

## 鋼材に印字された低画質文字列の認識

非会員 澤 和 弘 (三重大学)  
 正 員 大 山 航 (三重大学)  
 非会員 若 林 哲 史 (三重大学)  
 非会員 木 村 文 隆 (三重大学)

## Recognition of Degraded Character String on Steel Plate

Sawa Kazuhiro, Non-member, Ohyama Wataru, Member, Wakabayashi Tetsushi, Non-member, Kimura Fumitaka, Non-member (Mie University)

In this paper, we propose a new recognition procedure for a degraded character string printed on a steel plate for factory automation. The character string consists of alphanumerals and ' '. The procedure includes the image emphasis with Laplacian Gaussian filter, the extraction of the character string image, the segmentation-recognition with the dynamic programming, and the high accuracy character recognition. We evaluated the accuracy for 1036 images (8,806 characters) scanned by the monochrome video camera at an actual steel producing factory line. The average recognition rate was 99.2% for the character recognition, and 91.6% for the character string recognition.

キーワード : 文字切り出し, DP マッチング, 文字列認識, ラプラシアン・ガウシアンフィルタ

## 1. ま え が き

FA(Factory Automation)の普及に伴って生産工程における生産指示や生産管理の手段として文字認識技術が用いられるようになった。スプレー等によるマーキング文字は、バーコード等の符号に比べ人間には読みやすいが、機械認識においては、かすれや照明むら、インクむら、背景色のむらなどのために、二値化や文字の切り出しが困難である。このような低画質文字列認識においては、文字と背景を正確に分離するため、文字の線幅に関する閾値特性を考慮する二値化法<sup>(1)(2)</sup>や、弾性エッジ列整合法を用いる認識手法<sup>(3)</sup>などが提案されている。本論文では、ノイズや背景むらの激しい低品質の入力画像に対して、ラプラシアン・ガウシアン処理による文字画像の強調を行って文字列を切り出した後に、動的計画法による文字の切り出し・認識を行う手法を提案し、有効性を実験的に評価する。実験に用いる画像は、鋼材の製造工程における品質管理のために、生産された鋼材の一部を切り出した分析用サンプル鉄片に直接マーキングされた文字列をVTRにて撮影したものである。2.に背景の濃度むらを除き、濃度変化の緩やかなエッジを正確に求める前処理と二値化法について、3.に文字切り出しと認識について、4.にこの手法による実験結果とその有効性について述べる。

## 2. 前処理と二値化法

〈2・1〉 画像の強調 提案手法による認識処理の流れを図1に示す。まず、ラプラシアン・ガウシアンフィルタを用いて文字画像の強調を行い、半二値化により背景ノイズを除去する。得られた画像に対して高い閾値 $t_1$ で二値化し、文字列領域の位置と大きさを推定し、文字列の切り出しを行う。次に強調画像を閾値 $t_1$ より低い閾値 $t_2$ で二値化する。その後、動的計画法により文字認識の結果を利用しながら1文字ずつ切り出す。この文字認識では高速処理が要求されるので64次元の特徴ベクトルを用いる。最後に、切り出された文字を400次元の特徴ベクトルを用いて再認識し、認識精度を向上させる。入力画像は、5mm×5mm×30mmの鉄片にマーキングされた文字列を白黒カメラによって撮影し、その画像を256階調、640×100ドットのサイズでコンピュータに取り込んだものである。カメラで取り込んだ入力画像の例を図2に示す。これらの例から、対象とする画像は背景の濃度むらと文字輪郭線のボケが著しいことがわかる。このような、背景の濃度むらを除き、濃度変化の緩やかなエッジを正確に求めるためには、ラプラシアン・ガウシアンフィルタが適している<sup>(4)</sup>。図3に、ラプラシアン・ガウシアン処理によって背景むらとボケのある画像を正しく二値化する一次元の例を示す。入力画

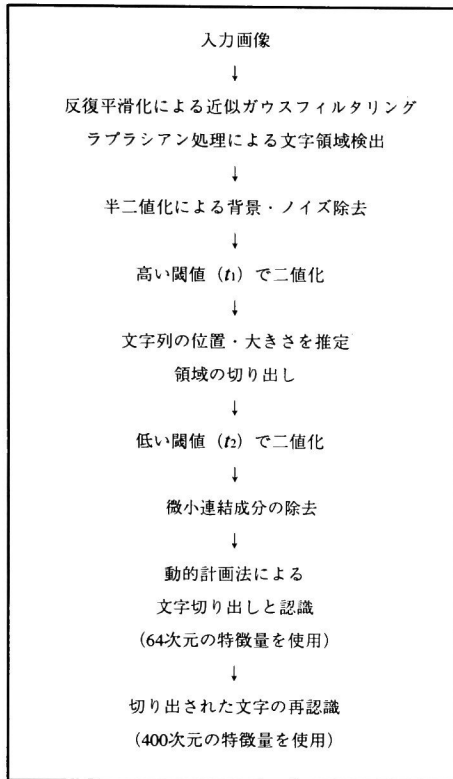


図1 認識処理の流れ

Fig. 1. Flow chart of character string recognition.

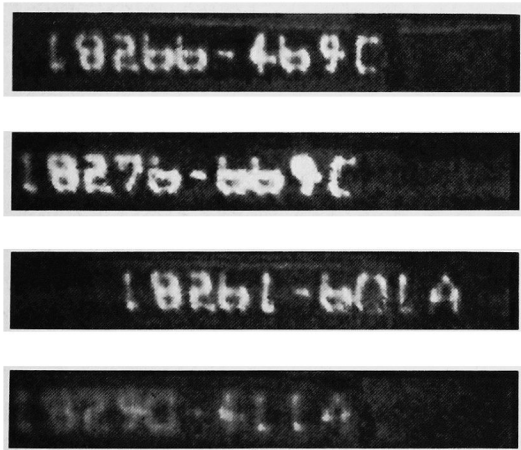


図2 鉄片に書かれた入力画像

Fig. 2. Input string images on steel plate.

像の濃度変化の変曲点を文字と背景の境界と考え、ラプラシアンフィルタの出力の零交差点（符号の変化する点）が境界となるように文字領域を検出する。この過程において、高周波成分に含まれるノイズはガウスフィルタにより除かれ、低周波成分に含まれる背景むらはラプラシアンによって除かれる。なお、1次元の場合は、ラプラシアンと2階

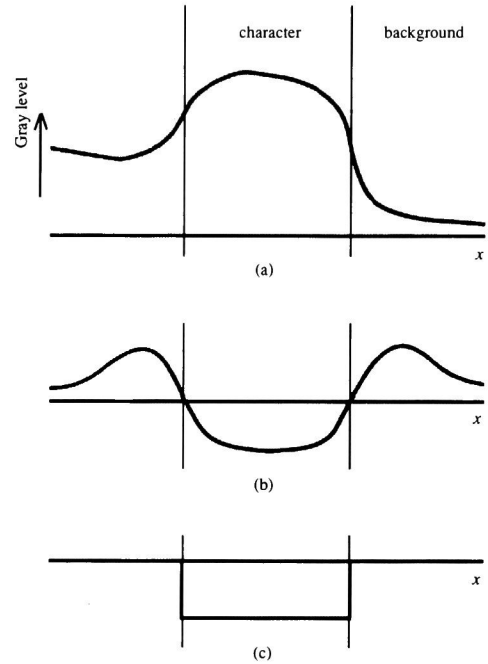


図3 入力画像の1次元の濃度分布 (a) ラプラシアン・ガウシアン出力 (b) 二値画像 (c)

Fig. 3. Input image (a). Output of Laplacian Gaussian (b). Binary image (c).

微分（2階差分）は等価である。しかし、対象とする入力画像に適した十分な大きさのラプラシアン・ガウシアンフィルタを直接適用するためには、多くの処理時間を必要とするため、本研究では、 $2 \times 2$ の平滑化フィルタ(1)を必要回数だけ反復適用した後、距離  $k$  のラプラシアンを適用することによりラプラシアン・ガウシアンフィルタを近似した。

$$M(i, j) = \frac{1}{4} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \dots\dots\dots (1)$$

まず、 $2 \times 2$ の平滑化フィルタを  $n$  回反復することによりガウスフィルタリングを近似する。平滑化フィルタを1回適用する毎にフィルタサイズが1増加して重みは  $1/4$  になるので、5回反復した場合は図4に示す平滑化フィルタと等価になる。次に、図5に示すような距離  $k = n + 1$  のラプラシアンフィルタを適用する。以上の処理により、文字領域は負の値、背景は正の値をとる画像が得られる。入力画像の濃度分布の例を図6に、ラプラシアン・ガウシアン処理後の濃度分布を図7に示す。

〈2・2〉文字列領域の推定と切り出し 文字列の切り出しのために、領域推定と文字列切り出しの2段階に分けて、二値化を行う。まず、ラプラシアン・ガウシアン処理後、背景を除くため、画像の半二値化を行う。画像にラプラシアン・ガウシアンフィルタを施した後の画像を  $f(x, y)$  とすると、半二値化された画像  $g(x, y)$  を次式で求める。

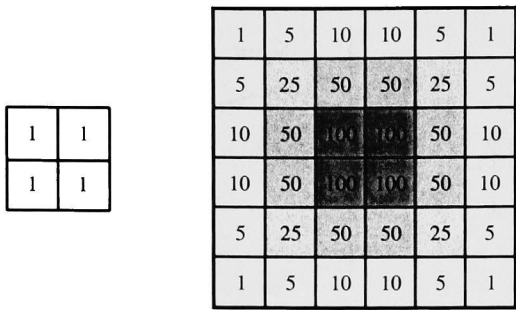


図4 2×2の平滑化フィルタ(左:重みは1/4)を5回反復適用する場合の等価フィルタ(右:重みは1/1024)

Fig. 4. A filter equivalent to 2×2 smoothing filter (left) repeatedly applied five times(right).

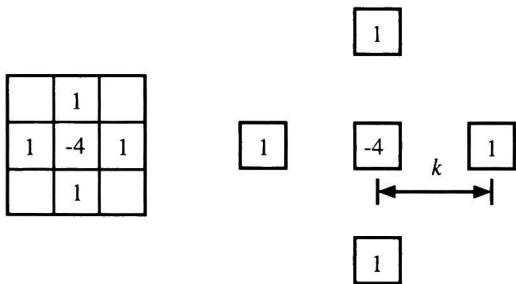


図5 距離1のラプラシアンフィルタと距離  $k = n + 1$  のラプラシアンフィルタ

Fig. 5. Laplacian filter of distance 1 and Laplacian filter of distance  $k = n + 1$ .

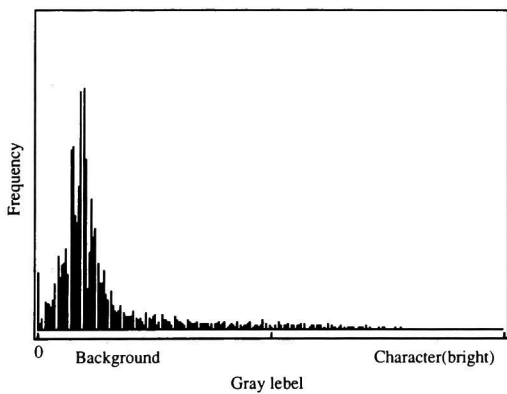


図6 入力画像の濃度分布

Fig. 6. Distribution of input gray level.

$$g(x, y) = \begin{cases} -f(x, y) & (\text{文字部分 } f(x, y) < 0) \\ 0 & (\text{背景部分 } f(x, y) \geq 0) \end{cases}$$

得られた半二値化画像  $g(x, y)$  に対して大津の閾値選択法<sup>(5)</sup>により、閾値  $t_1$  を求め画像  $f(x, y)$  を二値化する。その例

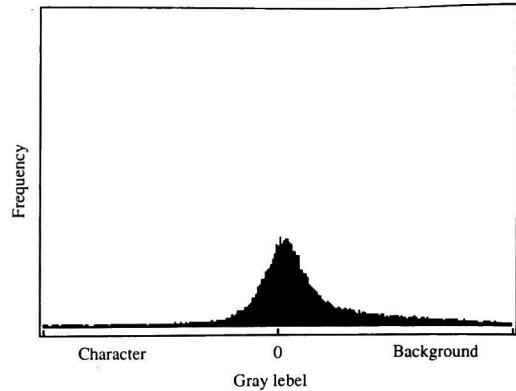


図7 ラプラシアン・ガウシアン処理後の濃度分布

Fig. 7. Distribution of gray level after Laplacian Gaussian filtering.

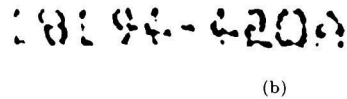
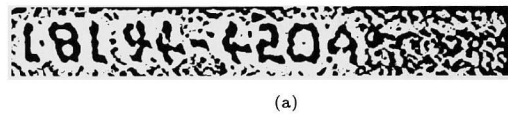


図8 半二値化画像 (a) と閾値  $t_1$  による二値化画像 (b)

Fig. 8. Partial binarization (a) Higher level binarization (b).



図9 閾値  $t_2$  による二値化の例

Fig. 9. Lower level binarization.

を図8 (a), (b) に示す。この画像に対して、水平方向および垂直方向の周辺分布を求め、文字列画像の位置と大きさを推定する。推定された文字列画像に対して、上下に10ドット、左右に5ドットのマージンを設けラプラシアン・ガウシアン処理後の画像からその領域を切り出す。次に、この領域に対して再度  $t_1$  より低い閾値  $t_2$  で二値化する。ここで、半二値化画像  $g(x, y)$  において  $g(x, y) > 0$  となる領域にはすべての文字領域が含まれるが、 $t_2 = 0$  で二値化すると多くのノイズが発生する。そこで、閾値  $t_2 = t_1/10$  を用いて二値化する。二値化の例を図9に示す。このようにして得られた二値画像に対して、ラベリングを施し、小さな連結成分(実験では50画素以下)はノイズとして取り

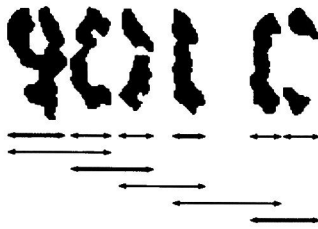


図 10 動的計画法による文字切り出し

Fig.10. Segmentation recognition by dynamic programming.

除く。

### 3. 文字切り出しと認識

文字の切り出しと認識には、二値画像を過分割して得られる複数セグメントに対して文字認識を行いながら動的計画法によって最適な文字の切り出しを行った後<sup>(6)(7)</sup>、再度、切り出された文字を 400 次元の特徴ベクトルを用いて再認識する 2 段階の認識手法を用いる<sup>(6)</sup>。文字切り出しでは、二値化された文字列画像において、連結成分の輪郭線の垂直方向の極小点（谷の部分）でセグメントに分割する。図 10 において、'9' と '0' の左側に相当するセグメントは、1 つの連結成分がその谷の部分で分割されたものである。このようにして得られたセグメントを 1 つ、または、2, 3 セグメントずつまとめた成分を文字認識して文字の尤度（文字らしさ）を求め、文字列の平均尤度が最大となるよう動的計画法を用いてセグメントの組み合わせを選ぶ（図 10 太線）。この文字切り出しにおける文字認識には加重方向指数ヒストグラム法<sup>(8)</sup>による 64 次元の特徴ベクトルを使用する。その後、切り出された文字に対して、濃度こう配から求められる 400 次元の特徴ベクトル<sup>(9)</sup>を用いて再認識をし、高精度化を行う。また、文字列の認識においては、使用する文字は英数字 '0' から '9' と '-' や特定の英字に限られるため、図 11 に示すようなテンプレートを用いた。ここで # には 0 から 9 の数字が入る。最初のテンプレートは数字 6 文字、5 番目のテンプレートは数字 5 文字に '-' がつづき、数字 3 文字の後に英字 'A' がくることを示す<sup>(10)</sup>。個々のテンプレートに対して、テンプレートで指定される条件を満たす文字の組み合わせの中から動的計画法により最適解が求められる。さらに、各テンプレートに対する最適解の中から、文字の平均尤度が最大の最適解を最終的な認識結果とする。文字の尤度計算には擬似ベイズ識別関数を用いた<sup>(9)</sup>。

### 4. 認識実験と結果

実験には、256 階調の白黒ビデオ画像からサンプリングした文字列データを用いた。サンプル数は 1036 で、1 画像に含まれる文字列は英数字横書き 1 行で 1 文字列当たりの平均文字数は約 8.5 文字である。全サンプルを 4 グループ

```
#####
####-##
#####
####-##-#
#####-###A
#####-###C
#####-###X
J#####
H#####
T#####
ZG####-##-#
ZJ####-##-##
ZU####-##-##
```

図 11 テンプレートの例

Fig. 11. Example of template.

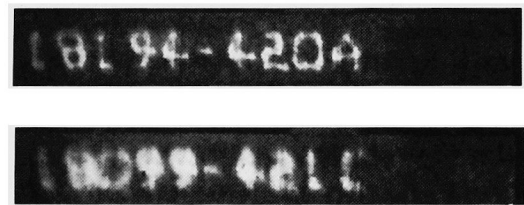


図 12 画質の良い例と悪い例

Fig.12. Examples with high and low image quality.

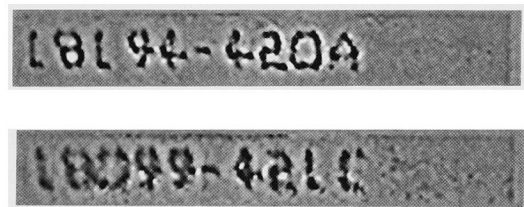


図 13 ラプラシアン・ガウシアン処理を施した画像例

Fig. 13. Laplacian Gaussian filtered images.

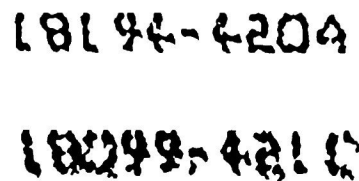


図 14 本手法により二値化した画像の例

Fig.14. Example of binary images obtained by proposed procedure.

に分け、1 つを認識評価用に、残りの 3 グループを合わせて辞書用とした。実験に用いた画像の中で画質が良好なものとそうでないものの例を図 12 に、また、これらにラプラシアン・ガウシアン処理を施した画像の例を図 13 に示す。ま

8194-420A

8099-4211

図 15 入力画像を直接二値化した画像の例  
Fig. 15. Example of direct binarization of input image.

表 1 文字および文字列の認識率

Table 1. Recognition rates of character and character string.

|        | 文字      |         | 文字列       |
|--------|---------|---------|-----------|
|        | 認識率 (%) | 認識率 (%) | 処理時間 (mS) |
| 64 次元  | 98.9    | 88.4    | 305       |
| 400 次元 | 99.2    | 91.6    | 779       |

表 2 文字列の誤読率を 1%以下に押えた場合の認識率

Table 2. Recognition rates of character and character string.

|        | 誤読率 (%) | 棄却率 (%) | 正読率 (%) |
|--------|---------|---------|---------|
| 64 次元  | 0.97    | 19.0    | 80.0    |
| 400 次元 | 0.97    | 11.2    | 87.7    |

た、本手法により二値化した画像と入力画像を直接大津の閾値選択法で選択した閾値により二値化した画像の例を図 14, 15 に示す。入力画像を直接二値化した画像は、ラプラシアン・ガウシアン処理を施したものに比べ、細くかすれた文字や太く滲んだ文字の二値化がうまくなされていない。64 次元および 400 次元の特徴ベクトルを用いた場合の各文字の認識率および、文字列の認識率と処理時間を表 1 に示す。64 次元では文字認識率 98.9%, 文字列認識率 88.4%, 400 次元では文字認識率 99.2%, 文字列認識率 91.6% が得られた。図 16 に第 1 候補から第 5 候補までの文字列の累積認識率を示す。ここで第 1 候補および第 5 候補とはそれぞれ、文字の平均尤度が最大の候補および 5 番目に大きい候補のことである。上位 5 候補の中に正解が含まれている累積認識率は、64 次元の特徴ベクトルで 92.1%, 400 次元の特徴ベクトルで 93.5% となった。また、誤読率を 1% 以下に押さえた場合の文字列認識率を表 2 に示す。64 次元では 80.0%, 400 次元では 87.7% の認識率が得られた。また、図 17 に誤認識の例、図 18 に棄却の例を示す。図 17 の例では文字部分にある傷が誤読の原因、図 18 の例では印字のかすれが棄却の原因と考えられる。

### 5. む す び

ノイズや背景むらの激しい低画質入力画像を対象とする、文字の切り出し・認識手法を提案し有効性を実験的に評価した。実験の結果、

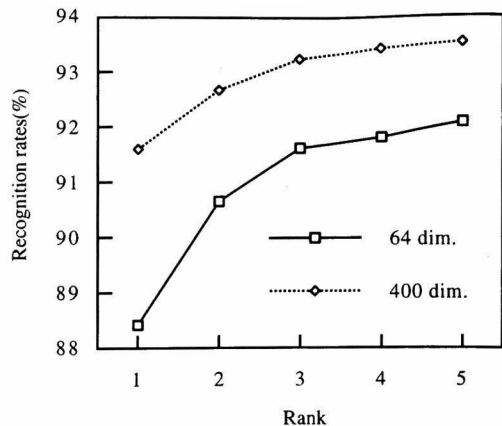


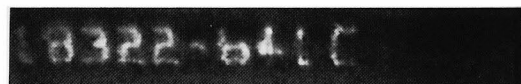
図 16 累積認識率

Fig. 16. Cumulative recognition rates.



18303-661C

図 17 誤認識の例 (18303-661C → 18302-661C)  
Fig. 17. Example of mis-recognition (18303-661C → 18302-661C).



8322-641C

図 18 棄却の例 (18322-641C)  
Fig. 18. Example of rejection (18322-641C).

- (1) ラプラシアン・ガウシアン処理による強調処理と、半二値化、二値化処理によって対象文字列の位置・大きさの推定と領域の切り出しが可能となり、
- (2) 細くかすれた文字と、太く滲んだ文字が混在する場合においても良好な二値化が可能となった。その結果、
- (3) 動的計画法による文字切り出し・認識により文字認識率 99.2%, 文字列認識率 91.6% の良好な結果が得られた。

今後、より多くのサンプルを用いて、平滑化フィルタの反復回数を変化させて評価実験を行い、照明条件の改善など

を含め、誤読率を改善し、鋼材のマーキング文字列以外への応用を検討する予定である。

謝辞 本研究を進めるにあたり画像データを提供していただいた新日本製鐵株式会社エレクトロニクス情報事業本部の皆様に感謝いたします。

(平成13年5月9日受付, 同13年7月25日再受付)

## 文 献

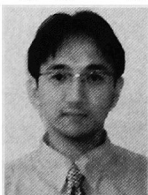
- (1) 中野, 小関, 山本:「低品質刻印文字認識のための最適2値化法」, 信学論, J69-D No.6, pp.958-966 (1986.6)
- (2) 大塚, 谷口, 西田, 大倉:「清酒リサイクル瓶の不鮮明文字と欠陥の識別」, 信学論, J78-DII No.7, pp.1058-1070 (1995.7)
- (3) 山田, 筒井:「弾性エッジ列整合法の階層的適用による2値化・切り出しを伴わないガラス刻印の位置決め・認識」, 信学技報, PRU95-103, pp.1-6 (1995.9)
- (4) 安居院, 中嶋正之:「画像情報処理」, 基礎情報工学シリーズ18, 2, 画像の空間フィルタリング, pp.15-26 (1991) 森北出版社
- (5) 大津展之:「判別および最小2乗基準に基づく自動しきい値選定法」, 信学論, J63-D No.4, pp.349-356 (1982.4)
- (6) F.Kimura, Y.Miyake, M.Shridhar: "Handwritten ZIP Code Recognition Using Lexicon Free Word Recognition Algorithm", ICDAR'95, pp.906-910 (1995.8)
- (7) F.Kimura, M.Shridhar, Z.Chen: "Improvements of a Lexicon Directed Algorithm for Recognition of Unconstrained Handwritten Words", ICDAR'93, pp.18-22 (1993.10)
- (8) 鶴岡, 栗田, 原田, 木村, 三宅:「加重方向指数ヒストグラム法による手書き漢字・ひらがな認識」, 信学論, J70-D No.7, pp.1390-1397 (1987.7)
- (9) 若林, 鶴岡, 木村, 三宅:「特徴量の次元数増加による手書き数字認識の高精度化」, 信学論, J77-DII No.10, pp.2046-2053 (1994.10)
- (10) F.Kimura, M.Shridhar: "Handwritten Address Interpretation Using Extended Lexicon Word Matching", IWFHR'96, pp.369-372 (1996.9)

澤 和 弘 (非会員) 1982年三重大学工学部電子工学科卒業。



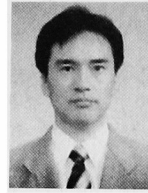
1984年三重大学大学院修士課程修了。三重県立津高等技術学校電子制御情報科勤務。文字認識の研究に従事。電子情報通信学会会員

大 山 航 (正員) 1998年三重大学工学部情報工学科卒業。



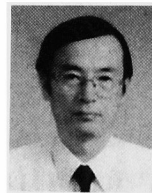
2000年三重大学大学院博士前期課程修了。同年三重大学工学部助手。主として医用画像処理, 生体信号処理に関する研究に従事。電子情報通信学会, 日本エム・イー学会, 各会員。

若 林 哲 史 (非会員) 1985年三重大学工学部電子工学科卒業。



1987年三重大学大学院修士課程修了。1991年三重大学工学部助手。1998年より三重大学工学部助教授。その間, 1998-99年米国レンセア工科大学客員研究員。手書き文字認識, 文書理解, 画像処理, コンピュータグラフィックスに関する研究に従事。工博。電子情報通信学会, 情報処理学会会員。

木 村 文 隆 (非会員) 1973年名古屋大学工学部電気工学科卒業。



1978年名古屋大学大学院博士課程修了。同年名古屋大学工学部助手。1983年三重大学工学部助教授。1989-91年米国ミシガン大学客員助教授。1998年より三重大学工学部教授。文字・パターン認識, 画像処理, コンピュータグラフィックスの研究に従事。工博。電子情報通信学会, 情報処理学会, 日本エム・イー学会, 人工知能学会, 各会員。