

## カメラを用いたタイピング練習システムにおける目の向きの認識

荒井正之, 下境浩<sup>†</sup>, 渡辺博芳, 武井恵雄  
帝京大学理工学部, <sup>†</sup>帝京大学大学院理工学研究科

カメラを用いてタイピング練習者を撮ることにより,タイピング練習を支援するための目の向き認識について研究を行っている.カメラを用いて,練習者の目の向きを認識することにより,練習者にアドバイスをすることなどが可能になる.これは,習得時間短縮などに効果があると考えられる.本論文では,タイピング練習者の目の向きの認識のための一手法を提案する.提案手法は,まずタイピング練習者の正面画像からモザイクテンプレートを用いて顔面部を抽出,次に黒目の分離フィルタを用いて,抽出した領域内を走査し,分離度によって目の向きの認識を行う.タイピング練習者 20 名に対する実験の結果,認識率は 83%となった.

### Recognition of Eye Direction Using a Camera for Touch-type Training System

Masayuki Arai, Hiroshi Shimosakai\*, Hiroyoshi Watanabe and Shigeo Takei  
*School of Science and Engineering, Teikyo University.*

*\*Master's program of Science and Engineering, Teikyo University.*

We study recognition of a trainee's eye direction using a camera to support touch-type training. The recognition enables us to give a trainee warnings and appropriate advice, and these warnings and advice assist a trainee to learn touch-typing quickly. We propose a method to extract the face area from a trainee's picture using mosaic templates and recognize a trainee's eye direction using separability filters. On the experimental results for twenty trainees' black-and-white pictures, we obtained an 83% recognition rate.

#### 1. はじめに

我々はカメラを用いてタイピング練習者を撮ることにより,タイピング練習を支援するシステムのための目の向き認識技術について検討を行っている.カメラを用いて練習者を撮り,練習者の目の向きを認識することにより,次のようなことが可能になると考える.(1)目の動きと打鍵データのログを分析することにより,練習者へ適切なアドバイスを行うこと.(2)リアルタイムで目の動きを認識することにより,練習法で禁止されている場所を見たときに警告を行うこと.(3)多くの練習者のデータを分析することにより,目の動きという視点から,タイピング練習者のモデルを確立すること.(1)(2)が可能になれば,練習者の習熟時間の短縮に,(3)が可能になれば,タイピング練習法の評価や練習者にあった練習方法の提示などに寄与できる.

計算はもとより読み書きなどもコンピュータで行う時代となった。パーソナルコンピュータなどの情報機器の使用は、キーボードによる文字入力を中心である。キーボードは、使うのが比較的易しいため、練習をせずに、独自の方法で入力を行う場合が多い。これには、ある程度の速さまで打鍵が可能になるが、それ以上速く打鍵できない、応用がきかないなどの問題がある。

このような状況から、タッチタイピング練習に関する研究が行われている[1~4]。これらの研究は、練習方法の提案と練習モデルの提案の2つに大別できる。前者の研究として次が挙げられる。大岩[1]は「文字と指の動きを関連づけること、タイピングミスの指摘は練習後に行うことが、練習のポイントである」と述べて、これらの条件を満足する練習法を提案している。増田[2]は「文字を条件反射的に正確に打つには、勝手に記憶が残ることが重要」と述べて、関連して記憶する練習方法を提案している。川島[3]はキーボードにカバーをして、練習者からキーボードを見えなくすることにより、タイピングの習熟速度が速まることを実験的に確かめている。練習モデルの提案の研究として、吉長ら[4]の研究がある。この研究では、タッチタイピングの習熟特性や習熟パターンを抽出する方法を提案している。

本研究の目的は、カメラを用いてタイピング練習者を撮ることにより、タイピング練習を支援するシステムに利用するための目の向きの認識技術を開発することである。本論文では、練習者の顔面部を抽出し、その領域に対して黒目の分離フィルタを用いて走査し、その分離度によって目の向きを認識する手法を提案する。

## 2. カメラを用いたタッチタイピング練習システム

### 2.1 想定するシステム構成

想定するタッチタイピング練習システムの構成は次の通りである。

#### ハードウェア

パーソナルコンピュータと市販の CCD カメラで構成する。CCD カメラはディスプレイ上の中央に置くものとする。

#### ソフトウェア

タッチタイピング練習と練習を支援する機能を持つ。練習支援機能には(a)本研究で提案する技術を利用して、練習者の目の向きの認識と分析を行う機能、(b)打鍵データのログの収集と分析を行う機能、(c)練習者へのアドバイス機能などを含む。

### 2.2 目の向き認識技術の利用例

タイピング練習に対して、目の向きの認識技術を用いることにより、次にあげることが可能になると考える。

#### (1) タイピング練習者へのアドバイス

種々のタイピング練習法に対して、どのような支援が可能になるか、具体例をあげて説明する。

大岩式練習法[1]では、文字と指の動きを関連づけること、タイピングミスの指摘は練習後に行うことなどを挙げている。練習者の目の向きを認識することにより、練習者がキーボードを見ている文字、すなわちタッチタイプができない文字が明らかになる。これにより、文字と指の動きを関連づけるために、どの文字を重点的に練習しなければならないかなどを、練習後にアドバイスすることが可能となる。

増田式練習法[2]では、ディスプレイを見て、打鍵をすることを奨めている。練習者の目の向きを認識することにより、打鍵時にディスプレイ以外の場所を見たときに、警告を発すること、練習後に目の動きと打鍵データのログを分析することにより、適切なアドバイスを行うことなどが可能となる。

川島[3]は練習者からキーボードを見えなくすることにより、習熟速度が速まることを実験的に確かめている。キーボードを物理的に見えなくするのではなく、目の向きを認識することにより、練習者がキーボードを見たときに、警告を発することが可能となる。

## (2) タイピング練習者モデルの研究

多くの練習者の目の動きを分析することにより、練習者の目の動きという視点から、タイピング練習者のモデルが研究できる。そのようなモデルの確立により、タイピング練習法の評価や練習者に合った練習方法の提示などが可能になる。

### 3. 目の向き認識の処理手順

本研究で提案する目の向き認識手法の処理手順と内容は次の通りである。

#### (1) 顔面部の抽出

練習の前に、練習者に一度ディスプレイを見てもらい、ディスプレイ上の中央に設置した CCD カメラから練習者の頭部または上半身の画像を入力する。その画像から、顔面部を切り出す。顔面部の切り出しは、最初に一度だけ行い、以後この領域内で(2)の処理を行う。

#### (2) 特徴抽出と目の向きの認識

抽出した顔面部から特徴抽出を行い、目の向きを認識する。

(1)(2)の処理について、次章以降で詳しく述べる。

### 4. 顔面部の抽出

本章では、モザイクテンプレート[8]を用いた顔面部の抽出方法について述べる。

#### 4.1 モザイクテンプレートの作成方法

モザイクテンプレートの作成は以下の手順で行う。

(1) 手操作により練習者の画像から、頭部、顔面部を切り出す。

(2) 切り出した頭部、顔面部にモザイク処理を行う。

(3) K平均クラスタリングにより、各部 n 個のテンプレートを作成する。

#### 4.2 テンプレートマッチング

モザイクテンプレートと練習者の画像とのマッチングについて述べる。タイピング練習者の上半身画像に対して図 1 に示すように  $W_x \times W_y$  画素のブロック単位でモザイク処理を行った画像と、頭部モザイクテンプレートとのマッチングを行い、頭部を抽出する。同様に、顔面部モザイクテンプレートを用いてマッチングを行い、顔面部を抽出する。

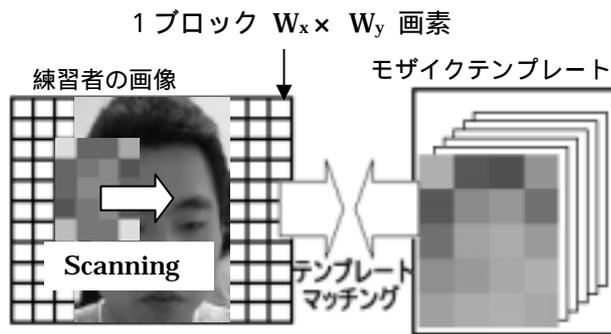


図1 モザイクテンプレートとのマッチング

## 5. 目の向き認識

目の向き認識は次に示す手順で行う。

### 5.1 分離フィルタによる黒目候補の抽出

抽出した顔面部に対して、分離フィルタを用いて、黒目の候補を抽出する。文献[7][9]では、黒目の分離フィルタの形状を真円としているのに対して、本研究ではディスプレイを見ている眼の形状に近づけるために、図2のような形状とする。このフィルタの位置や大きさを変えながら式(1)に示す分離度  $S(0.0 < S < 1.0)$  の高いものを黒目の候補とする。さらに、黒目を抽出した場合、中心が黒、周りが白になるので、図3に示すように分離フィルタの中心の平均濃度と、分離フィルタの周りの平均濃度の差が以上  $(|\bar{P}_I - \bar{P}_O|)$  のもの  $m$  個を黒目の候補とする。

$$S = \frac{n_1(\bar{P}_1 - \bar{P}_m)^2 + n_2(\bar{P}_2 - \bar{P}_m)^2}{\sum_{k=1}^N (\bar{P}_k - \bar{P}_m)^2} \quad (1)$$

ここで、 $N$ は領域内の全画素数、 $n_1, n_2$ は領域1、領域2の画素数、 $\bar{P}_k$ は位置  $k$ における濃度、 $\bar{P}_m$ は領域全体での濃度の平均値  $\bar{P}_1, \bar{P}_2$ は、それぞれ領域1、領域2での濃度の平均値を示す。

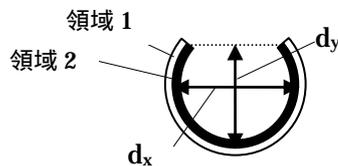


図2 分離フィルタ

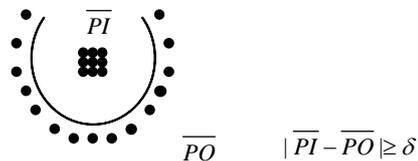


図3 濃度差による黒目候補の絞込み

### 5.2 黒目の同定

前節で抽出した黒目の候補の中から両黒目を同定する。図4に示すように、右黒目の中心の座標を  $(r_x, r_y)$ 、左黒目の中心の座標を  $(l_x, l_y)$  とすると、 $x_{max} > |r_x - l_x| > x_{min}$  かつ  $|r_y - l_y| < y$  を満足する2点を両黒目の中心とする。候補が見つからない場合は、キーボードを見ていると認識する。

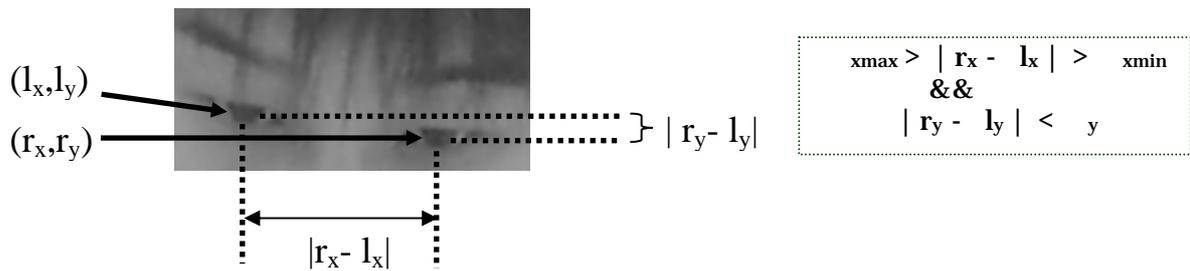


図 4 両黒目の同定

### 5.3 目の向き認識

練習者がディスプレイを見ている状態では、黒目がカメラに写るため、高い分離度で抽出されることが予想される。また、練習者がキーボードを見ている状態では、黒目はまぶたが覆うような形になり、カメラに写らない場合が多いため、分離フィルタにより候補が抽出されない、もしくは低い分離度で抽出されることが予想される。そこで図 5 に示すように、分離度が閾値 以上ならばディスプレイを見ている、以下ならばキーボードを見ていると認識する。

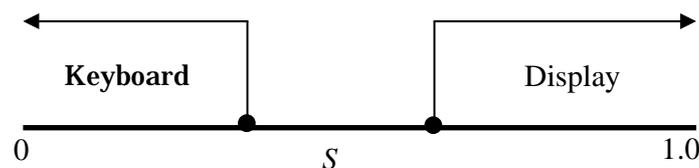


図 5 分離度による目の向き認識

## 6. 実験

タイピング練習者 20 名に対して認識実験を行った。練習者には、まず練習の前にディスプレイを見てもらい、そのときに撮った静止画像をモザイクテンプレートの作成に用いた。次に、タッチタイピング練習ソフトを用いて、実際に練習を行っている練習者の白黒 256 階調の濃淡動画を 1 分 30 秒撮った。全員眼鏡はかけてなく、内 15 名はタッチタイピングが行えない。動画を調べた結果、練習者はディスプレイを見ているか、キーボードを見ているかどうかであり、他の方向を見ている練習者はいないことが分かった。そこで、目の向き認識実験では、動画から静止画像に変換した練習者のディスプレイを見ている静止画像 20 枚と、タッチタイピングのできない練習者のキーボードを見ている静止画像 15 枚を用いて、leave-one-out により、2 つのクラスに分類する実験を行った。

なお、実験は CPU 1.7GHz の PC を用いて行った。

### 6.1 顔面中心部の実験条件

各部のモザイクテンプレートの作成条件は表 1 の通りとし、テンプレートマッチングで練習者の画像に対してモザイク処理を行う時の 1 ブロックの大きさ ( $W_x \times W_y$ ) は、頭部が  $75 \times 75$  画素、顔面部が  $25 \times 25$  画素とした。

表1 モザイクテンプレート作成条件

	モザイク処理	テンプレート数(n)
頭 部	横 4 ブロック×縦 5 ブロック	6
顔面中心部	横 12×縦 12	6

## 6.2 分離フィルタの実験条件

分離フィルタの大きさ(縦( $d_x$ ):横( $d_y$ ))は 20 画素 : 12 画素 , 24 画素 : 13 画素 , 26 画素 : 14 画素の 3 種類を使用した . その他のパラメータは ,  $m=20$  ,  $x_{max}=195$  ,  $x_{min}=100$  ,  $y=20$  , ( $=$ ) $=0.2$  ,  $=10$  とした .

## 6.3 実験結果と考察

まず , モザイクテンプレートを用いて , 顔面部の抽出を行った . すべての練習者の顔面部を抽出することができた . 1 人練習者の顔面部の抽出に要する時間は , 1 秒未満であった .

次に , 分離フィルタによる目の向き認識を行うため , 黒目候補の抽出を行った . 抽出率は真円の 55% に対して , 今回提案した分離フィルタでは 95% になった . 認識結果を表 2 に示す . 認識率はディスプレイを見ているクラスが 85% , キーボードを見ているクラスが 80% となった . キーボードを見ているクラスで正認識となった画像の一例を図 6 に示す .

ディスプレイのクラスで誤認識となった 3 データは , 練習者の目が細いなどの理由で , 現在使用している分離フィルタの形と大きさでは分離度が低くなったため , 誤認識となった . キーボードのクラスで誤認識となった 3 データは , 背景などに黒目に似た丸いものが存在し , それらを高い分離度で抽出したため , 誤認識となった . 1 枚の画像を認識するのに要する時間は約 1 秒であった .

表2 目の向き認識率

クラス	データ数	認識率
ディスプレイ	20	85%
キーボード	15	80%

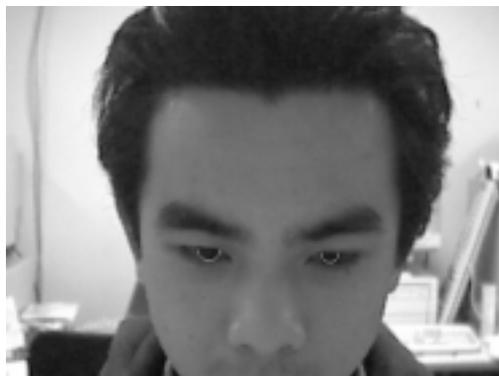


図6 ディスプレイを見ているクラスで正認識となった顔面部画像の例

## 7. おわりに

タイピング練習者の目の向きを認識するため、モザイクテンプレートによる顔面部の抽出手法と、分離フィルタによる目の向きの認識手法を提案した。実験の結果、顔面部の抽出率は100%、分離フィルタによる目の向きの認識率は83%であった。今後の課題として(1)分離フィルタの大きさと形の検討、(2)眼鏡をかけた練習者への対応、(3)まばたきの対応などがある。

## 謝辞

本研究の一部は科学研究費補助金13680253の補助による。

## 参考文献

- [1] 大岩 元：情報教育におけるキーボード，情報処理学会研究報告，Vol.CE-44-3, pp.11-18,(1997).
- [2] 増田 忠：キーボードを3時間でマスターする方法，日本経済新聞社，(1994).
- [3] 川島大司：ブラインド・タッチの学習方法について，情報処理学会第39回全国大会，4Y-8，(1988).
- [4] 吉長祐司，川畑洋昭：“情報教育におけるキーボードリテラシーの一考察”，情報処理学会論文誌，Vol.42，No.9，pp.2359-2367 (2001).
- [5] Arai,M., Watanabe,H. Oguri,K. and Takei,S. : Localization of a Feature Extraction Area for Touch-type Training Using a Camera, Proc. of The 8th International Conference on Computers in Education, Vol.2, pp.1544-1547 (2000).
- [6] 小栗賢志，荒井正之，渡辺博芳，武井恵雄：タイピング練習における目の向き認識のための特徴抽出法の検討，情報処理学会第60回全国大会，2ZB-1，pp.2-293 - 294，(2000).
- [7] 小栗賢志，荒井正之，渡辺博芳，武井恵雄：タイピング練習における目の向き認識のための特徴の検討，情報処理学会第62回全国大会，5L-02(2001).
- [8] 小杉 信：個人識別のための多重ピラミッドを用いたシーン中の顔の探索・位置決め，信学論 D- ，Vol.J77-D- ，No.4，pp.672-681(1994).
- [9] 山口 修，福井 和広：動画像を用いたPC顔認識システム”Smartface”，情報処理学会研究報告.CVIM，Vol.2000，No.50，pp.99-106(2000).