

顔表情識別による聴講者の 評価取得システムの設計と実装

岩井 祐太†

西山 裕之†

† 東京理科大学理工学部

1 はじめに

学業や研究などの発表の場において、情報を相手に正確に伝えられたか、伝え方が効率的であったかの評価を得るためには発表の聴講者からのフィードバックが必要となる。聴講者からのフィードバックを得るための従来の手段として、発表直後に聴講者から質問を受け付けるか、聴講者に発表の評価のアンケートを行なう方法が一般的である。

しかし、こういった従来の手段には聴講者が発表者に対して直接アプローチを仕掛けて評価をフィードバックしなければならないという共通の問題点がある。この問題点を解決する方法として、発表中の聴講者の表情を伺い知る手段が挙げられる。顔の表情は視覚を通じて情報を伝えることができる [1] とされているが、発表中に発表者が聴講者の表情を随時伺うことは容易ではなく、評価データとしての蓄積もまた難しい。

そこで本研究では、発表中の聴講者の顔表情をカメラにて撮影して識別し、そこから聴講者の評価を可能にするためのシステムを設計し、実装する。本システムは、予め複数の顔情報を学習させることで顔表情の識別を可能にする。

2 顔表情識別システムの設計

始めに本システムの利用イメージを図 1 内の発表の場として示す。ユーザは聴講者の評価を取得したい発表に対し、聴講者に対してカメラを設置し録画を行なう。録画した映像データを本システムに送り画像を解析して表情を識別することにより、最終的に聴講者の評価を取得する。

次に本システムの設計を述べる。システムに送られた映像データを 1 フレームずつの画像として切り出し、下記の処理を行なう。

- 画像内からの顔領域の検出
- 顔領域内からの顔器官の検出
- 顔器官内の特徴抽出

上記の処理は、画像内の無駄な情報を排除して、顔表情の識別に必要な特徴のみを抽出し、学習データとしての標準化を行なうための設計であり、本システムでは画像解析として構成する。

次に、抽出された特徴に対して、その特徴がどのような表情の際に表されるのかの意味付けをしながら学習することが要求される。この学習手法は教師あり学習と呼ばれ、本システムでは教師あり学習を用いた Support

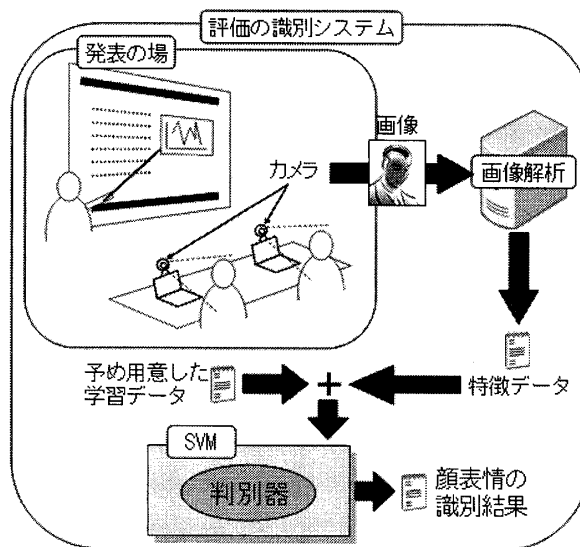


図 1: システム構成図

Vector Machine(以下 SVM) を用いる [2]。SVM は与えられたデータを二値分類する手法が一般的であるが、多値分類への拡張も可能である。そこで本システムでは複数の表情パターンを多値分類により識別可能にする。

以上により本システムでは、発表の場における聴講者の顔表情の録画、画像解析による顔表情の特徴点抽出、SVM による特徴点の学習および識別の 3 つの機能に構成される (図 1 参照)。

3 顔表情識別システムの実装

3.1 顔表情の録画

聴講者の顔表情を録画するにあたり、Web カメラによる固定点での撮影を行なう。なお、本システムでは 1 人の聴講者につき 1 台の Web カメラを用いることとし、画像の圧縮処理は画質の悪化につながるため無圧縮とし、解像度は Web カメラ固有の解像度を採用し、1 秒あたり 5 枚のフレームを取得し、その映像データを保存する。

3.2 顔表情の画像解析

本システムでは、画像解析としてコンピュータビジョンライブラリである OpenCV を用いる。初めに保存した映像データから 1 フレームを切り出し、Haar タイプの特徴量による手法 [3] を用いて 1 つの顔領域を検出する。検出された顔領域が 2 つ以上ある場合は各領域に対して HSV 表色系にて肌色抽出 [4] を行ない、肌色の

Design and implementation of the face expression identification system to enable the evaluation of the person of attendance.

Yuta Iwai†, Hiroyuki Nishiyama†

†Faculty of Sci. and Tech. Tokyo University of Science

割合が最も多かった領域を顔領域とした。また顔領域が検出されない場合はそのフレームを棄却し、次のフレームへと移行する。

顔領域が定まると次に顔器官を抽出する。抽出する器官は目と眉と口の3つとし、各器官の特徴である白色と黒色と赤色を YCbCr 表色系または RGB 表色系にて抽出する [5]。各器官の領域が抽出された後、それらの各特徴点の座標を抽出する。特徴点の位置は目と口は各領域の中心点と上下左右端とし、眉は前髪の影響を受け易いことを考慮して中心点のみとした。

ここで顔表情から各特徴点を抽出した表示例を図 2 に示す。各特徴点の座標を確認したところ、ほぼ同じような顔画像でも特徴点の座標が微妙に異なる揺らぎが確認されたため、左目と右目の各中心点の中間点を座標の原点とし、両目の中心点間の距離を基準とした相対的な長さ置き換えた座標の再計算を行なった。また、抽出に失敗したと思われる特徴点座標は特異点として棄却した。これらの画像解析により、顔の各器官の特徴点座標を出力する。

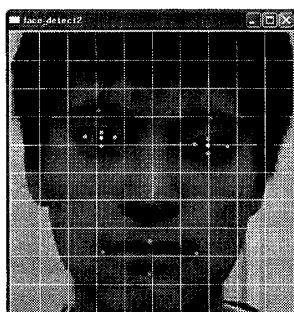


図 2: 顔器官の特徴点抽出の表示例

3.3 SVM による特徴点の学習および識別

本システムでは、学習手法として教師あり学習による SVM を利用し、多値分類での識別が可能な SVM-Multiclass を用いた。識別を行なうにあたって学習データが必要であるため、始めに学習を行なう。しかし学習を行なうには特徴点座標のデータと共に実際の評価が必要になるため、その際には聴講者の撮影時に発表の評価のアンケートを行なう。学習データを一度用意することによって、以後は特徴点座標のデータのみでその評価を識別することが可能である。

4 評価実験

本システムの評価実験として、4種類の顔表情パターンを想定し、学習のために各種類 100 枚ずつの顔画像を用意した。本実験で学習および識別に用いる顔画像のパラメータは、眉領域の中心点の高さ、目領域の上端と下端の高さの差異、口領域の上端と下端の高さの差異の3つとして、事前に学習を行なわせた。ここで4種類の顔表情パターンを図 3 に示す。また実験対象として、各種類 30 枚ずつを新たに用意し、それらを識別させた結果を表 1 に示す。

表 1 より、120 枚中 116 枚の正確な識別に成功した。これにより、聴講者の評価を正確に識別できると考えられる。しかし、撮影を行なう環境や聴講者によっては光量が足りなかったり個人差が大きかったりなどの理由で顔器官を判別できないことも多い。また、顔が横や下を向いてしまうと、評価は一切できなくなってしまふといった問題点もある。よって、本システムを利用する上では環境を限定する必要がある。

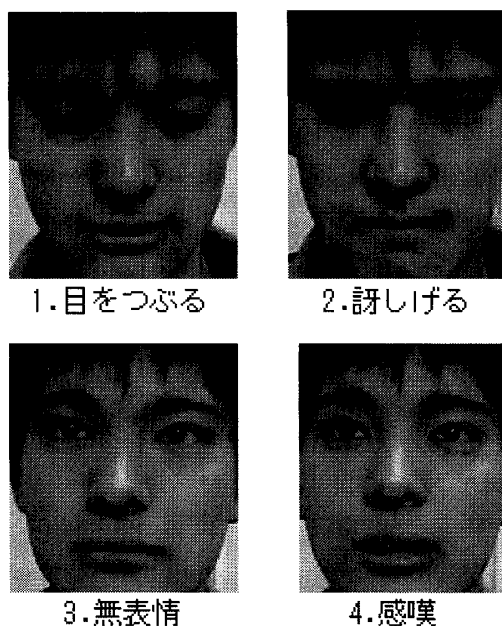


図 3: 4 種類の顔表情パターン

表 1: 評価の判別結果

	判別結果				
	1	2	3	4	正答率
1. 目をつぶる	30	0	0	0	100.0 %
2. 訝しげる	0	30	0	0	100.0 %
3. 無表情	0	4	26	0	86.7 %
4. 感嘆	0	0	0	30	100.0 %
合計	30	34	26	30	96.7 %

5 おわりに

本研究では、聴講者からの直接のアプローチを必要としない手法として、カメラから聴講者の表情を識別し、発表の評価を可能にするシステムの設計および実装を行なった。本研究には、人の表情を学習して理解し、識別するための基本的なシステムを備えている。これは、プレゼンテーションなどといった発表の場だけに本システムの利用が限定されないことを示している。例えば、広告や展示品等を見た人たちの反応やサービスシステムの利用者からのフィードバックを得るなどの応用も可能になると考えられる。

参考文献

- [1] 赤松 茂, "人間とコンピュータによる顔表情の認識", 電子情報通信学会誌, Vol.85, No.9, pp.680-685, 2002.
- [2] 山田寛康, 松本裕治, "Support Vector Machine の多値分類問題への運用法について", 情報処理学会研究報告, Vol.43, No.1, pp.44-53, 2002.
- [3] Paul Viola, Micael Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", IEEE CVPR, 2001.
- [4] A.R.Smith, "Color gamut transform pairs," Comput. Graphics 12, pp.12-19, 1978.
- [5] 叶 冠峰, "顔動画像からの特徴点抽出を用いた表情認識", 2005.