# A MATCHING METHOD BASED ON MARKER-CONTROLLED WATERSHED SEGMENTATION

*Yi Hu, Tomoharu Nagao*

Yokohama National University
Graduate School of Environment and Information Sciences
79-5, Tokiwa-dai Hodogaya-ku, Yokohama 240-8501, Japan

**ABSTRACT**

A new template matching method that is based on marker-controlled watershed segmentation (*TMCWS*) is presented. It is applied to recognize numbers on special metal plates on production lines where traditional image recognition methods do not work well. Different from previous matching algorithms, TMCWS firstly creates a marker image for each pattern, and then takes both the pattern image and its corresponding marker image as a template window and shifts this window across a gradient space of an unknown image pixel by pixel to do a search. At each position, the marker image is used to try to extract the contour of the target object with the help of marker-controlled watershed segmentation, and the pattern image is employed to evaluate the extracted shape in each trial. All the pattern images and their corresponding marker images are tried and the pattern that best matches the target object is the recognition result. TMCWS contains shape extraction procedures. Experiments are performed with this method on nearly 400 images of metal plates and the test results show its effectiveness in recognizing numbers in noisy images.

## 1. INTRODUCTION

In some factories, it is necessary and important to recognize a number engraved on a special metal plate since this number is a part of an identifier to processed components. Two example images of the metal plates are shown in Fig. 1. The metal plate images are photographed by a monochrome CCD camera without special lighting and their images have the same size (640 x 480 pixels). The number on each metal plate consists of four characters, each of which can be one of the twelve signs that are ten Arabic numerals plus "M" and "N". For simplicity we temporally coin the word "*metalchar*" to refer to each character, which will be represented with $\Omega_t$ respectively, where $t \in [0, Num-1]$ and $Num = 12$.

Thanks to Mr. Ichio Tomisawa, Mr. Yoshimasa Isobe and Mr. Toshiaki Arai from Mitsubishi Nuclear Fuel Co., Ltd for supporting this research.

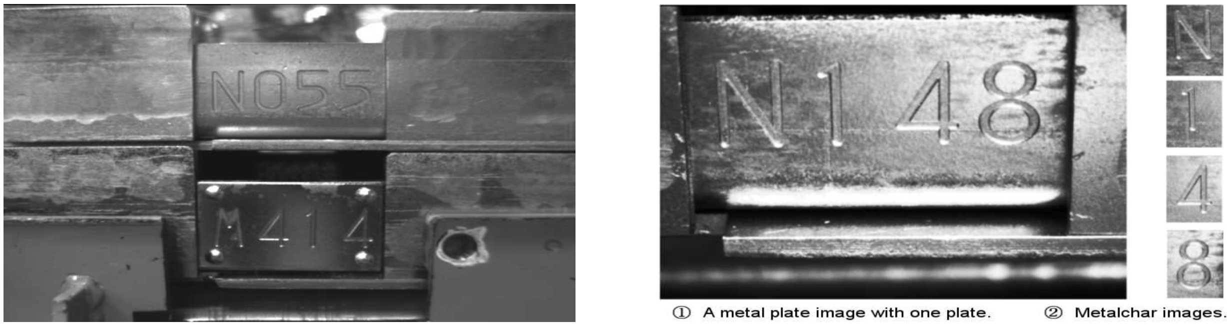
Two main factors make the recognitin difficult. One is that these plate images are much noised due to diffused reflection caused by inhomogenous plate surfaces. As shown by Fig. 1, the metalchars are characterized by their weak contours. The other is that the position of each metalchar in a plate is not all the same with other plates. It differs with plates. The translation difference is sometimes big to about a half-metalchar size. The projection methods (graylevel distribution along x-axis or y-axis) cannot determine the position of each metalchar because of their low contrast. This implies that the recognition method should have an ability of translation invariance.

In related works on recognizing characters on metal, Nakamura et al. developed a high-performance stamped character reader [1], where an image of stamped characters is acquired by scanning a metal rod with a special type of scanner. However, optimal photographic conditions are sometimes difficult or even impossible to meet in a production site due to some limitations. Because of the big difference in image quality caused by different image acquisition ways and different metal materials, their recognition method cannot be applied to our images. Advanced feature extraction tools such as the ones of direct grayscale extraction of features [2] only work on a small number of images. The correlation matching method gives low correlation coefficients causing questionable recognition results.

Here, we present a template matching method based on marker-controlled watershed segmentation (TMCWS). To the best of our knowledge, the approach to setting markers and the combination of repeated marker-controlled watershed segmentation attempts with pattern recognition has not been reported in literature.

A metal plate image is firstly partitioned into several subimages by image preprocessing, each of which isolates an individual character. This subimage is called a metalchar image. In Fig. 1 , (b)② shows 4 metalchar images, the partitioning result of (b)①. TMCWS will be applied to each of the metalchar images to give recogntion result. The paper is organized as follows. The proposed TMCWS method is described in Sect. 2. Section 3 shows some experimental

0-7803-8554-3/04/$20.00 ©2004 IEEE.




(a) (b)

**Fig. 1**. Two example images of metal plates where (a) is an image with two plates and (b)① is with one plate. In (b), ② shows four metalchar images, the image partitioning results from the image ① after preprocessing.

results. The last section is the conclusion.

## 2. THE TMCWS METHOD

In this section, we first introduce the concept of doing marker-controlled watershed segmentation with a pattern marker image, then give rules to create a marker image for a given metalchar pattern. After that, we scan the image exhaustively with a method combining a template matching technique and marker-controlled watershed segmentation to determine which pattern best matches the target object. Finally, we present the recognition algorithm.

### 2.1. Doing marker-controlled watershed segmentation with a pattern marker image

Marker-controlled watershed segmentation [3] is a powerful tool in image segmentation and contour detection, where marker extraction plays a key role in efficient morphological approach to segmentation. Many marker extraction methods, either gradient-based or intensity-based, and their applications have been reported. However, by analyzing several dozens of metal plate images, we found it was difficult to identify some properties that can be shared by most unknown images to be served as markers because the gray level intensity or gradient of metalchars differ with images.

Since metalchars are engraved on metals, they have fixed shapes. Their size information can be acquired from the preprocessing partition procedure (for example, in Fig. 1 (a), by detecing the dark rectangle between two plates can deduce the metalchar size). This allows us to think of these unknown metalchar images as instances of their pattern images. Assuming each unknown metalchar image is denoted by $X_i, (i = 1, 2, \cdots, N_x)$, where $N_x$ is the number of all unknown metalchar images, then those images $X_i$ containing the same metalchar $\Omega_t$ form a set $S_{\Omega_t}, t \in [0, Num-1]$.

$$S_{\Omega_t}=\{X_i|X_i \supseteq \Omega_t, i=1,2,\cdots,N_x\}$$

We establish a pattern image for $S_{\Omega_t}$ called $P_{\Omega_t}$, which represents all characteristics of $\Omega_t$ and is free of noise. Each element of $S_{\Omega_t}$ is denoted by $Y_{j,\Omega_t}$, where $j = 1, 2, \cdots$ .

For 2D discrete images, an image is a function that maps a finite rectangle subset $D_I$ of the discrete plane $Z^2$ into a discrete set $\{g_v\}$ of gray levels, where $g_v \in [minGraylevel, maxGraylevel]$. The $maxGraylevel$ and $minGraylevel$ are the greatest and least possible gray level intensity for each pixel in metalchar images respectively. Assuming $X_i$ is defined on domain $D_I$, $X_i = X_i(\lambda)$, where $\lambda \in D_I$, then for each element $Y_{j,\Omega_t}$ of $S_{\Omega_t}$, $Y_{j,\Omega_t}(\lambda)$ can be regarded as a sum of $P_{\Omega_t}(\lambda - \phi)$ and a noise function $\Psi_j(\lambda)$.

$$\forall Y_{j,\Omega_t} \in S_{\Omega_t} \quad Y_{j,\Omega_t}(\lambda)=P_{\Omega_t}(\lambda-\phi)+\Psi_j(\lambda) \tag{1}$$

where $\lambda \in D_I$, $P_{\Omega_t}(\lambda - \phi)$ is an image translated $\phi$ by $P_{\Omega_t}(\lambda)$.

$Y_{j,\Omega_t}$ is called an instance of $P_{\Omega_t}$. As mentioned above, it is difficult to find a general method to extract markers from each $Y_{j,\Omega_t}$. From Eq. 1, if $\Psi_j(\lambda)$ meets some certain conditions, we hope we can find markers which setting are decided by $P_{\Omega_t}$ and if applying them to each instance $Y_{j,\Omega_t}$ of $P_{\Omega_t}$, performing marker-controlled watershed segmentation with them can extract contours of $\Omega_t$ from most instances of $P_{\Omega_t}$. The image constituted of these markers is called a pattern marker image of $P_{\Omega_t}$, denoted by $M_{\Omega_t}$.

Different from previous techniques of applying watershed transformation to segment objects where markers are extracted or selected from an unknown image, here we do not extract or select markers. We directly create a marker image $M_{\Omega_t}$ from a pattern $P_{\Omega_t}$. This created marker image $M_{\Omega_t}$ is in fact independent of unknown images $X_i$. It is created in advance. We use a different technique to apply marker-controlled watershed segmentation with it to an unknown metalchar image.

### 2.2. Creating a pattern marker image for each metalchar pattern

In this section, we discuss the case of $\phi = 0$ of Eq. 1. This means that the $\Omega_t$ in both $P_{\Omega_t}$ and its instances $Y_{j,\Omega_t}$ have no translation. If the marker image $M_{\Omega_t}$ is suitably designed, then a direct application of marker-controlled watershed segmentation with $M_{\Omega_t}$ to the instances $Y_{j,\Omega_t}$ of $P_{\Omega_t}$ should detect its contours.

Markers stand for some minima to be imposed on a gradient image. Dual reconstruction of a gradient image with a marker image imposes all markers as minima of the gradient while removing all the other minima [4]. The inner marker should be in the inner of an object needs to be detected and the outer marker should situate in the outer part of this object. Based on this, we set the skeleton of a metalchar pattern as its inner marker and use a rectangle enclsoing the pattern as its outer marker. Given a pattern image $P_{\Omega_t}$ of a metalchar $\Omega_t$, its pattern marker image $M_{\Omega_t}$ is dependent on $P_{\Omega_t}$ and is created by the following steps:

(1) $M_{\Omega_t}$ is initialized to have the same size with $P_{\Omega_t}$ and have its every pixel with a value of $maxGraylevel$. It looks like a white empty image.

(2) An inner marker of $M_{\Omega_t}$ is a replica of a skeleton of $P_{\Omega_t}$ copied to $M_{\Omega_t}$.

(3) An outer marker of $M_{\Omega_t}$ is a rectangle. If it is copied into $P_{\Omega_t}$, then it should circumscribe $\Omega_t$. It looks like a black rectanglar frame.

(4) If an inner marker contains loops, then a small disc should be created at the center of each loop and this disc becomes an outer marker.

(5) The value in pixel of all markers, whether inner markers or outer markers, is set as $minGraylevel$.

According to the above rules, Fig. 2 shows the corresponding marker image of each metalchar pattern.

### 2.3. Scanning with a template matching technique

In the case of $\phi \neq 0$ of Eq. 1, the marker image $M_{\Omega_t}$ created above cannot be used directly. And another problem

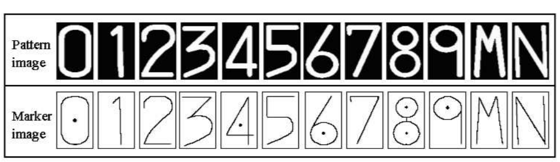



**Fig. 2**. The pattern images and their corresponding marker images of metalchar patterns.

is, given an unknown image $X_i$, we have no knowledge whether $X_i$ is an instance of $P_{\Omega_t}$.

The search method used in template matching technique can help us solve the above two problems. As shown by Fig. 3, assuming $X_i$ is an instance of $P_{\Omega_t}$, we regard the above created $M_{\Omega_t}$ as a template window and shift this window pixel by pixel across the gradient space $G_i$ of $X_i$. At each position, we do marker-controlled watershed segmentation with $M_{\Omega_t}$ on a subimage of $G_i$ under $M_{\Omega_t}$, and measure the similarity between the watershed cachment basin $W_c$ formed by this segmentation and the pattern image $P_{\Omega_t}$. The similarity is denoted as $R_{tm}(X_i, P_{\Omega_t}, W_c)$. The maximal $R_{tm}$ during this search will test our assumption. If the assumption is correct, then $R_{tm}$ is high, or else it is low. Given an unknown image, we try all the possible cases of $P_{\Omega_t}$ and $\phi$ by scanning the image exhaustively and finally pick the highest similarity value. The $\Omega_t$ corresponding to the highest similarity is the recognition result.

The images $P_{\Omega_t}$ and $W_c$ are binary images where a pixel with a value "1" means it is on the object and "0" means not. To emphasize the shape comparison, we employ weighting coefficients. The similarity evaluation function between $P_{\Omega_t}$ and $W_c$ is designed as:

$$R_{tm}(X_i,\Omega_t,\phi)=R_{tm}(X_i,P_{\Omega_t},W_c)$$
$$= \frac{\sum_{u=0}^{N_T-1}\begin{cases} dw1 & (p_u == 1 \;\delta\delta\; w_u == 1) \\ dw2 & (p_u == 1 \;\delta\delta\; w_u == 0) \\ dw3 & (p_u == 0 \;\delta\delta\; w_u == 1) \\ dw4 & (p_u == 0 \;\delta\delta\; w_u == 0) \end{cases}}{\sum_{u=0}^{N_T-1}\begin{cases} dw1 & (p_u == 1 \;\delta\delta\; w_u == 1) \\ dw4 & (p_u == 0 \;\delta\delta\; w_u == 0) \end{cases}} \quad (2)$$

where $dw1$,$dw2$,$dw3$ and $dw4$ are weighting coefficients, $\delta\delta$ is a logic conditional AND operator, $dw1 > 0$, $dw2 < 0$, $dw3 < 0$, $dw4 > 0$, $dw1 + dw4 = 1.0$, $p_u$ and $w_u$ are pixels of $P_{\Omega_t}$ and $W_c$ separately and $N_T$ is the total number of pixels in $P_{\Omega_t}$.

If we define $R_{tm}(X_i)$ and $\Omega_{tm}$ as below, then $\Omega_{tm}$ is the reconigtion result of an unknown metalchar image $X_i$.

$$R_{tm}(X_i,\Omega_t)=max\{R_{tm}(X_i,\Omega_t,\phi) \mid all\ possible\ \phi\}$$
$$R_{tm}(X_i)=max\{R_{tm}(X_i,\Omega_t) \mid t\in[0,Num-1]\}$$
$$\Omega_{tm}=\{\Omega_t | R_{tm}(X_i,\Omega_t)=R_{tm}(X_i),t\in[0,Num-1]\}$$

### 2.4. The recognition algorithm

Given an unknown metalchar image $X_i$, the recognition algorithm of TMCWS is concluded as below:

*Read an unknown metalchar image* $X_i$
*Do morphological gradient on* $X_i \rightarrow$ *a gradient image* $G_i$
$G_i + 1 \rightarrow G_i$
*For (*$t \leftarrow 0$*;* $t < Num$*;* $t \leftarrow t+1$*)* {
*Take a pattern image* $P_{\Omega_t}$ *and its marker image* $M_{\Omega_t}$

**Table 1**. The comparison of applying the TMCWS method and the correlation matching method to images in Fig. 1.

| Unknown images | Recognition results | | | | | | | | Correct results |
|---|---|---|---|---|---|---|---|---|---|
| | The correlation matching method: $\Omega_{cor}(R_{cor}(X_i))$ | | | | The TMCWS method: $\Omega_{tm}(R_{tm}(X_i))$ | | | | |
| (a) | M(0.33) | 7(0.44) | 7(0.38) | 9(0.33) | N(0.88) | 0(0.85) | 5(0.83) | 5(0.82) | N055 |
| | M(0.53) | 4(0.44) | 9(0.39) | 4(0.52) | M(0.79) | 4(0.91) | 1(0.98) | 4(0.86) | M414 |
| (b) | N(0.36) | 9(0.26) | 1(0.32) | 8(0.33) | N(0.73) | 1(0.92) | 4(0.83) | 8(0.73) | N148 |

*Try all possible* $\phi$ *by shifting* $M_{\Omega_t}$ *across* $G_i$ {
*The subimage in* $G_i$ *under* $M_{\Omega_t} \rightarrow tmpImg1$
$min(tmpImg1, M_{\Omega_t}) \rightarrow tmpImg2$
*Do dual reconstruction with* $M_{\Omega_t}$ *on* $tmpImg2 \rightarrow tmpImg3$
*Do watershed transformation on* $tmpImg3 \rightarrow$ *watershed basin* $W_c$
*Calculate* $R_{tm}(X_i, \Omega_t, \phi)$ }
*Calculate* $R_{tm}(X_i, \Omega_t)$
}
*Calculate* $R_{tm}(X_i)$
*Get the recognition result* $\Omega_{tm}$ *of* $X_i$

## 3. EXPERIMENTAL RESULTS

The above recognition algorithm searches a gradient space to give recogntion result. Different gradient operators make different gradient image spaces. Here, two gradient operators $Rar(X_i)$ and $Dyr(X_i)$ shown below are used to organize the gradient spaces in a two-layer space from coarse to fine.

$$Dyr(X_i)=(X_i \oplus B)-(X_i \ominus B)$$
$$Ter(X_i)=(X_i \bullet B)-(X_i \ o \ B)$$
$$Rar(X_i)=Dyr(X_i)-Ter(X_i) \tag{3}$$

where B is a flat structuring element, $\oplus$, $\ominus$, $o$, and $\bullet$ represent morphological dilation, erosion, opening and closing operator respectively.

The space decided by $Rar(X_i)$ is first searched. If the value of $R_{tm}(X_i)$ is less than 0.70, then the other space

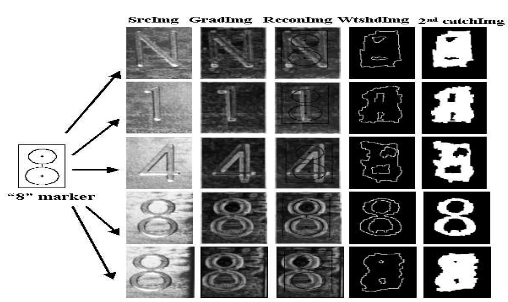



**Fig. 3**. The segmentation comparison of moving a same marker image $M_{\Omega_8}$ onto different metalchar images.

decided by operator $Dyr(X_i)$ will be explored. The maximal $R_{tm}(X_i)$ during the two trials decides the final recognition result. Parameters in Eq. (2) are set as $dw1 = 0.8$, $dw2 = -0.2$, $dw3 = -2.2$ and $dw4 = 0.2$. The structuring element B in Eq. 3 and in image dual reconstruction [4] is set as a 4 x 4 flat structuring element.

As a comparison, the correlation matching method is tested. Table 1 lists the results with the two methods on images of Fig. 1, showing that TMCWS recognizes all characters. In addition to this, 395 images of metal plates (2692 metalchars) were tested and 99% recognition ratio was acquired. The test shows that TMCWS method achieves higher recognition ratio than the correlation matching method does (78%). This can be attributed to that the TMCWS method contains shape extraction procedures.

## 4. CONCLUSION

The proposed TMCWS method is a template matching method that uses different pattern images and their corresponding marker images as probes to explore a gradient space of an unknown image to determine which pattern best matches the target object in it. It requires few parameter tuning. Nearly 400 images are tested and the results show its effectiveness in recognizing numbers in noisy images.

## 5. REFERENCES


[1] Y. Nakamura et al., "Development of a high-performance stamped character reader," *Proc. of IECON'84*, pp. 1011–1014, Oct. 1984.

[2] Li Wang and Theo Pavlidis, "Direct gray-scale extraction of features for characters recognition," *IEEE Trans. Patt. Anal. Machine Intell*, vol. 15, no. 10, 1993.

[3] Meyer F. and Beucher S., "Morhological segmentation," *Journals of Visual Communication and Image representation*, vol. 1, no. 1, pp. 21–46, Sep. 1990.

[4] Luc Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," *IEEE Trans. Image Proc.*, vol. 2, no. 2, pp. 176–201, April 1993.