

Kinect を用いた表情による入力インターフェイスの提案

井川 大輔^A 笹岡 久行^F旭川工業高等専門学校^A

1. はじめに

現在、インターネットを利用したチャットやメールなどを利用することで世界中の人とやり取りを容易に行うことができる。そして、顔文字を利用することにより、その時の感情を表現することで、相手に的確に伝わりやすくなる。しかし、顔文字は一つの感情でも多くの種類があり、個人の好み、感情の大きさなどで利用する顔文字が異なる。また、顔文字は予測変換に登録することで顔文字を簡単に利用することが可能だが、入力したい顔文字を候補の中から探す必要がある。

そこで本研究では、ユーザの表情およびその変化量を入力とし、顔文字の入力支援を行う手法の提案を行う。具体的には、Microsoft 社製の Kinect[1]を利用し、時間ごとの顔パーツの座標の変化量、顔の表情等を入力とし、表情と表情の大きさを推定し、これを基に利用者が利用したい顔文字を予測し、入力支援を行う。

2. Kinect SDK を用いた顔の認識処理

Kinect には、RGB カメラ、赤外線カメラの二種類のカメラが搭載されており、これにより画像と深度のデータを取得することが可能である。また、画像と深度のデータより骨格のデータを取得することも可能である。

そして、「Kinect SDK 1.5」以降より「Face Tracking SDK」を利用して顔の「目」、「鼻」、「口」などの位置を取得することができる [2]。これら以外にも、顔の向いている方向を取得することが可能となった。そのため、今回は「Face Tracking SDK」を利用することによって「口」などの位置データを取得し、この位置データをもとに表情を決定する。

3. 表情推定処理

表情には喜び、驚き、悲しみ、怒り、嫌う、等がある。本研究では文字入力に対する内省から、喜び（笑い）、驚き、困る（嫌悪）、無反応（退屈）の 4 種類の感情の判定を行こととした。判定方法は喜び、驚き、困る（嫌悪）のそれぞれの表情についての測定を行い、これらの表情がない場合においては無反応（退屈）として処理を行う。

また、それらの表情において、表情と顔文字の例をあげる。

• 喜び（笑い）

表情の例) 口が左右に広がる。
目が少し細めになる。

顔文字の例) $(\geq \nabla \leq)$
 $\smile (^ - ^) \smile$
 \wedge

• 驚き

表情の例) 目が上下に広がる。
口を開ける。

顔文字の例) $\Sigma (^ \pi ^ 111)$
 $(^ o ^ ; ;$
 $(*^o^*)$

• 困惑（嫌悪）

表情の例) 首をかしげる。
口が左右非対称。
眉間にしわがよる。

顔文字の例) $\text{♪} \sim (\overline{\varepsilon} ;)$
 $(-_-)$
 $(-_- ;)$

• 無反応（退屈）

表情の例) 顔に変化がない。

顔文字の例) $(^ \cdot \omega \cdot)$
 $(^ \circ ^)$
 $\wedge ;$

4.表情の大きさ

一人の一つの感情だけを取った場合でも、感情の大きさによって表情が異なる。特に、大きな感情である場合、表情のピーク値と表情の継続した時間が長くなる。そのため、表情の変化を取得するために、表情と認識を開始した時間と表情でないと認識した時間の差と表情と認識している間のピーク値の検出を行う。また、表情の大きさの平均値を測定することで、表情の大きさの誤差を小さくする。表情の時間、ピーク値、平均値を利用することで表情の大きさを測定する。

5.提案手法

図1のように、KinectのFace Trackingによってデータを取得し、そのデータをもとに、各表情大きさの測定を行う。その後認識した表情の候補から利用者が顔文字の選択を行う。

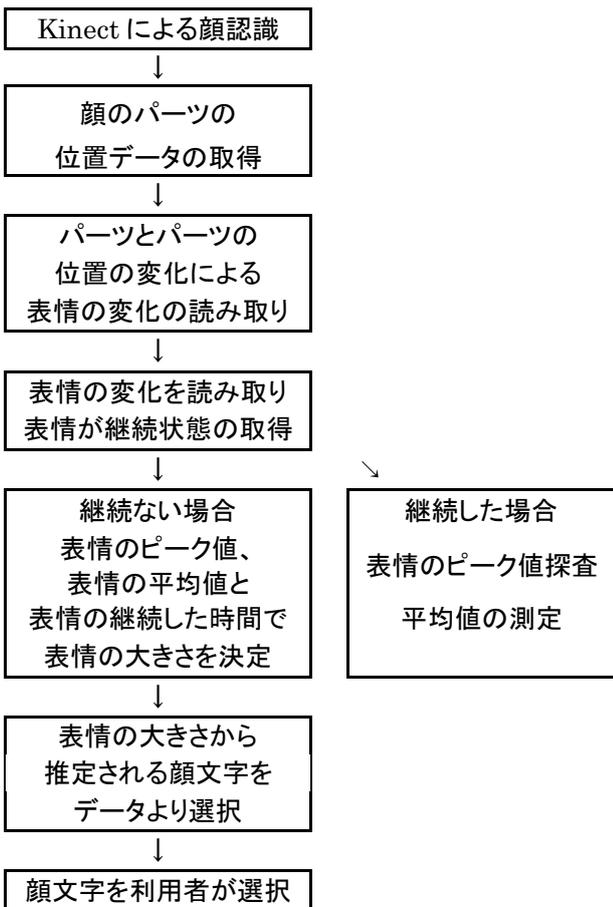


図1 提案手法の概略

6.処理例 (喜び)

図2のように変化する口の左右の開く大きさを測定し、開いた大きさが平均値よりも差がある場合に笑顔であると認識を行い、継続時間と平均値の測定を行い喜びの大きさを決定する

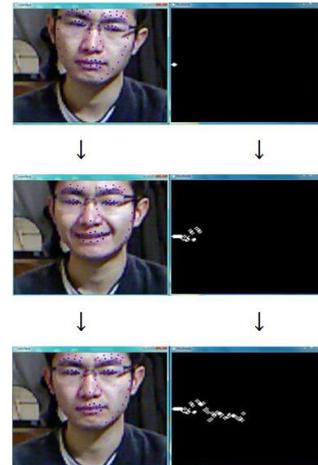


図2 処理の一例

7.おわりに

1つの表情に一つの部分観測を行い動作の確認を行っているため、表情の認識はうまくはいかない場合がある。

また、感情以外にも文章の終わりに個人の個性のために独自の顔文字を添える場合があるため、この場合においては文字予測が有効である。今回、Kinect SDKのFace Trackingを用いた提案手法により簡単に精度よい顔の認識し、顔のパーツの位置データを取得することができることを確認した。位置データの取得によって、口周辺、目周辺などの画像データより、画像処理を行うことでより表情の大きさの精度を良くすることが可能。

Kinectには、マルチアレイマイクを搭載されている。今後は、声の大きさ、言葉などからも表情の認識をすることも可能だと思われる。

参考文献

[1]Kinect for windows SDK
<http://www.microsoft.com/en-us/kinectforwindows/0>
 [2]杉浦 司,岩崎修介 “Kinect for windows SDK プログラミングガイド” ,工学社,2012年