色の横軸は「うなずく」の判定に使う Pitch の閾値を示し、縦軸は「首を傾げる」の判定に使う Roll の閾値を示している。首の傾きが左右どちらでも対応できるように Roll の閾値は2つ設けている。図3(a)は顔が傾いているため、現在の Pitch と Roll の値と閾値とにずれが生じているので、赤色と緑色の直線がずれている。閾値を現在の位置に合わせるために、カメラ映像から100フレーム分の画像を取得し、それらの Pitch と Roll の値の平均値が閾値となるように調整する。調整後の状態を図(b)に示す。赤色の直線の両側に緑色の直線が移動していることが確認できる。この状態の Pitch と Roll の値を動作の判断に使う閾値とする。最初の閾値の調整は、システムを使用する前に、図4のように顔の位置がカメラの正面となるように座ってもらい約5秒程度、顔を静止した状態で行う。システムを使用中には常に閾値の調整を行っている。

5. 評価実験

本システムの動作の判定精度についての評価実験を行った。評価は、システムの目的を説明した後に、「うなずく」と「首を傾げる」の動作に対して、「浅く短く」・「浅く長く」・「深く短く」・「深く長く」の4種類の動きを意識して行ってもらった。各動作はそれぞれ5回ずつ行ってもらい、実際の回数とシス



(a) 閾値を調整中 (b) 閾値を調整後

図3 閾値を表示した状態

テムが判定した回数を比較することで、システムの動作の判定の精度の評価とした。被験者は21名とし、浅く、深く、短く、長くなどの動作の程度はとくに指示はしていない。評価実験の様子を図4に示す。図4の手前が被験者である。システムが判定した動作の回数は、ディスプレイの上側に表示し、確認できるようにになっている。

実験の結果では、「うなずく」動作は「浅く短く」は約39%が正しく動作の回数を判定できているものの、約33%が実際の半分以下の判定となっていた。「深く長く」は、約78%が正確にできており、それ以外についても1回程度の判定ミスであったことから、判定精度が高いことが確認できた。しかし、「首を傾げる」動作は被験者によって精度が大きく異なり、特に「深く長く」は実際の回数よりも多く検知することを確認した。これは、「首を傾げる」動作を「深く長く」行くと、傾げた状態から戻したときに若干反対側に顔が振れてしまうことが多く、左右どちらに傾けても判定ができるようにしていることで、1回の動作が左右両方で検知され、重複してカウントすることが原因だと考えられる。

6. まとめ

カメラの映像から顔の特徴点を用いて顔方向を求め、「うなずく」と「首を傾げる」動作について判定するシステムの開発を行なった。評価実験では「首を傾げる」動作については、短時間で傾きが左右両方で検知した場合には動作を無効にするなどの処理を加えることで、正しい判定が行えると考える。今後は、各動作を検知した回数などから、相手の反応を確認できる機能などを実装し、実用的に使えるシステムへと改良を進める予定である。

参考文献

- (1) Face landmark detection in a video, OpenCV, “https://docs.opencv.org/3.4.15/d8/d3c/tutorial_face_landmark_detection_in_video.html”, (参照 2021.11.04).
- (2) HeadPoseEstimationusing OpenCV and Dlib, Learn OpenCV, “<https://learnopencv.com/head-pose-estimation-usingopencv-and-dlib/>”, (参照 2021.11.04).



図4 評価実験の様子