

単眼カメラを用いた教室内受講者の視線検出  
A Study on Sight Line Detection of Class Students using a Single Camera

○吉田功<sup>1</sup>, 清水雅夫<sup>2</sup>

\*Isao Yoshidai<sup>1</sup>, Masao Shimizui<sup>2</sup>

Abstract: The aim of our project is to develop a system that can detect the of class students using a single camera with an easy calibration, and presents the sight lines to the lecturer in real time to help to know the attention degree. In this study we investigate a horizontal sight line angle detection using a template matching technique. With multiple templates of specific sight line angles of an averaged eye image, our experimental results show a possibility to detect an inter-template-angel of sight line.

1. はじめに

授業中の教室にカメラを1台設置するだけで、授業を聴いている学生がどこを見ているかを先生が把握できるようなシステム構築を目標とする。教壇に画面を設置して、図1右のような見取り図を得ることが目標である。最新の視線検出方法<sup>[1]</sup>でも、かなり近い距離でしか視線の検出はできないが、本システムでは複数の手法を統合してロバストな視線検出を目標とする。

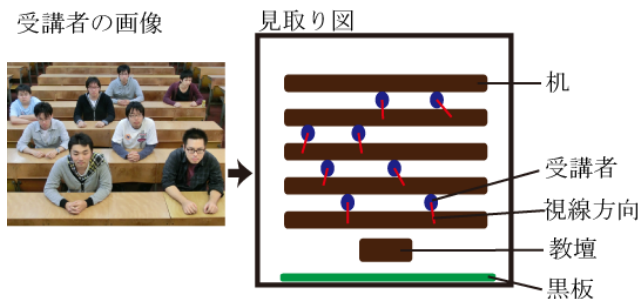


図1 監視モニターの例

本検討では、画像処理により人の視線方向を求める。最初から大きな教室で大人数の視線を検出することは困難なので、まずは少数での実験を行う。また、教室の最前列にいる人を仮定して、撮影するカメラから比較的近い距離に被験者を配置して実験を行い、視線検出方法の妥当性を検討した。

2. 原理

視線方向を検出する際に、撮影した画像に対して目の向きを検出しようとしても、目以外の部分がほとんどであるので検出対象を絞る必要がある。最初に撮影した画像から顔の部分だけを抽出する。次に抽出された顔の範囲に対して目の部分だけを抽出する。最後に視線の向きを求

めるようにした。以下に検出原理を示す。

2.1 顔と目の検出

顔検出と目検出にはHaar-like特徴分類手法<sup>[2]</sup>を用いる。人間の顔やロゴマークのような形状の検出に適している。事前に大量の顔及び目のサンプル画像に対して図2に示すような矩形の分類器の白と黒の領域間の明度差を特徴量として取得し、学習データの作成をしておく。次に、先ほどの分類器を検出対象画像と重ね合わせ、同様にして特徴量を計算する。事前に作成した学習データと近ければ顔および目と認識することができる。



図2 分類器の一例

2.2 テンプレートマッチング

テンプレートマッチングとは、あらかじめ用意された画像パターンが探索対象画像のどこにあるかを探す方法である。具体的には、テンプレートと探索対象画像の部分との間で類似度を計算し、最も類似度が高いとされる部分の位置を求める。類似度として次式の正規化相互相関を利用した<sup>[3]</sup>。

$$R_{X,Y} = \frac{\sum_{y=0}^{h-1} \sum_{x=0}^{w-1} (T_{x,y} - \langle T \rangle) (I_{X+x, Y+y} - \langle I_{X,Y} \rangle)}{\sqrt{\sum_{y=0}^{h-1} \sum_{x=0}^{w-1} (T_{x,y})^2 \sum_{y=0}^{h-1} \sum_{x=0}^{w-1} (I_{X+x, Y+y})^2}} \quad (1)$$

ただし、探索対象画像(X, Y)位置の演算結果を  $R_{X, Y}$ , テンプレートの幅と高さをそれぞれ  $w$  と  $h$ , テンプレートの位置(x, y)におけるピクセル値を  $T_{x, y}$ , 探索対象画像の(X+x, Y+y)におけるピクセル値を  $I_{X+x, Y+y}$  とする。また、 $\langle T \rangle$  と  $\langle I_{X, Y} \rangle$  は次式で求まる領域の平均値である。

1: 日大理工・院(前)・精機, 2: 日大理工・教員・精機

$$\langle T \rangle = \frac{1}{wh} \sum_{j=0}^{h-1} \sum_{i=0}^{w-1} T_{i,j}$$

$$\langle I_{X,Y} \rangle = \frac{1}{wh} \sum_{j=0}^{h-1} \sum_{i=0}^{w-1} I_{X+i,Y+j}$$

次にテンプレートの作成について述べる。被験者1人がカメラから1050mm離れた場所に座り、左側40度方向から右側40度方向へ10度刻みに視線だけを移動させる。その際に顔は正面を向いたままとする。それぞれの角度ごとに上半身を撮影する。撮影した画像から右目と左目を切り出す。被験者は7人、1人当たり3回撮影を行ったので、右目と左目のデータはそれぞれ21枚になる。それぞれの目の画像において、各画素位置の平均画素値を求め、テンプレートとする。テンプレートとなる平均眼を図3に示す。この平均眼は画像左側の眼(実際は右目)である。以後、負の角度は右側、正の角度は左側を示すものとする。

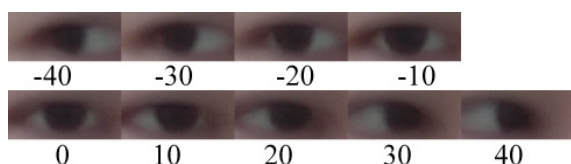


図3 被験者7人の画面左目平均眼テンプレート

### 3. 実験結果

#### 3.1 実験方法

実験で撮影に用いたカメラはCASIO製EX-ZR100である。本研究では1人撮影される状況で1人当たり3回、被験者は7人の実験を行った。被験者は壁から1850mm離れた場所に椅子に座った。カメラは壁から800mm離れた場所に置き、カメラと被験者の顔が一直線になるように配置した。被験者は左側40度方向から右側40度方向へ10度刻みに向き、顔は正面を向いたままにして撮影した。ここで、被験者が向くべき位置が分かるように壁にシールを貼った。最後に、撮影した画像から顔を抽出し、抽出された顔から目を抽出した。抽出された目画像が左側40度から右側40度のどのテンプレートと最も類似度が高いかを計算し、最も高かったテンプレートの角度を視線方向とした。

#### 3.2 実験結果

図4はHaar-likeによる顔・目検出及びテンプレートマッチング結果の1つを抜き出したものである。図4下のグラフは横軸に各テンプレートの角度を示し、縦軸に各々の類似度を示している。緑色で示した部分が類似度最大を示している。図4の場合は左目、右目共に-30度で類似度最大になっているのが分かる。

図5にテンプレートマッチングを行った結果を示す。横軸には被験者が実際に視線を向けた角度を表し、縦軸はテンプレートマッチングを行い最も類似度の高かった値を示している。データは7人分で1人当たり3回測定を行い、右目と左目があるので、1つの角度ごとに42のデータが含まれている。赤四角の範囲に測定した値の50%が含まれており、赤四角の中になる黒線は中央値を表す。

-40度から20度の間において、実際に向いた視線角度とテンプレートマッチング結果の中央値が一致している。また、赤四角部分に注目すると、ばらつきも少なく各角度の範囲が独立している。-40度から20度の間では、10度刻みであれば視線検出が可能である。

30度と40度においては、その他の角度と比べて明らかにばらつきがあり、視線検出が不可能という結果になった。

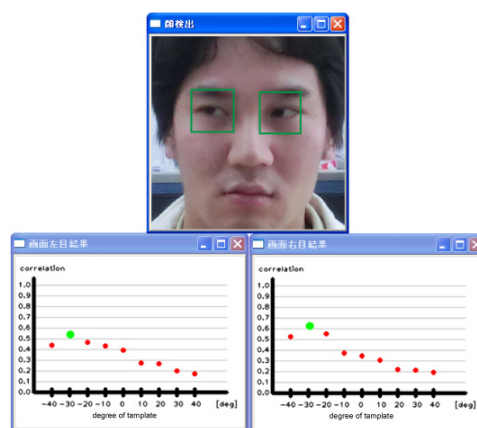


図4 テンプレートマッチング結果の一例

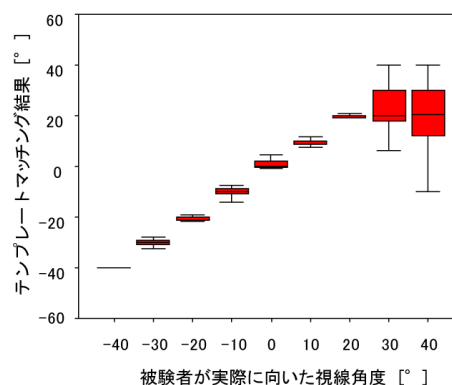


図5 テンプレートマッチング結果

### 4. 参考文献

[1] 入江淳, “高速な視線および目開閉推定による人のモニタリング”, SSII2011 予稿集(CD-ROM), 2011.  
 [2] Paul Viola and Michael J. Jones : "Rapid Object Detection using a Boosted Cascade of Simple Features", IEEE CVPR, Vol.1, pp511-518, 2001.  
 [3] 永田雅人: “実践 OpenCV”, カットシステム, 2009.