DOI: 10.2352/ISSN.2470-1173.2016.11.IMAWM-459

HAZMAT SIGN LOCATION DETECTION BASED ON FOURIER SHAPE DESCRIPTORS

Kharittha Thongkor†, Albert Parra Pozo‡, Bin Zhao‡, Thumrongrat Amornraksa† and Edward J. Delp‡

†Multimedia Communication Laboratory, Computer Engineering Department,
King Mongkut's University of Technology Thonburi, Bangkok, Thailand
‡Video and Image Processing Laboratory (VIPER), School of Electrical and Computer Engineering,
Purdue University, West Lafayette, Indiana, USA

ABSTRACT

In this paper we describe a method for hazardous material (hazmat) sign location detection based on Fourier shape descriptors. The proposed method uses matching from both the magnitude and phase of the Fourier descriptor. The contribution of this paper includes a contour extraction method based on color channel clipping followed by image binarization. The experimental results show that our method is robust to geometric distortion, low resolution, blur, lighting conditions, and perspective.

***Index Terms*—** sign location detection, Fourier shape descriptors, color channel clipping

1. INTRODUCTION

Hazardous materials can cause harm to humans, animals and the environment if there is exposure to the materials due an accident or spill. A federal law in US requires vehicles transporting or containing hazardous materials to be marked with a standard sign (i.e., a "hazmat sign") identifying the type of material the vehicle is carrying [1]. A typical hazmat sign is shown in Figure 1a and has identifying information described by the sign shape, color, symbols, and numbers. In the event of an emergency, first responders have to browse the Emergency Response Guidebook (ERG), published by the US Department of Transportation (DOT) [2], to identify the material and determine what equipment, procedures and precautions should be used to mitigate the situation. This process is slow and difficult for these who are not familiar with the guidebook. Although there exist several mobile-based applications that provide easy access to this guidebook [2, 3], they only provide manually browsing functionality. Hence, an integrated mobile-based system that makes use of location-based-services and image analysis methods has been developed called MERGE (Mobile Emergency Response GuidE). It is capable of detecting hazmat signs from an image acquired using the mobile device and querying an internal database to provide accurate and useful information to first responders in real time and to minimize the amount of work of the users especially in the emergency situation. B. Zhao *et al.* [4] proposed to use saliency maps to denote regions that are likely to contain hazmat signs and A. Parra *et al.* [5] proposed to use geometric constraints and a convex quadrilateral shape detector to find hazmat sign candidates in such regions. These methods have an overall sign detection accuracy of 64.50%. In this paper we propose a new approach to hazmat sign location detection based on Fourier shape descriptors. This approach helps to improve the accuracy of existing methods presented in [4, 6]. Our proposed technique uses contour shape representation and matching based on the magnitude and phase of the Fourier descriptor. The Fourier shape descriptors are generated from morphologically modified closed contours, which are obtained using color channel clipping followed by image binarization. Since we will implement it in a mobile telephone, we are concerned about speed. The classical tools, i.e. binarization and morphological techniques, are thus being considered and applied to generate close contours because they are simple and fast.

This work was partially supported by the U.S. Department of Homeland Security's VACCINE Center under Award Number 2009-ST-061-CI0001. Address all correspondence to Edward J. Delp (ace@ecn.purdue.edu). The first author would like to thank the Thailand Research Fund through the Royal Golden Jubilee Ph.D. Program (Grant No. PHD/0111/2553) and King Mongkut's University of Technology Thonburi for financial support.

2. REVIEW OF EXISTING METHODS

We estimate sign location based on shape, which is an important low level image feature [7, 8]. Humans are able to recognize objects using only shape information [9]. Shape can be described using "shape descriptors," which can be generally classified into two methods: contour-based methods and region-based methods [10]. Contour-based methods only exploit boundary information while region-based methods exploit all the pixels within a region. Contour-based methods are widely used in many applications because of their simplicity [11]. Although shape signatures obtained through contour-based methods are not generally robust to noise [11] the Fourier descriptor (FD) overcomes noise sensitivity by usually using only the first few low frequency coefficients to describe shape. The FD is also compact and

DOI: 10.2352/ISSN.2470-1173.2016.11.IMAWM-459

easy to normalize. Because of its properties the FD is one of the most used shape descriptors [6, 10]. In addition, it has been shown that the FD outperforms many other shape descriptors [11].

Previous work on FDs includes methods for generating descriptors invariant to geometric transformations and matching methods for shape similarity and image retrieval. For example, Kunttu *et al.* [12] proposed a new Fourier descriptor for image retrieval by exploiting the benefits of both wavelet and Fourier transforms. A complex wavelet transform is first used on the shape boundary, and then the Fourier transform of the wavelet coefficients at multiple scales is examined. Since FD is used at multiple scales, the shape retrieval accuracy improves with respect to using ordinary FDs. In 2004, Tahir *et al.* [9] analyzed FDs as feature vectors for pedestrian shape representation and recognition. The results showed that only ten descriptors of both low and high frequency components of pedestrian and vehicle shapes are enough for accurate recognition. In 2007, Kaick *et al.* [13] used the shape context from [14] to generate descriptors and proposed a matching method that uses correspondences between two shapes based on ant colony optimization. Jie *et al.* [15] described simple shapes using FDs based on chain codes and the Fourier transform. The first ten coefficients are used to approximate the shapes. In 2011, Larsson *et al.* [6] proposed a method that uses the Fourier transform of local regions using a maximally stable external region detector. They proposed a FD matching method that uses the phase information to extract the orientation of the shape and used the FDs for recognizing road signs. However, this method fails when signs have low resolution. A. Parra *et al.* [5] and B. Zhao *et al.* [4] proposed methods for hazmat sign detection and recognition. These methods were based on segment detection and grouping using geometric constraints and a saliency map. These methods are fast but straight edges at 45 degrees can be missed in low resolution images.

3. FOURIER DESCRIPTORS

Fourier Descriptor (FD): The Fourier descriptor describes the shape of an object through the use of the Fourier transform of the object's contour. Assuming the contour of a shape has N pixels, numbered from 0 to $N-1$, a set of coordinates describing the contour can be defined as

$$b(k) = (x(k), y(k)) = x(k) + iy(k), \tag{1}$$

where $k = 0, 1, 2, \ldots, N-1$. The Fourier transform of the contour function, $A(v)$, is the FD:

$$A(v) = F(b(k)) = \frac{1}{N} \sum_{k=0}^{N-1} b(k) exp^{-\frac{i2\pi vk}{N}}, \tag{2}$$

where $v = 0, 1, 2, \ldots, N-1$. To describe the shape of a boundary the Fourier coefficients have to be normalized to make them invariant to translation and scale [6, 9, 12, 15, 16]. To normalize the coefficients we set the first coefficient, $A(0)$, to zero and then we normalize to 1 the energy of the remaining coefficients:

$$A'(v) = \frac{A(v)}{\sqrt{\sum_{v=1}^{\infty} |A(v)|^2}}, \tag{3}$$

where $A'(v)$ is the normalized FD. The low frequency components of $A'(v)$ contain information about the general shape and the high frequency components contain finer details. Therefore, the first P Fourier descriptor coefficients can be used to create an approximate reconstruction of the contour $b(k)$,

$$\widehat{b}(k) = \frac{1}{P} \sum_{v=0}^{P-1} A'(v) exp^{\frac{i2\pi vk}{N}}, k = 0, 1, 2, \ldots, N-1. \tag{4}$$

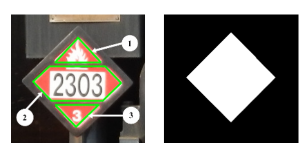
Correlation-Based Matching: Correlation-based matching estimates the cost between two normalized FDs. The cost is defined as

$$e = 2 - 2 \max_l |r_{IJ}(l)|, \tag{5}$$

where $r_{IJ}(l)$ is the cross-correlation between the normalized FDs of contours I and J, which is defined as $r_{IJ}(l) = F^{-1}\{A'_I A'_J\}(v)$. $A'_I(v)$ and $A'_J(v)$ are the normalized FDs of the two contours.

4. PROPOSED METHOD

In order to determine if a contour obtained from an image belongs to a hazmat sign we need to compare its FD against the FD of a predefined shape template or shape contour. In this paper the shape template is a diamond shaped binary image resembling a hazmat sign (see Figure 1b).



(a) (b)

Fig. 1: (a) Example of a hazmat sign divided into three regions and (b) a diamond shaped binary image used as a shape template.

Contour matching can be done in the spatial or frequency domain. We use matching in the frequency domain for two reasons. First, matching in the frequency domain is scale independent, as opposed to spatial domain matching. Second, matching in the spatial domain involves scanning an image multiple times modifying the scale and rotation of the shape template. Since the normalized FDs are invariant to scale,

©2016 Society for Imaging Science and Technology
DOI: 10.2352/ISSN.2470-1173.2016.11.IMAWM-459

and the correlation matching in frequency domain is invariant to rotation, the matching is less computationally expensive. The frequency domain matching has also been shown to be more efficient [17] and allows easy recognition for rotated and scaled noisy sign images [18].

Hazmat signs mounted on vehicles are usually contained in a metallic frame containing two horizontal slides to hold the sign, dividing the sign into three regions as shown in Figure 1a. Therefore, the methods used to detect road signs in [6] cannot be directly used to locate hazmat signs because each shape of sign will be detected separately. Hence, it would be difficult to know which shapes belong to the same hazmat sign. In our method we need to use morphological operations to merge shapes and regions that may belong to the same sign. Figure 2 shows the block diagram of our proposed hazmat sign location detection technique.

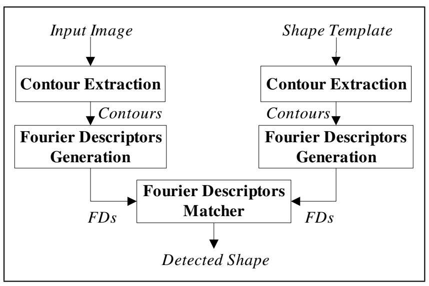


Fig. 2: Proposed sign location detection technique.

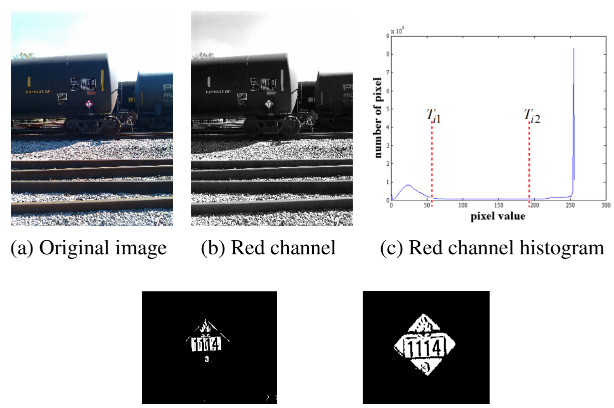
Contour Extraction: The hazmat signs in our dataset contain either one or two of the following colors: black, white, red, blue, green and yellow. In our method we then divide the input image into six color channels and we process them as separate images. The red, green and blue channels are obtained from the RGB color space. The yellow channel is obtained from the CMYK color space. The black and white channels are obtained from the luminance channel. Since images containing hazmat signs are likely to be acquired with various lighting conditions, each of the six images are binarized separately. However, the ordinary binarization method that directly uses Otsu's thresholding method on a channel can fail in some situations, i.e. it does not produce accurate results when images contain variable illumination [19]. For this purpose we propose the use of color channel clipping followed by Otsu's thresholding to obtain the final binary image [20]. For each of the six color channel images, I_i, $i \in [1, 6]$, we first select two parameters for channel clipping, T_{i_1} and T_{i_2}. The histogram of each color channel can be analyzed for minima/valleys which can then be used to determine two thresholds, T_{i_1} and T_{i_2}, as follows.

$$T_{i_1} = min\left(\frac{255}{4}, h_{i_1}\right); T_{i_2} = max\left(3\frac{255}{4}, h_{i_2}\right), \quad (6)$$

where h_{i_1} is the location of the first valley of the histogram of the i^{th} color channel, and h_{i_2} is the location of the last valley of the histogram of the i^{th} color channel. The first valley is the minimum point between the first two significant peaks. A significant peak is a local maximum of the histogram that is greater than its closest local maximum on the left and on the right. The color channel image I_i is then clipped by:

$$I'_i(x, y) = \begin{cases} 0 & I_i(x, y) \leq T_{i_1} \vee I_i(x, y) \geq T_{i_2} \\ I_i(x, y) & \text{Otherwise} \end{cases} \quad (7)$$

Each image I'_i is then used as input for Otsu's thresholding method to automatically generate a threshold T_{i_b}. Each color channel will get the different threshold T_{i_b}. Finally, each original color channel image I_i is then binarized using T_{i_b}. Figure 3 illustrates a comparison using Otsu's method with and without our proposed color channel clipping method. Note how Otsu's method fails to find the optimal threshold because of the high density of pixels in the sky region having high intensity values in the red channel. As we mentioned above we use morphological techniques



(a) Original image (b) Red channel (c) Red channel histogram

(d) Otsu's method (e) Proposed method

Fig. 3: Example of image binarization using Ostu's method with and without our proposed color channel clipping.

to merge areas in the binary image found above that may belong to the same hazmat sign. First, we use a flood-fill operation to fill holes in the binary image [21]. A hole is a set of background pixels surrounded by foreground pixels. Next, we use morphological dilation with a 5 pixel diamond shape structuring element S_d to enlarge the boundaries of foreground areas [21]. Then, we remove small objects by using a morphological opening with a 20 pixel diamond shape structuring element S_o. We also remove objects containing less than $T_c = 0.03\%$ of the total number of pixels in the image. We chose 0.03% because it is the minimum number of pixels contained in a hazmat sign in our image test set. Finally, we obtain closed contours by tracing the exterior boundaries of objects in the resulting binary image [22]. Note that the size of the structuring elements are empirically obtained from the

DOI: 10.2352/ISSN.2470-1173.2016.11.IMAWM-459

ground-truth data in our dataset. They came from searching the best values that give the maximum number of signs before tracing the exterior boundaries of objects.

Fourier Descriptors: Each contour found from the previous step is used to generate a FD using Equation (2). FDs for both the shape template and image contours are obtained in this step. FD matching is usually done by using only the magnitude and ignoring the phase information. By discarding the phase information we achieve rotation and starting point invariance [23]. However, different shapes can have similar magnitude but completely different phase information, thus making magnitude-based matching less accurate [6]. Therefore, we use a correlation-based matching cost function that uses both magnitude and phase information [6]. The cross-correlation between the shape template contour T and the image contour I, $r_{TI}(l)$ is

$$\begin{aligned} r_{TI}(l) &= (T * I)(l) = \int_0^K T(k)I(l+k)\,\mathrm{d}k \\ &= \sum_{v=0}^{\infty} A'_T(v)A'_I(v)exp^{-\frac{i2\pi vl}{K}} \\ &= F^{-1}\{A'_T A'_I\}(v). \end{aligned} \quad (8)$$

By using normalized contours and complex FD matching we approximately compensate for scaling, rotation, translation and starting point. We say "approximately" because we are only using the first few Fourier coefficients to describe the shape of the contour. To find the appropriate number of Fourier coefficients, n, needed for matching we examined the effect of varying the number of low-frequency coefficients we used from our shape template. Using more Fourier coefficients than necessary leads to increasing computation time and does not significantly improve the matching performance [11]. Thus, only the first eight Fourier coefficients were used in our experiments ($n = 8$). To decide if a contour extracted from an image corresponds to a hazmat sign we check if the correlation-based matching cost e between the normalized FD of our shape template and the normalized FD of the extracted contour is below a threshold T_e. To obtain the value of T_e we calculate the correlation-based matching cost e between our shape contour (Figure 1b) and each of various shape template contours. Since the cost of matching our shape template against a diamond shape (including rotation) is not greater than 1.75 we set $T_e = 1.75$. Table 1 shows all the parameters/thresholds we used including empirically derived parameters.

5. EXPERIMENTAL RESULTS

We tested the accuracy of our proposed technique using as image dataset that consisted of 50 images, each containing one or more hazmat signs (62 hazmat signs in total). The images were acquired by first responders using three different cameras: a 8.2 Mpx Kodak Easyshare C813, a 16 Mpx Nikon Coolpix S800c, and a 5 Mpx camera on an HTC Wildfire mobile telephone. The images were acquired in the field, under various lightning conditions, distances, and perspectives. Among the 50 images, 23 were acquired at 10-50 feet, 23 at 50-100 feet, and 4 at 100-200 feet. Among the 62 hazmat signs, 2 had low resolution, 11 had projective distortion, 8 were blurred, and 6 were shaded.

Table 1: Parameters and thresholds used in our proposed method. T_c is based on the total number of pixels in the image.

Parameter	Description	Value
S_d	Structuring element for dilation	5
S_o	Structuring element for opening	20
T_c	Connected components threshold	0.03%
n	Number of Fourier coefficients for matching	8
T_e	Correlation-based matching cost threshold	1.75

Table 2: Hazmat Sign Location Detection Results.

Total Images	Total Signs	Detected Signs (Accuracy)		
		Method in [6]	Method in [4]	Proposed Method
50	62	9 (14.52%)	40 (64.52%)	45 (72.58%)

We implemented the methods in [6] and [4] and then compared their accuracy against our method. Table 2 shows the results. Our method has a hazmat sign location detection rate of 72.58%, while the detection rates for [6] and [4] are 14.52% and 64.52%, respectively.

6. CONCLUSIONS AND FUTURE WORK

We proposed a method for hazmat sign location detection based on Fourier shape descriptors. The experimental results show that our method is more accurate than existing techniques for hazmat signs. We can further increase the sign location accuracy by using techniques to merge multiple contours belonging to the same hazmat sign. The false positives include diamond shape contours that are not hazmat signs. We can decrease the false positive rate by analyzing the inside of the contours and looking for specific text or symbols. Text recognition may require the use of Optical Character Recognition techniques, and symbol recognition may require template matching using different template images. Note that in the case of low resolution images the above methods may not be efficient.

©2016 Society for Imaging Science and Technology
DOI: 10.2352/ISSN.2470-1173.2016.11.IMAWM-459

7. REFERENCES

[1] United States Department of Transportation, *Code of Federal Regulations, Title 49, DOT Hazmat*, Labelmaster, 2012 edition, October 2012.

[2] ERG, Available: http://www.phmsa.dot.gov/hazmat/library/erg.

[3] WISER, Available: http://wiser.nlm.nih.gov.

[4] B. Zhao, A. Parra, and E. J. Delp, "Mobile-based hazmat sign detection system," *Proceedings of the IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pp. 735–738, December 2013, Austin, TX.

[5] A. Parra, B. Zhao, A. Haddad, M. Boutin, and E. J. Delp, "Hazardous material sign detection and recognition," *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, September 2013, Melbourne, Australia.

[6] F. Larsson, M. Felsberg, and P.-E. Forssen, "Correlating fourier descriptors of local patches for road sign recognition," *IET Computer Vision*, vol. 5, pp. 244–254, January 2011.

[7] T. Pavlidis, "A review of algorithms for shape analysis," *Computer Graphics and Image Processing*, vol. 7, no. 2, pp. 243–258, 1978.

[8] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 4, pp. 509–522, 2002.

[9] N. M. Tahir, A. Hussain, and M. M. Mustafa, "Fourier descriptor for pedestrian shape recognition using support vector machine," *Proceedings of the IEEE International Symposium on Signal Processing and Information*, pp. 636–641, December 2007, Cairo, Egypt.

[10] O. R. Mitchell and Timothy A. Grogan, "Global and partial shape discrimination for computer vision," *Optical Engineering*, vol. 23, no. 5, pp. 484–491, October 1984.

[11] D. Zhang and G. Lu, "Evaluation of MPEG-7 shape descriptors against other shape descriptors," *Multimedia System*, vol. 9, pp. 15–30, July 2003.

[12] I. Kunttu, L. Lepisto, J. Rauhumaa, and A. Visa, "Multiscale fourier descriptor for shape-based image retrieval," *Proceedings of the IEEE Conference on Pattern Recognition*, pp. 765–768, August 2004, Cambridge, United Kingdom.

[13] O. van Kaick, G. Hamarneh, H. Zhang, and P. Wighton, "Contour correspondence via ant colony optimization," *Proceedings of the Pacific Conference on Computer Graphics and Applications*, pp. 271–280, October 2007, Maui, HI.

[14] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 509–522, April 2002.

[15] M. Jie, Z. Zhiwei, T. HongMei, and Z. QuanMing, "Fast fourier descriptor method of the shape feature in low resolution images," *Proceedings of the IEEE Conference Wireless Communications Networking and Mobile Computing*, pp. 1–4, September 2010, Chengdu, China.

[16] C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane closed curves," *IEEE Transactions on Computers*, vol. 21, no. 3, pp. 269–281, March 1972.

[17] F. Essannouni and D. Aboutajdine, "Fast frequency template matching using higher order statistics," *IEEE Transactions on Image Processing*, vol. 19, no. 3, pp. 826–830, 2010.

[18] E. Persoon and K. S. Fu, "Shape discrimination using fourier descriptors," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 7, no. 3, pp. 170–179, March 1977.

[19] C. Correa, C. Valero, and P. Barreiro, "Row crop's identication through hough transform using images segmented by robust fuzzy possibilistic c-means," *Proceedings of the Spanish Association for Artificial Intelligence*, November 2011, La Laguna, Spain.

[20] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62–66, January 1979.

[21] G. Anelli, A. Broggi, and G. Destri, "Decomposition of arbitararily-shaped morphological structuring elements using genetic algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 2, pp. 217–224, 1998.

[22] H. Park and R.T. Chin, "Decomposition of arbitararily-shaped morphological structuring elements," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 1, pp. 2–15, 1995.

[23] I. Bartolini, P. Ciaccia, and M. Patella, "WARP: Accurate retrieval of shapes using phase of fourier descriptors and time warping distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 1, pp. 142–147, 2005.