

An X-ray Inspection System for Illegal Object Classification based on Computer Vision

Yu Wang², Xiaomin Yang^{1*}, Wei Wu¹, Bingshan Su¹ and Gwanggil Jeon³

¹*School of Electronics and Information Engineering, Sichuan University,
Chengdu, Sichuan, 610064 P.R.China*

²*Aviation Engineering Institute, Civil Aviation Flight University of China,
Guanghan, Sichuan, 610064 P. R. China*

³*Department of Embedded Systems Engineering, University of Incheon, 12-1
Songdo-dong, Yeonsu-gu, Incheon406-772, Republic of Korea
{wuwei, arielyang}@scu.edu.cn; {2307574, 625864564}@qq.com;
gjeon@inu.ac.kr*

Abstract

Security checks at airport are of importance to any safe flight. Traditional method for security check is to check the luggage manually. However, this method needs lots of human labor and time. It is desired to automatically check x-ray image of the luggage with computer vision. In this paper, the illegal object classification system is implemented. First, we introduce a computer vision based x-ray inspection system for Illegal object classification. Then we propose a method by combining Taruma feature based on Contourlet transform and histogram. Finally, we apply the random forests classifier to classify these features from the illegal objects. Experimental results show that the proposed method can effectively distinguish different kinds of illegal objects.

Keywords: *Illegal object classification, X-ray image, Contourlet transform, Taruma texture, Random forest*

1. Introduction

Airport, which is a public use facility, is very important for commercial flights. Security checks at airport are of importance to any safe flight [1]. One important tasks of each security check is to detect and classify illegal objects in the luggage. This would put great pressure on custom. Traditional method for security check is to check the luggage manually. However, this method needs lots of human labor and time. An alternative method is to check x-ray image of the luggage, which is obtained with a x-ray generator, by human rather than to open the luggage and check it manually. Since it is not necessary to open the luggage, lots of human labor and time can be saved. Therefore, to check x-ray image of the luggage instead of to open the luggage is a good solution for security check.

X-radiation is a form of electromagnetic radiation. The absorption coefficient and the thickness of a material determine the x-ray absorption of an object. Due to the penetrating ability of x-rays, x-rays are widely used in various applications, which include but not limited to: airport security, medical uses, and computed tomography. Based on the penetrating ability of x-rays, the inside objects in luggage can be easily obtained with x-ray image. Compared with traditional human labor based method, which is to open the luggage and check the luggage manually, the x-ray image based method is much fast, since opening luggage is not necessary. Figure 1 shows the two examples of X-ray images and their corresponding images.

* Corresponding Author

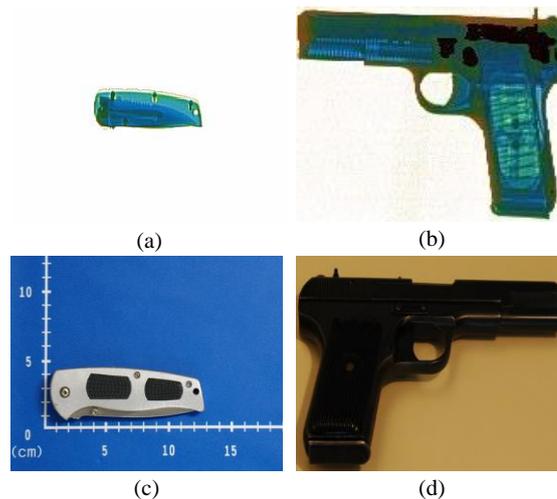


Figure 1. X-Ray Images and their Corresponding Visible Images (a) Visible Image of Knife; (b) x-Ray Image of Knife; (c) Visible Image of Pistol; (d) x-Ray Image of Pistol

However, existing inspection system requests that a worker must keep high concentration on the screen to check the x-ray image of the luggage. It is a very hard work for any worker, who would make a mistake when he is distracted. With the development of computer vision, it is desired to check or classify x-ray images of different illegal objects automatically.

In [2], the authors proposed the use of a method based on multiple x-ray views to detect some regular prohibited items with very defined shapes and sizes. The method consists of two steps: structure estimation and parts detection. In [3], the authors stated that security check task in the airport consists in selection of infrastructure, technical equipment, allocation of personnel, and financial means that are necessary to perform all functions of the security check task. Then they proposed a fuzzy system to support the configuration of the whole security check system. In [4], the authors extended the previous work by exploring the effect of attribute selection on the classification performance. In [5], the authors proposed to describe the cargo x-ray image by joint shape and texture feature, which could reflect both interior details and cargo stacking mode. In addition, the authors discussed x-ray image property and explore some reasons why cargo classification under x-ray is very difficult.

Though, many efforts have been made on the x-ray image processing, most of those efforts mainly focus on x-ray image enhancement and medical image recognition [4]. There are few researches on x-ray image recognition with computer vision for the security check. This paper deals with x-ray image recognition with computer vision for illegal objects recognition.

The main work of this paper is to propose a prototype of illegal objects recognition system based on x-ray image automatically for civil aviation airport. In addition, we proposed a method by using contourlet transform [6, 7, 8] and Tamura feature [9, 10, 11] to extract the features of x-ray image. This work not only would provide the help for illegal objects recognition system with x-ray image, but also provides a reference for the automatic identification of irregular objects.

Since illegal objects maybe in various shapes, it is very difficult to employ a knowledge model to extract the features of an illegal object. To solve this problem, in this paper we extracted texture features of illegal objects, since texture features have the properties inherent in the objects, and is widely used in image analysis and recognition. Specifically, contourlet transform [6] and Tamura feature [11] are adopted in this paper.

Contourlet transform [6, 7, 8], which is with the properties of directionality, multi-scale, and anisotropy, can effectively capture the edge information of the image comparison with wavelet transform. Thus, contourlet transform is widely used to extract the transform coefficients maps. In addition, this paper extracted Tamura features of the transform coefficients maps, since the Tamura texture features [9, 10, 11] has similar texture model with human visual system. By using contourlet transform and Tamura texture features, a recognition system may have a better performance. Furthermore, image histogram feature was extracted, since image histogram feature is not affected by image rotation and translation. Combining Tamura texture feature based on contourlet transform and image histogram feature, we can obtain the final feature of illegal objects.

Compared with other classifiers, the random forests [12, 13, 14] have fewer parameters and strong classification ability, which is very suitable for engineering application. Therefore, this paper employed the random forests to classify the illegal objects by using the extracted features.

The structure of this paper is as follows. Section 2 introduces the framework of computer vision based x-ray inspection system. Then, Sections 3 and 4 present the feature extraction and classification respectively. The experimental results are shown in in Section 5. Section 6 give summary of this paper.

2. Computer Vision Based X-ray Inspection System

In this paper, computer vision based x-ray inspection system is proposed. The system automatically recognizes the categories of illegal objects of the x-ray image. The block diagram of the system is shown in Figure 2.

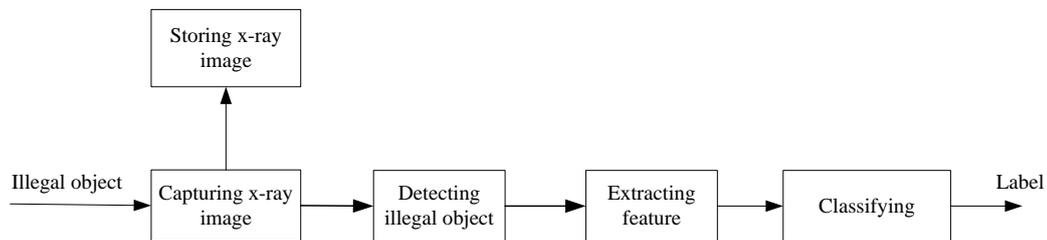


Figure 2. The Block Diagram of Computer Vision based X-ray Inspection System

The whole system includes four steps: image acquisition, illegal object localization, feature extraction, and classification. The first step is to capture the x-ray images with an x-ray generator. The second step *i.e.*, illegal object localization step is to detect the illegal objects in the x-ray image. The third step is to extract the feature of the illegal objects. Specifically, we extracted the feature of x-ray image by contourlet transform, Tamura texture and image histogram features. We first performed contourlet transform [7] to obtain the transformation coefficient maps of x-ray image. And then we extracted the Tamura features from transformation coefficient maps [7]. Also at the same time, we extracted image histogram feature. Finally, we combined all those features to form a feature vector. The last step is to classify the illegal objects. In this step, the feature vector was feed into random forest [12, 13, 14] to obtain the final results.

It should be noted that this paper mainly exploit the last two steps, *i.e.*, feature extraction and classification. Though illegal object localization is a very critical issue in the inspection system, the location of the illegal object is beyond the scope of this paper.

3. Feature Extraction

For an image recognition or classification system, feature extraction generates features, which intend to be informative and non redundant. Generally, the extracted features has a lower size than original image. This means that feature extraction is related to dimensionality reduction. The main purpose of feature extraction is to help to improved the performance of the subsequent classification. It has very important significance for a recognition system. In this paper, three techniques, which include contourlet transform, Tamura texture, and image histogram, are adopted to extract the features of illegal objects. The following subsections will describe the details of those three techniques.

3.1. Contourlet Transform

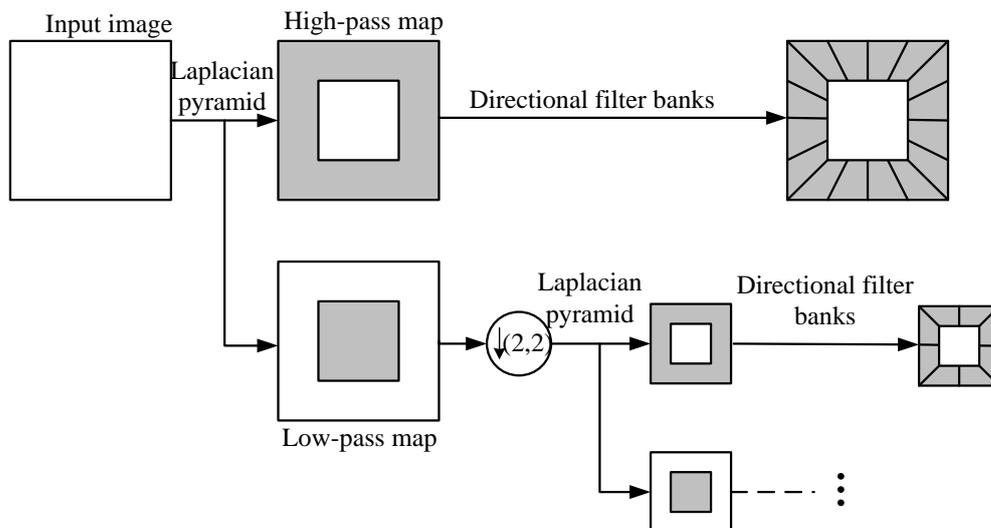


Figure 3. The Structure of Contourlet Transform

Wavelet transform has been widely used in many image processing applications. Wavelet is not capable of capturing edges along other directions, since wavelet could only catch vertical and horizontal edges in an input image. To solve this problem, contourlet transform rather than wavelet transform is employed in this paper. Contourlet transform, which is was proposed in [6], is similar with the human visual system. Contourlet transform stems from curvelet transform, which can capture the smoothness images in different directions. Thus, contourlet transform could extract the image information at different orientations and at multiple scales.

The contourlet transform consists of two steps: the multi-scale decomposition and the directional decomposition. The Structure of contourlet transform is as shown in Figure 3.

In the multi-scale decomposition step, Laplacian pyramid is performed to decompose the input image into a low-frequency and a set of high-frequency maps. In the directional decomposition step, directional filter banks (DFB) are employed to obtain the high-frequency sub-image. After performing contourlet transform, multiscale and multi-orientation maps can be obtain from an input image. Figure 4 shows an example of two-level decomposition of contourlet transform for a x-ray image of stun gun.

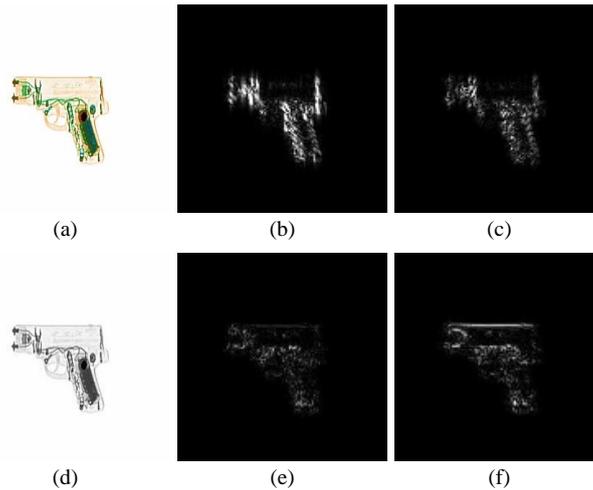


Figure 4. An Example of Contourlet Transform for a X-ray Image of Stun Gun (a) the Original X-Ray Image of stun gun; (d) the Intensity Image of (a); (b), (c), (e) and (f) are the High Frequency Coefficients Maps after Performing Contourlet Transform

3.2. Tamura Texture Features

The texture feature is one of the properties inherent in the image. It is related to the properties of the object, and it reflects the changes of intensities of image. The texture feature is widely used in image analysis and recognition. We extracted Tamura features from the contourlet transform coefficient maps, since the Tamura texture features [11] have strongly similar psychological measurements with human visual system,

Specifically, the Tamura features [11] include six features *i.e.*, coarseness, contrast, directionality, line-likeness, regularity, and roughness. Among six features, direction, roughness and contrast, which are particularly important for feature extraction, are employed to represent contourlet transform coefficient maps of the x-ray image. The following paragraphs will give the details of directionality, roughness and contrast [11].

A. Directionality

Directionality describe the total degree of concentration along certain orientations of an image globally rather than locally. To calculate the directionality, we needed to calculate the modulus and edge directionality of each pixel. The modulus and directionality can be calculated as:

$$|\Delta G| = \frac{(|\Delta_x| + |\Delta_y|)}{2} \quad (1)$$

$$\theta = \tan^{-1} \left(\frac{\Delta_x}{\Delta_y} \right) + \frac{\pi}{2} \quad (2)$$

where ΔG and θ refer to modulus and directionality of a pixel respectively. Δ_x and Δ_y refer to the horizontal and vertical gradients respectively. Δ_x and Δ_y can be computed by convolution of the image with the following two 3×3 operators.

$$\begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \quad (3)$$

After computing all the directionality θ of each pixel, we quantized directionality θ . Then, we counted all the pixels with $\Delta G > t$, where t is a predefined threshold. Subsequently, a histogram can be built with the number of the pixels satisfying $\Delta G > t$.

The histogram can be calculated as:

$$H(k) = \frac{N_{\theta}(k)}{\sum_{i=0}^{n-1} N_{\theta}(i)}, \quad k = 0, 1, \dots, n-1 \quad (4)$$

where n refers to the quantization level. $N_{\theta}(k)$ is the number of the pixels satisfying $\Delta G > t$ at each level k .

The directionality can be calculated with the sharpness of the histogram. This directionality can be expressed as:

$$F = \sum_p^{n_p} \sum_{\theta \in \omega_p} (\theta - \theta_p)^2 H_D(\theta) \quad (5)$$

where ω_p represents all the bins which include it, and θ_p is the bin, which has the highest peak value. And p is the peak value of histogram, n_p are all the peak values of the histogram.

If a histogram is with a obvious peak, *i.e.*, high value of F , this demonstrates that the image is an image with obvious direction. Similarly, If a histogram is very flat, *i.e.*, high value of F , this illustrates that the image is an image without obvious direction.

B. Coarseness

The coarseness measures the coarseness of the image. Specifically, the coarseness is to assess the size of the texture elements. Generally, if an image is very rough in the texture, the image would has a big coarseness value. On the contrary, the image would has a small coarseness value, if an image is smooth in the texture.

To calculate the roughness, the average intensity value of local area of size $2^k \times 2^k$ centered by each pixel is calculated. This can be expressed as:

$$A_k(x, y) = \frac{\sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} g(i, j)}{2^{2k}} \quad (6)$$

where $k = 1, 2, \dots, 5$. And $g(i, j)$ denotes the intensity value of a pixel at position (i, j) .

Then, for each pixel, on horizontal and vertical directions we can obtain the difference between the average intensity values of two neighbor areas, which are not overlapped with each other. This can be defined as:

$$\begin{cases} E_{k,h}(x, y) = |A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y)| \\ E_{k,v}(x, y) = |A_k(x, y + 2^{k-1}) - A_k(x, y - 2^{k-1})| \end{cases} \quad (7)$$

For each pixel, we selected the best size $S_{best}(i, j)$, which gives the highest output value. The best size can be calculated by:

$$k = \arg \max_{k=1,2,\dots,5} \max(E_{k,v}(i, j), E_{k,h}(i, j)) \quad (8)$$

$$S_{best}(i, j) = 2^k + 1 \quad (9)$$

Finally, we took the average over the best size for each pixel as a coarseness measure to the image:

$$F_{crs} = \frac{1}{M \times N} \sum_{i=1}^m \sum_{j=1}^n S_{best}(i, j) \quad (10)$$

where $M \times N$ is the size of the image.

The core idea of coarseness is to measure the coarseness value of each pixel with various sizes. Then, the average value of coarseness over the whole image is computed as the coarseness of the whole image.

C. Contrast

Contrast measures the intensity distribution over the whole image. The contrast of an image is calculated by

$$F_{om} = \sigma / \alpha_4^{1/4} \quad (11)$$

$$\alpha_4 = \mu_4 / \sigma^4 \quad (12)$$

where μ_4 is the fourth moment, and σ^2 is the variance of the intensity of the image.

3.3. Image Histogram

An image histogram consists of many bins. The histogram is calculated by assigning each pixel in the image to a bin depending on the pixel intensity. Each pixel in the image is assigned to its corresponding bin. After processing all the pixels in the image, the histogram can be obtained. Image histogram is a global feature for image. The advantage of this feature is that it is not affected by image rotation and translation. Also it is not only insensitive to image scale, but also insensitive to the spatial location of the object in the image. Especially, image histogram feature is suitable for object, which is difficult to be automatically segmented. Figure 5 and Figure 6 are two examples, which illustrate the discriminability of image histogram feature. Figure 5 is a comparison of two histograms of illegal objects from different types. From Figure 5, we can observe that different types of illegal objects have totally different histograms. Figure 6 is a comparison of different histograms of illegal objects from same type. Different with Figure 5, from Figure 6 we can find that same type of illegal objects has similar histograms. All this demonstrates that the histogram feature has good discriminability for classification.

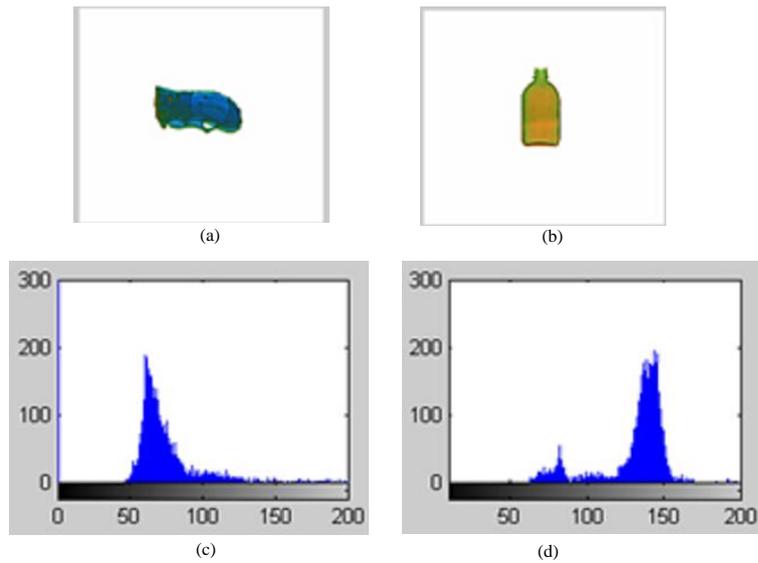


Figure 5. A Comparison of Different Histograms of Illegal Objects from Different Types: (a) x-Ray Image of a knife; (b) x-ray image of a bottle of acetone; (c) Histogram of (a); (d) Histogram of (b)

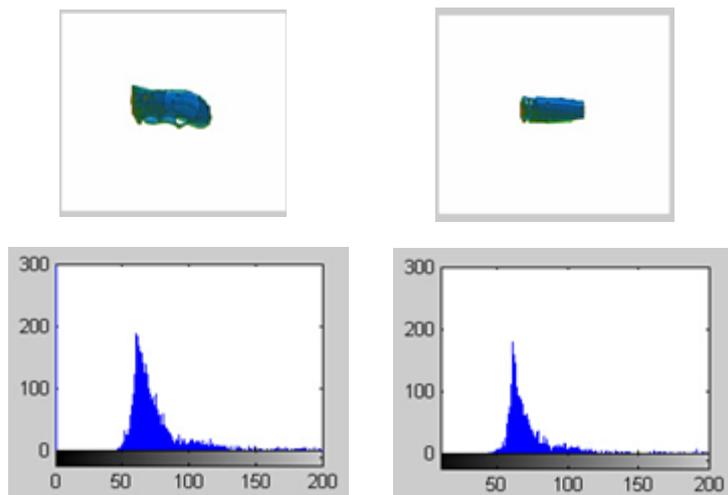


Figure 6. A Comparison of Different Histograms of Illegal Objects from Same Type: (a) x-Ray Image of a knife; (b) x-Ray Image of Another Knife; (c) Histogram of (a); (d) Histogram of (b)

4. Classification

With the development of the statistical theory, many classifiers, which include, but not limited to, SVM, random forests [14], have been proposed in recent years. Compared with other classifiers, random forests have fewer parameters and strong classification capabilities. This leads to that random forests are very suitable for engineering applications. Thus, this paper used random forests to classify the extracted features. Random forests is an ensemble classification method, which usually demonstrates superior accuracy compared to those of single classification methods.

Generally, there are two types of off-the-shelf ensemble classification methods. One type of method is AdaBoost, another is bagging. The bagging method adopts some

randomization heuristic methods to build independently of one another. Random forests, which consist of L decision trees, is a bagging method. Each tree in random forests is independently built by using a different set of extracted features from train set. Random forests have the following advantages:

- (1) they can produce highly recognition rate.
- (2) the training and testing processes need a small amount of calculation.
- (3) they can evaluate the importance of each feature variable.

A framework of random forests is shown in Figure 7.

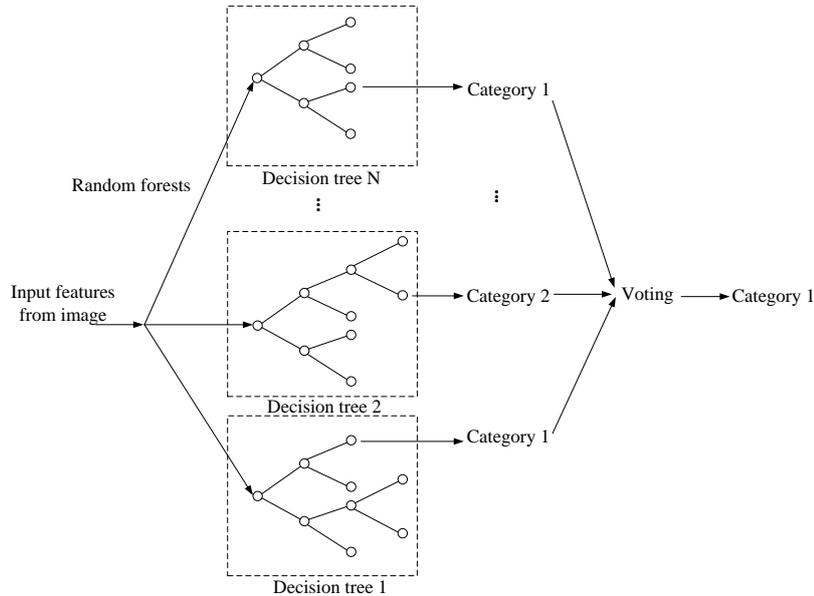


Figure 7. A Framework of Random Forests

5. Experimental Results

To verify the effectiveness of the proposed method, 276 x-ray images of illegal objects are used in our experiments. These x-ray images include four categories of illegal objects, *i.e.*, guns, ammunitions, edge tools, and inflammable and explosive objects. These images are randomly divided into two parts. One part, which includes 117 images, is for training. The rest, which includes 159 images, is for testing. All images are normalized to same resolution $256 * 256$. Part of the x-ray images is shown in Figure 6.

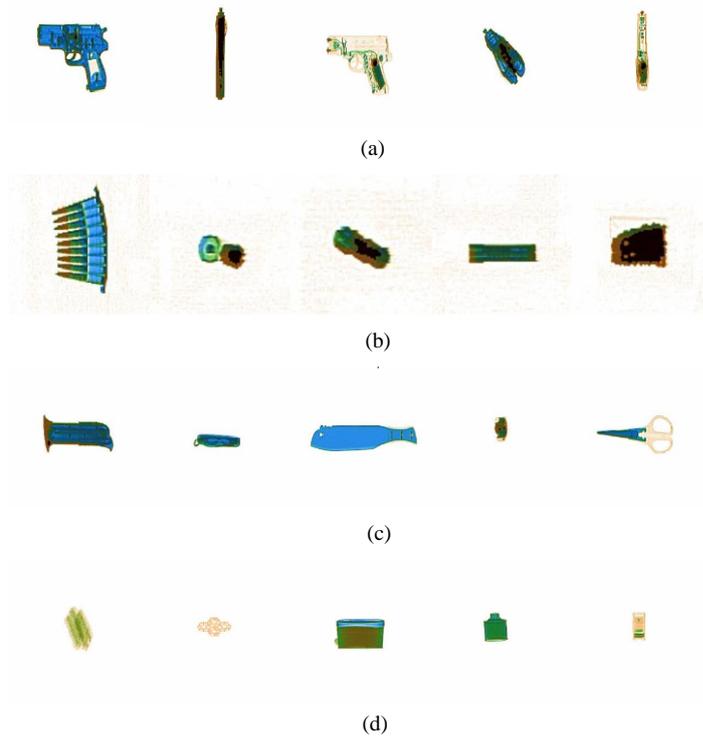


Figure 8. Part of X-ray Images of Illegal Objects: (a) x-ray Images of Guns; (b) x-ray Images of Ammunitions; (c) x-ray Images of Edge Tools; (d) x-Ray Image of Inflammable and Explosive Objects

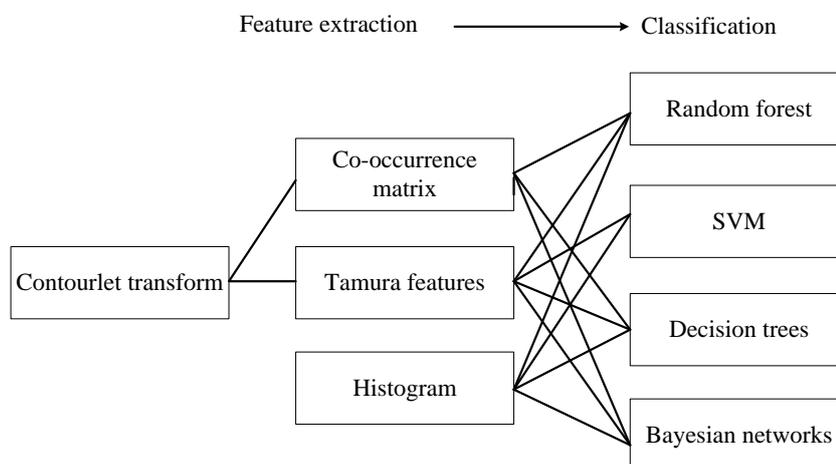


Figure 9. The Experiment for Feature Extraction and Classification

Both the feature extraction methods and the classifiers have impact on the performance of the proposed method. In order to analyze the performance of the proposed method, different feature extraction methods and different classifiers are combined, as shown in Figure 9. Feature extraction methods include co-occurrence matrix features, Tamura texture features, and image histogram features. The classifiers

used in the experiment include the random forests, Bayesian networks, decision trees, and SVM. Table 1 summarizes the experimental results.

Table 1. The Recognition Rates of Different Method

Feature Extraction	Classification results (correct classification rate%)			
	Random Forest	Decision Tree	Bayesian networks	SVM
Contourlet transform and co-occurrence matrix	71.07	68.55	65.40	75.47
Contourlet transform and Tamura feature	74.84	62.89	66.67	67.92
Histogram feature	79.24	76.10	76.73	57.86
Combination feature	84.28	79.24	81.76	57.86

The combination feature, which was used in this paper, combine histogram feature and Taruma feature based on contourlet transform. From Table 1, we can find that the random forests classifier achieves the better results than other classifiers when using any feature. This illustrates that the random forests classifier has a better performance than other classifiers. The reason is that the ensemble methods generally have a better performance than single classifier methods. The random forests classifier is a typical the ensemble method, while other classifiers are single classifier methods.

In addition, we can see that the combination feature has a better performance compared with other features on any classifier. This demonstrates that the combination feature is the best feature among the compared features. The reason for this is that the combination feature combines the advantages of both histogram feature and Taruma feature based on contourlet transform.

6. Summary

Security checks at airport are of importance to any safe flight. Traditional method for security check is to check the luggage manually. However, this method needs lots of human labor and time. It is desired to automatically check x-ray image of the luggage with computer vision. In this paper, the illegal object classification system is implemented. Firstly, we performed contourlet transform to obtain contourlet transform coefficient maps. Then, Tamura feature was extracted from those coefficient maps. At the meantime, image histogram feature was obtained. Thus, by combining these features, we concatenated these features into a feature vector. Finally, the concatenated feature vector was feed into random forests to obtain the result. Experiments show the effectiveness of the proposed method. Our work in this paper not only provides a prototype system for the automatic recognition system for x-ray image, but also provides a reference for object recognition.

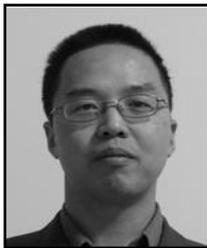
Acknowledgments

The research is sponsored by the National Natural Science Foundation of China (No. 61271330 and No. 61411140248), the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry.

References

- [1] G. Zentai, "X-ray Imaging for Homeland Security", IEEE International Workshop on Imaging Systems and Techniques, (2008).
- [2] D. Mery and V. Riffó, "Automated Object Recognition using Multiple X-ray Views", Materials Evaluation, vol. 72, no. 11, (2014), pp. 1362-1372.
- [3] J. Skorupski and P. Uchrońska, "A fuzzy system to support the configuration of baggage screening devices at an airport", Expert Systems with Applications, vol. 44, no. 2, (2016), pp. 114-125.
- [4] D. Unay, O. Soldea, A. Ekin, M. Cetin and A. Ercil, "Automatic Annotation of X-Ray Images: A Study on Attribute Selection", Lecture Notes in Computer Science, Medical Content-Based Retrieval for Clinical Decision Support, vol. 5853, pp. 97-109.
- [5] J. Zhang, L. Zhang, Z. Zhao, Y. Liu, J. Gu, Q. Li and D. Zhang, "Joint Shape and Texture Based X-Ray Cargo Image Classification", IEEE Conference on Computer Vision and Pattern Recognition Workshops 2014, pp. 266-273.
- [6] M. N. Do and M. Vetterli, "The contourlet transform, an efficient directional multiresolution image representation", IEEE Transactions on Image Processing, vol. 14, no. 12, (2005), pp. 2091-2106.
- [7] W. Wu, Z. Liu and Y. He, "Classification of Defects with Ensemble Methods in the Automated Visual Inspection of Sewer Pipes", Pattern Analysis and Applications, vol. 18, no. 2, (2015), pp. 263-276.
- [8] W. Wu, Z. Liu, W. Gueaieb and X. He, "Single-image super-resolution based on Markov random field and contourlet transform", J. Electron. Imaging, vol. 20, no. 2, (2011), pp. 023005-023005.
- [9] T. Deselaers, D. Keysers and H. Ney, "Features for Image Retrieval: An Experimental Comparison", Information Retrieval, vol. 11, no. 2, (2008), pp. 77-107.
- [10] H. Tamura, S. Mori and T. Yamawaki, "Texture features corresponding to visual perception", IEEE Transactions on Systems Man and Cybernetics, vol. 8, no. 6, (1978), pp. 460-473.
- [11] Y. Liu, Z. Li and Z. Gao, "An Improved Texture Feature Extraction Method for Tyre Tread", Lecture Notes in Computer Science, Patterns Intelligence Science and Big Data Engineering, vol. 8261, pp. 705-713.
- [12] J. J. Rodriguez, L. I. Kuncheva, and C. J. Alonso, "Rotation Forest: A New Classifier Ensemble Method", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 10, (2006), pp. 1619-1630.
- [13] A. Verikas, A. Gelzinis and M. Bacauskiene, "Mining data with random forests: A survey and results of new tests", Pattern Recognition, vol. 44, no. 2, (2011), pp. 330-349.
- [14] P. O. Gislason, J. A. Benediktsson and J. R. Sveinsson, "Random Forests for land cover classification", Pattern Recognition Letters, vol. 27, no. 4, (2006), pp. 294-300.

Authors



Wei Wu, he is currently an Associate Professor in College of Electronics and Information Engineering, Sichuan University, China. He received his BS degree from Tianjin University, China in 1998 and received his MS degree and PhD degrees in communication and information system from Sichuan University China in 2003 and 2008, respectively. He worked in a National Research Council Canada as a post-doctorate in 2009 - 2010. He has published over 50 scholarly papers in academic journals and conferences. He serves as an associate editor of Journal of Pattern Recognition Research. His research interests are image process, computer vision and pattern recognition.



Xiaomin Yang, she is currently an Associate Professor in College of Electronics and Information Engineering, Sichuan University. She received her BS degree from Sichuan University, and received PhD degrees in communication and information system from Sichuan University. She worked in University of Adelaide as a post doctorate for one year. She has published over 30 scholarly papers in academic journals and conferences. Her research interests are image process and pattern recognition.



Gwanggil Jeon, he received the BS, MS, and PhD (summa cum laude) degrees in Department of Electronics and Computer Engineering from Hanyang University, Seoul, Korea, in 2003, 2005, and 2008, respectively. From 2008 to 2009, he was with the Department of Electronics and Computer Engineering, Hanyang University, from 2009 to 2011, he was University of Ottawa, as a postdoctoral fellow, and from 2011 to 2012, he was with the Graduate School of Science & Technology, Niigata University, as an assistant professor. He is currently an assistant professor with the Department of Embedded Systems Engineering, University of Incheon, Incheon, Korea. His research interests fall under the umbrella of image processing, particularly image compression, motion estimation, demosaicking, and image enhancement as well as computational intelligence such as fuzzy and rough sets theories. He was the recipient of the IEEE Chester Sall Award in 2007 and the 2008 ETRI Journal Paper Award.

