Automatic Segmentation of Pallet Images Using the 2-D Wavelet Transform and YUV Color Space

Aida Fooladivanda, Narges Chehrerazi, Saeed Sadri, Rasoul Amirfattahi, Mohammad Ali Montazeri
Digital signal processing research lab., Department of Electrical and computer engineering Isfahan University of tech. 84156-83111.
Isfahan.Iran.
Email:{a.fooladivanda, n.chehrerazi}@ec.iut.ac.ir ,{sadri, fattahi, montazer}@cc.iut.ac.ir

Abstract— **Segmentation is an important stage in automatic digital image processing. A special case of segmentation is to segment objects from their background. Among different segmentation algorithm for object detection, learning based approach is widely applied. In steel industry, pallets are moving on a rail. They have high resolution details in their structure and the image of a pallet taken by a camera in real time suffers from severe noise and illumination variations. The purpose of this paper is to segment the pallet from a frame of a sequence of video images, such that the pallet is segmented without degradation of resolution. We use the pallet image in YUV color space together with wavelet transform (WT) for detection. For classification Support Vector Machine (SVM) is incorporated to the images. It is shown that the above procedure segments the pallets successfully without degradation of resolution.**

Keywords: Segmentation; YUV Color Space; Wavelet Transform; SVM.

I. INTRODUCTION

Object detection is a crucial stage in industry inspection. For instance object detection is used in driver assistance systems, traffic monitoring, robotics, security systems, video surveillance and medical image processing.

In recent years, for industry inspection, due to its importance, different algorithms have been developed for object detection and classification. For this task image processing techniques is widely used. The first step of these algorithms is to detect and segment the main object from the background.

In steel industry, in pelletizing unit usually there is a chain of pallets moving on a rail, carrying the pellets into furnace for heat operations. In Mobarake Steel Industry (Isfahan, Iran), each pallet contains four rows of grate-bars, each row containing 90 grate-bars. Each pallet has a length of 3.6 meter and width of 1.5 meter. Each of the grate-bars which are composed of incombustible brick is 0.3 meter in length and 0.035 meter in width. During heat operation different kinds of defects is imposed on pallets and their grate bars. These defects cause an intensive amount of economical deficit for the steel industry, so intensive attention has been paid for detecting and classifying different defects of the pallets and maintaining them by using laborers as inspectors.

There are a vast amount of algorithms for object detection in industry inspection. For instance in [1] an automated flaw detection method in aluminum castings based on the tracking of potential defects in a radioscopic image sequence was presented. Special filtering and masking were used to segment the casting defects. In [2] a hybrid image segmentation method based on edge detection and Fisher discriminant for the steel surface defects of ship plate was used. In [2] at first a gradient operator detects the edges of defections. Then grayscale of gradient image enhances for more accurate edges. Secondly, Fisher discriminant is adopted in order to find optimum threshold, to segment the defects and then the noise is also filtered by morphology method. In [3] Gabor wavelets are applied to detect the defects in fabric images. Defects can be automatically segmented from the regular texture by applying the Gabor wavelets and Proper thresholding. In [4] a solution to the problem of defect detection on semiconductor wafer-die images is proposed. The segmenting of defects is done by using wavelet transformation and morphology-related properties of the associated wavelet coefficients. In [5] morphological multistage watershed segmentation is used for detection of flaws in radiographic weld images. First, watershed transform is applied to an X-ray image and the resultant mosaic image pattern is further thresholded by Otsu's thresholding method and converted into the binary image. Then top-hat of morphology is applied on the binary image to separate partially overlapping objects. Euclidean distance map is calculated for each basin to label resultant segments uniquely and to separate ridges. This follows the second stage of watershed segmentation to obtain better-defined boundaries while removing over-segmented regions.

In present work a digital camera takes the image of each pallet during its motion on the rail sequentially. Our main purpose in this paper is to segment the pallet within each frame of the image (Classification of

Proceedings of ICEE 2010, May 11-13, 2010
978-1-4244-6760-0/10/$26.00 ©2010 IEEE

defects will be present in another paper). The paper continues as following: Part 2 gives the specification of the captured pallet images. Part 3 discusses the essential theories used in this paper. Part 4 presents our proposed algorithms for segmentation of the pallet in each image frames and at least in section 5 the results and discussion is presented.

II. The Specification of the Captured Pallet Images

Our dataset is composed of 303 images captured by a digital camera. The images are generally taken in similar conditions under usual changes of the environment, change of illumination value, change of direction of illumination and usual dust noise.

Each image has 2448×3264 pixels and consists of three parts as represented in Fig. 1. Among three parts, pallet area has 1400×3264 pixels. The pictures are captured in RGB color space.

III. Primitives

We will use an example-based learning approach where a model of an object class is derived implicitly from a training set of examples. Example based algorithm have two stages: feature extraction and classification. Statistical parameters, edges, corners, colors, frequency domain information and wavelet coefficients are among the features which are extensively used for object detection. In this paper we propose to incorporate wavelet transform in YUV color space.

A. YUV color space

The YUV is chosen as a standard for TV monitors in European countries. In YUV color space the Y component shows the imaging luminance and U and V together represent chromaticity. The YUV color space may be transform into YIQ and YCbCr by a determined rotation.

The reason upon which we have chosen the YUV color space for detection and segmentation of the pallet in each frame is as following:

1) The luminance component(Y) of YUV is independent of the color, so can be adopted to solve the illumination variation problem [6].

2) In YUV color space, U and V components represents the difference between component Y and blue and red colors respectively [7].

3) The distance between the pixel colors of pallets in the images and yellow is small and the color of pallet boarders with its background is close to black. So the YUV color space successfully discriminates the pallet from its background.

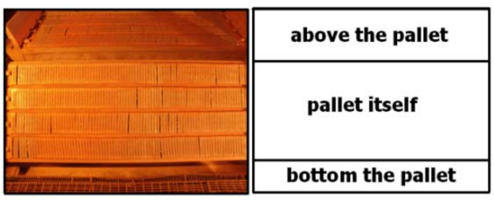

Figure 1. Each image frame has three parts.

4) Computational complexity of transforming RGB to YUV is small [6].

We convert RGB images to YUV by the following transformation [7]:

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1)$$

$$where \;\; 0 \le R \le 1, 0 \le G \le 1, 0 \le B \le 1$$

The U component in YUV color space represents the difference between blue color and luminance. This idea is proved as below:

$$Y = 0.299R + 0.587G + 0.114B$$

$$U = -0.147R - 0.289G + 0.437B$$

$$Y + 2.03U = (0.299 - 0.298)R + (0.587 - 0.586)G + (0.114 + 0.887)B = 0.001R + 0.001G + 1.001B \cong B$$

$$2.03U = B - Y \;\rightarrow U = 0.49(B - Y) \quad (2)$$

The hue of images captured from pallets is close to yellow. We know that yellow is complement of blue color. The U component also consist yellow color information. This is why we have chosen YUV color space for our algorithm. Among the three components of YUV color space the U component not only consist of the useful image information like edges, boundaries and hue of color, it also has a compression rate of three in comparsion to the YUV image.

To show the above mentioned claim we transform the captured images from RGB to CMY and consider the histogram of the three components statistically. Fig. 2, 3 and 4 shows the average of histograms of 100 images. It can be seen that concentration of yellow pixels is located at high intensity like 230 while concentration of C and M components happens at 14 and 135 respectively.

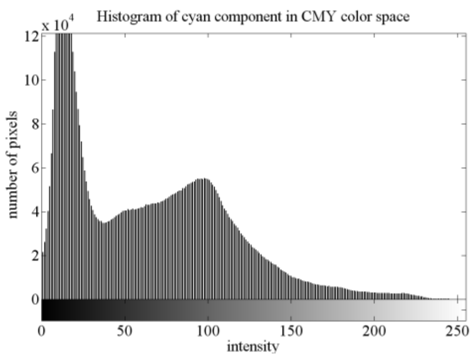

Figure 2. The average of histograms of 100 images for cyan component in CMY color space

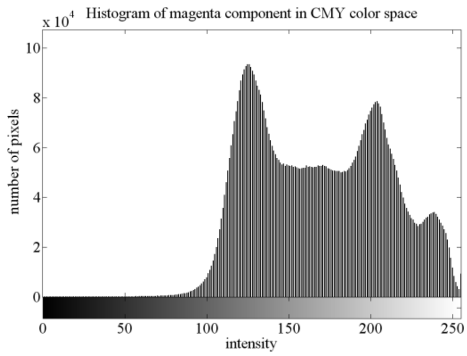

Figure 3. The average of histograms of 100 images for magenta component in CMY color space

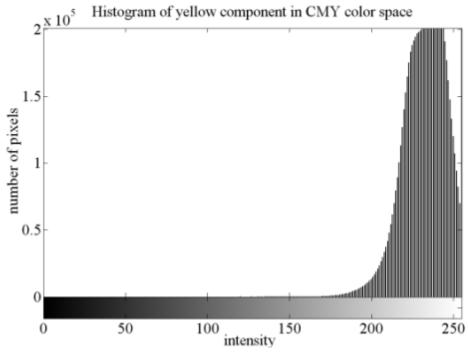

Figure 4. The average of histograms of 100 images for yellow component in CMY color space

B. *Using wavelet transform for segmentation*

Fig. 5 shows the image of a pallet and a part of one horizontal line at below of it. For successful segmentation our aim is to filter small fluctuations and preserve large changes.

We may compute high level decomposition wavelet coefficients of the signal. For instance consider the signal in Fig. 6, the following procedure extracts its main changes and omits small fluctuations: compute haar wavelet coefficients of the signal in a high decomposition level (for instance in level5) and then scale the detailed coefficients d-5 to the main signal length. This procedure eliminates small fluctuations of the signal and maintains the large changes of it [8].

This is similar to incorporating nonlinear diffusion filtering (NDF) to the image. To compare wavelet transform and NDF we mention that NDF has four different parameters which must be chosen according to the image properties. The computational complexity of NDF is also high in comparsion with that of wavelet transform [9].

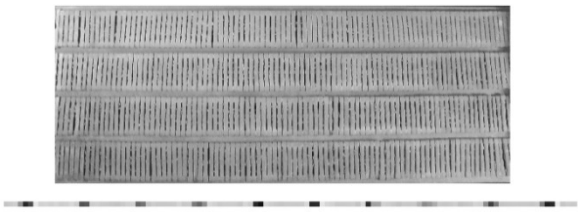

Figure 5. A part of one horizontal line of a pallet image

In Fig. 7 the two different kinds of edges with their noisy variations are shown. The frequency spectrum of the noise is high for both of the edges. If we desire to detect the edges it is evident that if the detailed coefficients of edges in (1) for instance occupies the frequency band [π/16, π/8], the detail coefficients of edges (2) occupies the next dyadic band [π/8, π/4] or higher bands. This may be explained according to Fig. 8. The nature of pallet edges is such that the output d-5 preserves large changes of the signal and eliminates the high frequency components of it.

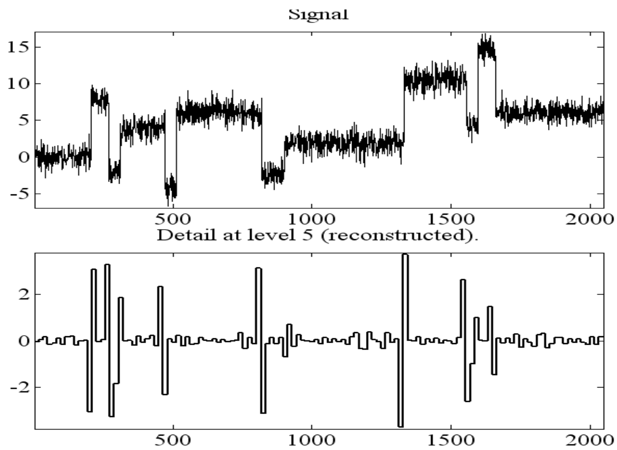

Figure 6. noisy signal and its extracted main changes

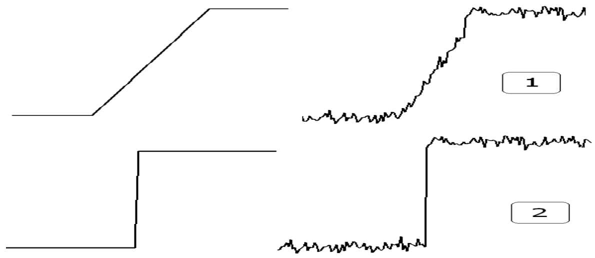

Figure 7. Two noisy edges with different frequency band.

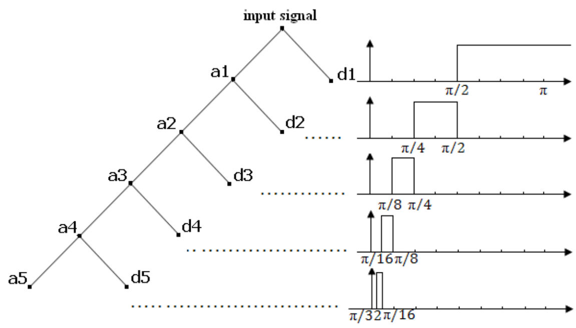

Figure 8. Decomposition tree and frequency bands of details.

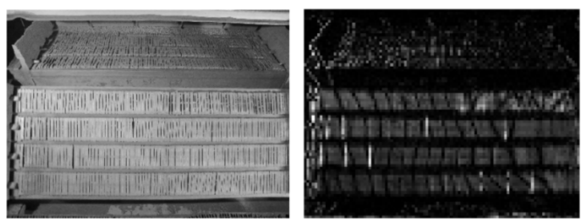

Figure 9. The main image and its level 5 detail component in the vertical direction.

According to above explanation to extract the grate -bars in a pallet image we decompose its image by haar wavelet to level 5 and selected the detail component in the vertical direction and then use it as one of our features vectors for classification. The main image and its level 5 detail component in the vertical direction are shown in Fig. 9.

C. Support Vector Machine (SVM)

SVM is a supervised learning technique for classification. One of the main attractive features of using SVMs is that it is capable of learning in high-dimensional spaces with very few training examples. Using the SVM formulation, the general form of the decision function for a point x is [10]:

$$f(x) = \theta\left(\sum_{i=1}^{l} \alpha_i y_i K(x, x_i) + b\right) \tag{3}$$

where l is the number of training data points, α_i are Lagrange parameters obtained in the optimization step and $\theta(.)$is a threshold function, and K(.,.) is a kernel that defines a dot product between projections of the two arguments in some feature space where a separating hyperplane is then found. The kernel K can be a Gaussian (Radial Basis Functions), or an nth degree polynomial$(x.y + 1)^n$, and is generally any positive definite function [10].

The main feature of SVM is that it finds, among all possible separating surfaces of the "Equation (3)", the one which maximizes the distance between the two classes of points. The support vectors are the nearest points to the separating boundary and twice the distance of a support vector from the boundary is called margin. The margin is an important quantity because it can be taken as an indicator of the separability of the data and consequently of the "goodness" of the representation used. It is possible that different input representations and/or kernel functions lead to quite different geometries of the data (i.e. different margins), which in turn influences the performance of the SVM [11].

IV. METHODS

We divided our dataset in two groups of 150 and 153 images. The first group was used in training procedure and the second group was used in testing. As it was mentioned before there are three different parts in each of the images.

In our algorithm, in each image the part which contains the pallet itself is considered to be a positive pattern and the other excessive parts (usually two or more) are considered as negative. So in learning stage among the first group we select the 150 pallet sections as positive patterns and 450 excessive parts as negatives patterns.

Now we have to define features which separate positive patterns from negative patterns. In the following we present some of the main features that we have used them for segmentation of pallets.

A. Signed and Unsigned WT coefficient of gray-level image of pallets

The wavelet coefficients which are the result of decomposition in the vertical direction of pallet images in level 5 are more useful for image analysis. These coefficients and their absolute values which solely describe the strength of the intensity differences are used as features.

B. PCA approach on WT coefficient

The PCA of WT coefficients of positive and negative patterns are computed and their eigenvectors and eigenvalues are found. Then the projection of the WT coefficients of images on the computed eigenvectors are chosen as features.

C. WT coefficient of color image of pallets

Color images contain more information than grayscale images. We use RGB and YUV color space and apply WT on each component of these spaces separately. Then for each space we compute the energy of detail coefficients of level 5 decomposition for three components and select the largest.

D. hough transform on color image of pallets

As another feature we transform the RGB to YUV color space. Then for each of Y, U and V images the following steps is done:

- An edge detector operator, sobel, detects the edges.
- The output of previous step is binarized by Entropic Thresholding algorithm.
- Combine the three binary images to obtain one binary image.
- Compute the hough transform of the binary image.
- Detect the main horizontal line edges of the image according to peak point of hough transform image.
- Use these horizontal lines as features.

E. Integral of horizontal projection of gray-level image

Consider the arrangement of grate-bars and symmetry of shapes. Another feature is integral of horizontal projection (IHP) of images. To compute IHP we add the values of all pixels of each row and is displayed as a number. The output is a vector whose length is equal to number of rows. This vector is shown in Fig. 10 for a sample image.

F. IHP of WT coefficient of gray-level image

IHP of WT coefficient at level 5 of decomposition is obtained. As it is shown in Fig. 11 there are four peaks each of which is corresponds to one row of grate bars in the pallet.

G. IHP of WT coefficient of U component in YUV color space

The procedure in section 4-6 is exactly done on U component of the image and its output is also considered as a feature.

V. RESULTS

SVM classification is used for learning, the number of SVM classes is chosen to be two and the kernel is Gaussian. TABLE I presents the list of features defined in previous section. These features are measured and fed as input to SVM. The results of classification are given by Receiver Operating Characteristic (ROC) curves. Then the areas under ROC curves (AUC) are presented in column two of TABLE I. The run time is shown at column three. The numbers of AUC column show the performance of each feature in detecting and classifying the pallet part of each image. The greater the AUC is the higher the performance. It is evident that integral of horizontal projection (IHP) of WT coefficients of U component in YUV color space is the most efficient feature for detection and classification of pallets. IHP of WT coefficients of grayscale image and hough transform of YUV color space have the next ranks. The ROC curves for theses three features is shown in Fig. 12.

We conclude that YUV color space is preferred to grayscale and the U component alone is more reliable for pallet detection and classification. Also absolute values of WT coefficients are preferred to normal WT coefficients. In Fig. 13 the result of segmentation algorithm on a pallet image is shown.

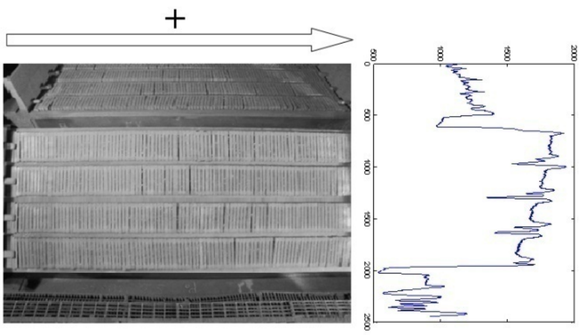

Figure 10. Plot of IHP for a sample image.

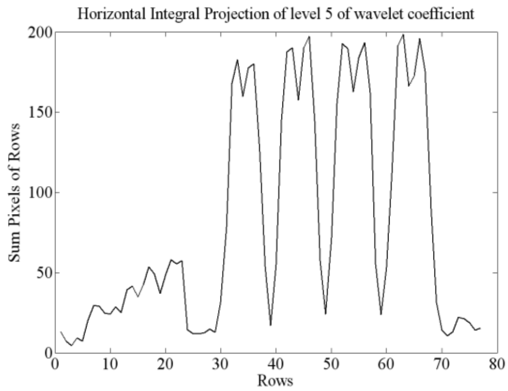

Figure 11: IHP of WT coefficient of pallet image

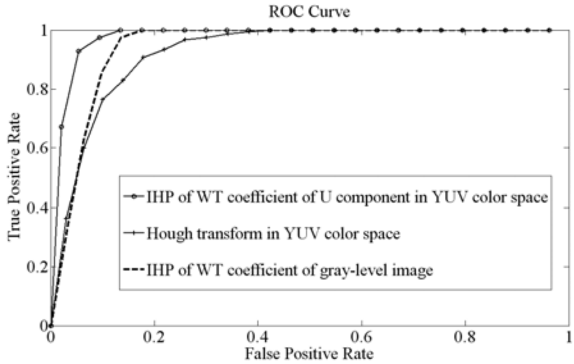

Figure 12. ROC curves of pallet detection to compare different features

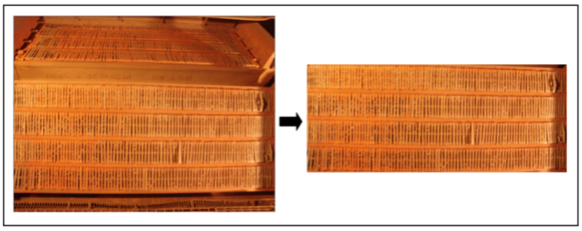

Figure 13. result of segmentation algorithm on a pallet image.

TABLE I: Compare efficiency of different features

feature	AUC	Run time
Signed WT coefficient of gray-level image	0.6691	105.1950s
Unsigned WT coefficient of gray-level image	0.9391	20.3822s
PCA approach on WT coefficient	0.7342	17.4474s
WT coefficient of RGB color space	0.7857	287.9549s
WT coefficient of YUV color space	0.9474	289.2658s
hough transform in YUV color space	0.9595	62.4215s
IHP of gray-level image	0.9240	27.1676s
IHP of WT coefficient of gray-level image	0.9619	8.2238s
IHP of WT coefficient of U component in YUV color space	0.9874	56.0123s

VI. CONCLUSION

We used of an example-based approach for pallet detection and segmentation. It has two main stages: feature extraction and classification. Whereas the pallets have high resolution details in their structure, we applied wavelet transform as a multi-resolution technique. Also based on the color nature of pallets we used YUV color space. We compare different features and concluded that the integral of horizontal projection (IHP) of WT coefficients of U component in YUV color space is the most efficient feature for detection and classification of pallets. This feature is fed as input to SVM for classification.

Acknowledgements

The authors appreciate from pelletizing unit management and its personnel for their cooperation in taking images of the pallets. Also we respect international IRISA company for their sponsorship for the project.

References

[1] D. Mery and D. Filbert, "Automated flaw detection in aluminum castings based on the tracking of potential defects in a radioscopic image sequence," *IEEE Trans*. Robotics and Automation, vol. 6, pp. 890–901, August. 2002.

[2] J. H. Guo, X. D. Meng and M. D. Xiong, "Study on defection segmentation for steel surface image based on image edge detection and fisher discriminant," 2006 J. Phys.: Conf. Ser. 48, pp. 364–368.

[3] S. Arivazhagan, L. Ganesan and S. Bama, "Fault segmentation in fabric images using Gabor wavelet transform," Machine Vision and Applications, vol. 6, pp. 356–363, 2006.

[4] N. G. Shankar and Z. W. Zhong, "Improved segmentation of semiconductor defects using area sieves," Machine Vision and Applications, vol. 1, pp. 1-7, 2006.

[5] Alaknanda, R. S. Anand and P. Kumar, "Flaw detection in radiographic weldment images using morphological watershed segmentation technique," NDT&E International, pp. 2–8, 2009.

[6] Ch. Lin, "Face detection in complicated backgrounds and different illumination conditions by using YCbCr color space and neural network," Pattern Recognition Letters, vol. 28, pp. 2190–2200, 2007.

[7] A. Koschan and M. Abidi, Digital color image processing, Canada: Wiley-Interscience, 2008, pp. 50-51.

[8] S. Mallat, A wavelet tour of signal processing, USA: Elsevier, 1998.

[9] P. Mrazek, "Nonlinear diffusion for image filtering and monotonicity enhancement," Ph.D. thesis, Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University, Czech Republic, 2001.

[10] C. Papageorgiou and T. Poggio, "A trainable system for object detection," IJCV, vol. 38, pp. 15-33, 2000.

[11] S. Abe, Support vector machines for pattern classification, London: springer-verlag, 2005, pp. 3-21.