Joint Shape and Texture Based X-Ray Cargo Image Classification

Jian Zhang[1], Li Zhang[1,2], Ziran Zhao[1,2], Yaohong Liu[1], Jianping Gu[1], Qiang Li[1], Duokun Zhang[1]
[1]Nuctech, Beijing, China
[2]Tsinghua University, Beijing, China
[1]{zhangjian1, zhangli, zhaoziran, liuyaohong, gujianping,liqiang1,zhangduokun}@nuctech.com
[2]{zli, zhaozr}@mail.tsinghua.edu.cn

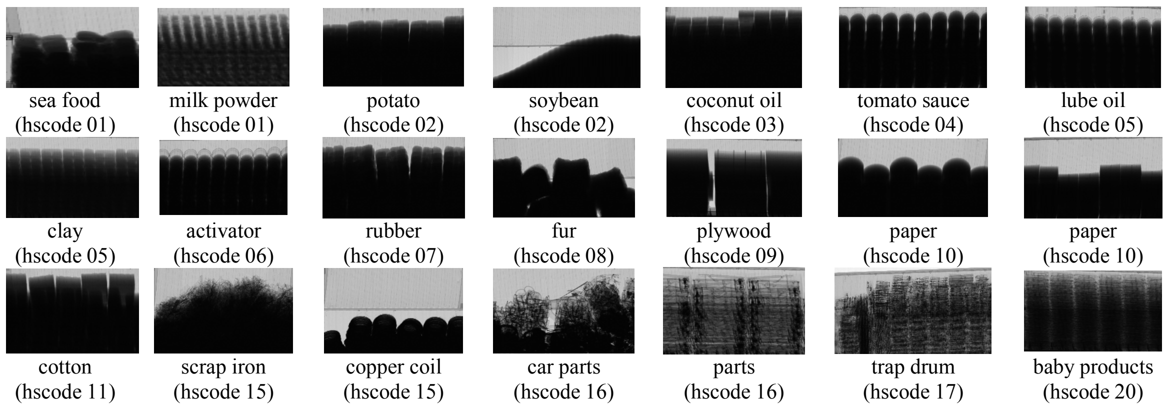


Figure 1: cargo X-Ray images

***Abstract*—Security & Inspection X-Ray Systems is widely used by custom to accomplish some security missions by inspecting import-export cargo. Due to the specificity of cargo X-Ray image, such as overlap, viewpoint dependence, and variants of cargo categories, it couldn't be understood easily like natural ones by human. Even for experienced screeners, it's very difficult to judge cargo category and contraband. In this paper, cargo X-Ray image is described by joint shape and texture feature, which could reflect both cargo stacking mode and interior details. Classification performance is compared with the benchmark method by top hit 1, 3, 5 ratio, and it's demonstrated that good performance is achieved here. In addition, we also discuss X-Ray image property and explore some reasons why cargo classification under X-Ray is very difficult.**

Keywords-cargo X-Ray image classification; edge based BOW; joint shape and texture feature

I. INTRODUCTION

Severe security and smuggling problems put great pressure on custom. Security & Inspection X-Ray Systems (Figure 2) is often used to accomplish some security missions by inspecting import-export cargo. Some computer-aided analysis technologies, such as image denoising, contrast enhancing, material recognition, and background peeling off, have been used to help screeners improve work efficiency. However, it requests that screeners must keep high concentration on the screen. Any little distraction would leave hidden dangers.

Nowadays, the computer vision technologies like object detection and image labeling have become hot spots. Someone would argue that why won't we apply these technology in the X-Ray image directly. Unfortunately, it's not easy like it sounds. Perspective results in the loss of exterior object structure and appearance information, which makes X-Ray image critically deviated from the natural one and couldn't be understood easily by human. In addition [1], due to (1) amounts of objects with different appearance and shape, (2) overlap between objects, (3) large variability of image depending on viewpoints, there is a long way to go in the X-Ray image field.

D.Mery et al. [1, 2, 3, 4] has made some explorations, which mainly focuses on threatening items, such as gun and knife in baggage screening. T Franzel et al. [5] use X-Ray multi-views information to detect objects. Muhammet Bastan et al. [6] brings visual words to describe baggage X-Ray image. These works give us some motivations to go ahead along this way.

In this paper, we classify cargo X-Ray images into 22 categories (Figure 1), according to HSCODE established by WCO (World Customs Organization) [7]. We design a special feature descriptor corresponding to cargo X-Ray image, which could desbribe both cargo stacking mode and interior details, and give good performance in the experiment.

In the rest of this paper, we give some related works in the image classification field in section 2, analyze X-Ray image property in section 3 and design the joint shape and texture feature in section 4. In addition, we provide a feature profile visualization method by two-layers SOM (Self Organization Map) [8, 9]. In section 5, some classification comparative tests are provided and good performance is achieved here. In section 6, we discuss the remaining challenges and future work to address them.

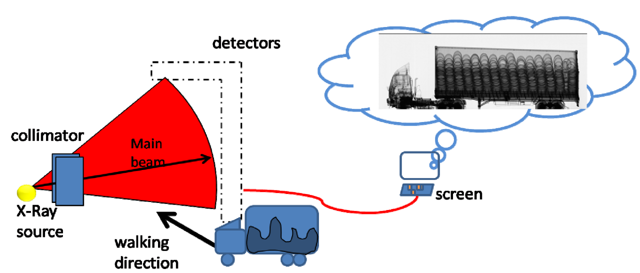


Figure2: Security & Inspection X-Ray Systems Diagram

II. Related Works

In recent 10 years, researchers have made much contributions to both representing image content, such as gray histogram and texture, and designing learning model, such as SVM (Support Vector Machine), LDA (Latent Dirichlet Allocation) and CRF (Conditional Random Field).

It's no doubt that image representation always plays a key role in the matter of improving classification performance. Image features usually fall into two modes - low-level and high-level description. Low-level description is one that describes individual components and provides detail rather than overview, such as color. Due to its poor discrimination power, the corresponding learning model with feature selection or voting strategy has to be provided, such as boosting [10] and random forest [11]. High-level description is one that is more abstracted, describes overall goals and systemic features, and is typically more concerned with the system as a whole, such as HOG, texton [12], boundary fragment [13].

BOW (Bag-of-Words) [14] representation has received wide application with its good performance in object categorization. The regular BOW utilizes k-means to create visual codebooks, which possesses proper inter-class discrimination and intra-class similarity. In addition, variants of BOW also have been proposed, such as semantic texton forest (Jamie Shotton [15]), and randomized clustering forest (Frank Mocosmann [16]) , which adopt a different coding mode. A method called shape based BOW [17] could describe human posture, which is most related to our implementation.

CRF model in object classification field proposed by Nils Plath[18], Justin Domke[19] brings semantic property easily, which emphasizes that neighbor pixel or superpixel should hold the same label. LDA [20, 21] was originally applied in the text analysis field. Recently, it's also brought into image classification by Liangliang Cao [22], which constructs a generative model by combing both spatial and content information. The hypothesis behind this probability topic model is that each data point can exhibit multiple components or 'topics'. To some extent, this could be considered as probability generalization of BOW.

To our best knowledge, the latest description about the application of computer vision technology in X-Ray field is D.Mery et al.[1, 2]. He gives the review in this field and threatening items detection algorithm by combing key points under different multi-views.

In this paper, our major contributions are: (1) by HSCODE established by WCO (World Customs Organization) [7], classifying cargo X-Ray image into 22 categories, (2) proposing joint shape and texture feature capturing both cargo stacking mode and interior details and (3) proposing feature profile visualization method.

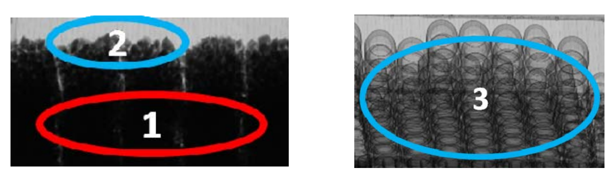
III. Cargo X-Ray Image Property

X-Ray beam is subjects to the exponential attenuation law:

$$I_1 = I_0 \exp(-\rho_m d_m)$$

, where I_0 is the original intensity of the beam; I_1 is the intensity of the beam at mass distance d_m into the substance; ρ_m is the mass attenuation coefficient. ρ_m is strongly correlated with substance material property and X-Ray energy.

Considering the limited dosage, X-Ray couldn't penerate some cargo regions with large mass-thickness. These image regions usually accompany low gray and lie at the bottom of container (region 1 in figure 3), which hardly provide some valid information. Under the sameX-Ray energy, gray and contrast imply cargo material property.



(a) coke (b) tire

region 1: high mass-thickness region
region 2: transition region
region 3: low mass-thickness region

Figure 3: cargo X-Ray image

The transition region (region 2 in figure 3) between cargos reflects the stacking mode. The low mass-thickness region (region 3 in figure 3) of cargo reveals much interior details. Cargo stacking mode, interior structure information and material property are three key ingredients to reveal cargo category.

Perspective, overlap, and viewpoint dependence are three main factors to disturb human's understanding for X-Ray image.

Perspective leads to the loss of object exterial structure and apperance details. Multual overlap between objects under X-Ray critically increase image complexities. In figure 4, two X-Ray images are plastic chair and car parts. It's not easy to map them to some our familiar objects.

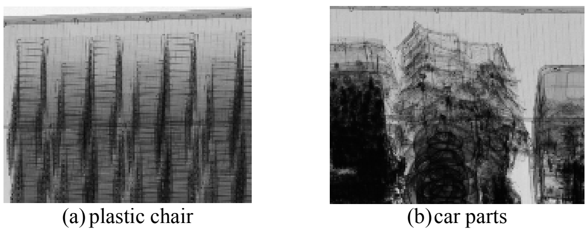


Figure 4: loss of exterial strucuture and apperance information

In addition, if X-Ray penetrates the cargo in container under different viewpoints, the object information reflected in X-Ray image would be different. Cargo (solder material) X-Ray images in figure 5 are collected under two different viewpoint 0° and 10° respectively. As far as we know, there is no 'viewpoint invariant' feature for X-Ray image.

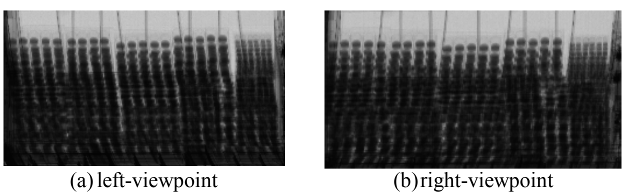


Figure 5: solder material X-Ray image at different viewpoint

Though facing so many challenges, there are certain traits we could take advantage of. Similar background, clear stacking pattern and simple primitives of cargo X-Ray image also bring classification task a new hope. How to describe cargo stacking pattern and image primitives effectively is the key.

IV. JOINT SHAPE AND TEXTURE FEATURE

Dense sampling BOW (Bag-of-Words) is the benchmark of extracting region feature. With different feature description on local pixels, it could describe statistical property in image region from different perspectives. Considering the contribution of every sampling position in image region equally, dense sampling BOW has weak capability to describe object shape in the region.

For cargo X-Ray images, object shape in region represents the cargo stacking mode and region texture reflects cargo interior details. By experience, both these two aspects reveal cargo category characteristics in X-Ray image.

In this section, edge based sampling BOW is proposed, which could describe object shape effectively when revealing region statistical information. On every sampling position, Leung-Malik (LM) Filter Bank [23] (Figure 6) with size 10 * 10 is used. It consists of 48 filters - 2 Gaussian derivative filters at 6 orientations and 3 scales, 8 Laplacian of Gaussian filters and 4 Gaussian filters, which describes direction and gray informaiton under multi-scales in the local region around sampling position.

Therefore, for X-Ray image, edge based sampling BOW with LM would reveal cargo stacking mode, interior structure informaiton and material property in region. That is why it is called "joint shape and texture feature".

All features are extracted in superpixel region. Superpixel are generated by SRM (Statistical Region Merging) [24] with complexity 0.2, which would gaurantee that every superpixel covers most parts of cargo and transimation regions. By experience, the superpixel with mean gray larger than 110 is invalid (non cargo region).

In order to compare the feature proposed here and the benchmark, we use the same group of visual words. Visual words are generated by dense sampling with step size 5 pixels and LM texture descriptor with size 10 * 10, which constructs a complete texture dictionary with 50 elements (Figure 7).

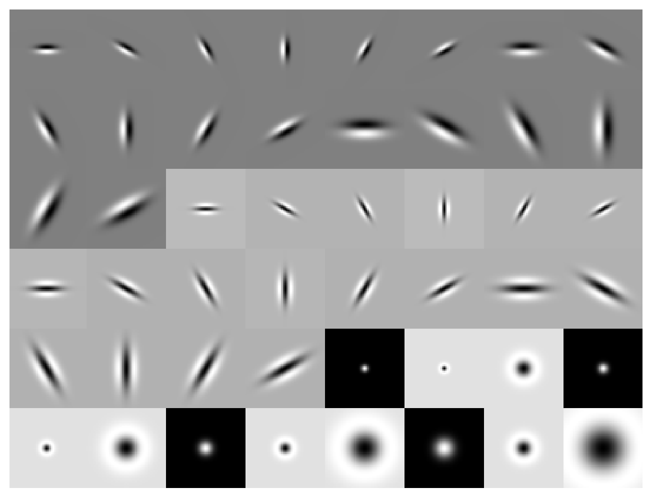


Figure 6: Leung-Malik (LM) Filter Bank

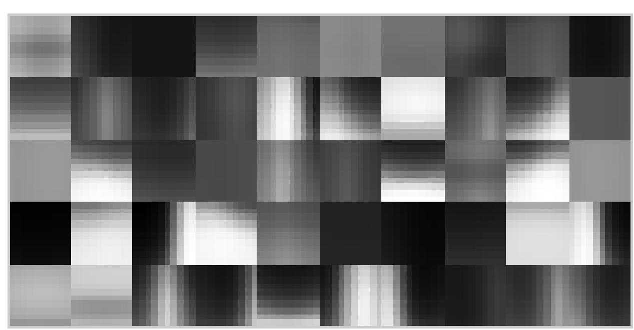
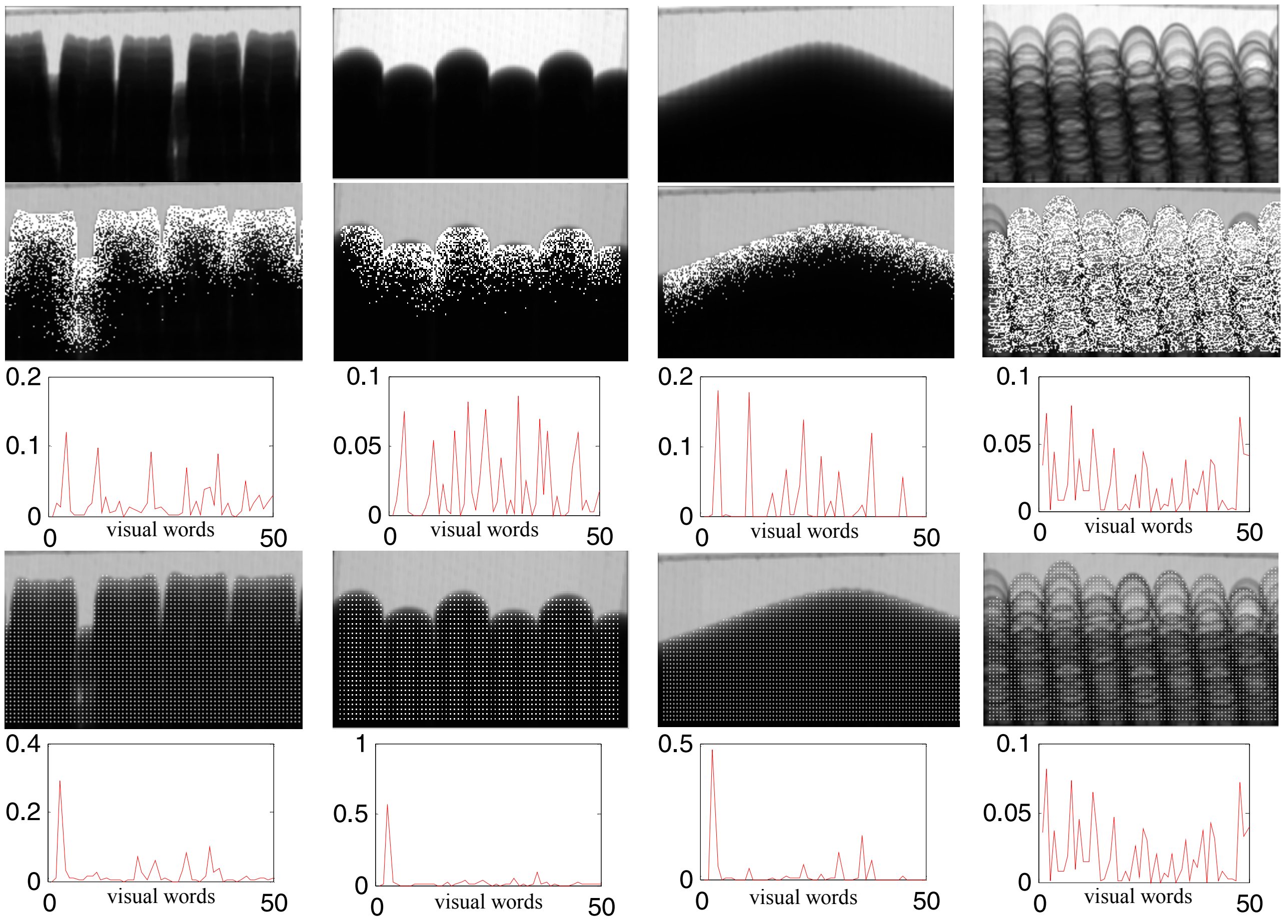


Figure 7: 50 visual words, which constructed by dense sampling with step size 5 pixels and LM texture descriptor with size 10 * 10. Visual words index is arranged in order from left to right and from top to bottom.



(a) activated carbon (b) paper (c) brewers grains (d) tire

Figure 8: superpixel representation. First row: original X-Ray image. Second row: white points represent edge based sampling positions. Third row: superpixel representation by edge based sampling. Fourth row: white points represent dense sampling positions. Fifth row: superpixel representation by dense sampling

In the rest of this scetion, firstly we describe the benchmark briefly in 4.1, then we give some details about edge based sampling BOW with LM in 4.2, finally we use 2- layers SOM (Self Organize Map) to visualize the joint shape and texture feature proposed here.

A. Benchmark

Dense BOW considers equally the contribution of every sampling position to encode superpixel content. It puts much attention on texture information inside superpixel. In every valid superpixel, we sample every 5 pixels (Figure 8 row 4) and extract the corresponding LM feature (LM size is 10 * 10). By the Nearest Neighbor, finding the corresponding one in the visual words set for every sampling feature. Finally, coding sueprpixel by visual words statistical histogram.

In figure 8 fifth row, we display superpixel representation result. For (a) activated carbon, (b) paper and (c) brewers grains, the peak of visual word distributions mainly locates in the visual word 3 (Figure 7), which illustrates that plain texture pattern takes up the main part and other texture patterns have weaker significance. The main reason is that most sampling points fall into the plain gray region, which gives poor discriminality. For (d) tire, interior structure details are shown clearly in X-Ray image, almost all sampling positions provide valuable texture information.

B. Edge Based Sampling BOW

For nautural images, edges usually represent shape characteristics of object. For X-Ray images, edges mainly locate in the transition region among cargos (Figure 3 region 2) and low mass-thickness region (Figure region 3), which reflect cargo stacking mode and interior structure information.

Unlike dense sampling, edge based sampling BOW with LM focuses on texture information around edges in the form of probability, which describes both shape characteristics and texture patterns effectively. The implementation details are shown in Algorithm 1.

The step 4 of Algorithm 1 is the key one. The way of Gaussian sampling around edges emphasizes edge shape characteristics and leaves some chances for those positions away from edges to encode sueprpixel content (Figure 8 row 2). As the variance of Gaussian distribution becomes larger, further texture information would be considered. The maximum sampling number controls statistical precision.

In figure 8 row 3, some classical cargo features are shown. For (c) brewers grains, sampling position mainly distributes at the upper edge of grain heap. The direction distribution is more simple than (a) activated carbon and (b) paper. For (a) activated carbon, the peak of visual word distribution maily locates at visual word 4,11,23,31,38 (Figure 7), these visual words present horizontal texture. Besides those possessed by (a) activated carbon, peaks of (b) paper lie in visual words 16,19,27. These three additional visual words present semi-circular texture. For (d) tire, due to complex direction distribution, its visual words distribution is critically different from the other three.

Algorithm 1 Edge Based Sampling BOW

Input: image, max sammpling number, variance
Output: superpixel feature description
1: splitting image to superpixels
2: getting rid of invalid superpixel
3: extracting edges by Canny
4: **for** point in edges of every superpixel
    Gaussian random with mean zero and variance
    while(count < max sampling number)
        point = point + gaussian random
        extracting feature at point
        finding the corresponding one in visual words set by Nearest Neighbor
        count = count + 1
    end while
  end for
5: **output** visual words distribution of every superpixel

C. *Feature Profile*

Is the calssification of cargos possible in 'single-energy single-view' X-Ray image? We argue that the answer is 'yes'. To illustrate the feasibility of this new application, we employ 2-layers SOM [19] to visualize the feature profile from some clues.

In figure 9, we describe 2-layers SOM structure. Distance metric is Euclidean distance. In the first-layer, discarding feature label, SOM is linked to feature data directly. Depending on data characteristics, data is mapped into some node regions (see A in figure 9). The lower the overlap ratio of node regions mapped by feature data with different class label, the stronger the feature descriptive power.

In order to give intuative feeling about overlap ratio, the second-layer SOM is built. The implementation details are shown in Algorithm 2.

Algorithm 2 Building The Second-Layer SOM

Input: The first-layer SOM model, feature and class label data
Output: feature profile
1: **for** feature in class label
    find the winning node and assign saliency score with exp(-distance) by the first-layer SOM mode
    the winning node: hue = class hue
    the winning node: saturation += saliency score
    the winning node: lightness = max(lightness, saliency score)
  end for
2: **for** node in the first-level SOM
    normalize saturation of node
    normalize lightness of node
  end for
3: transform hue, saturation and lightness to RGB space
4: using RGB of every node as feature data to train the second-layer SOM
5: **output** feature profile by the second-layer SOM

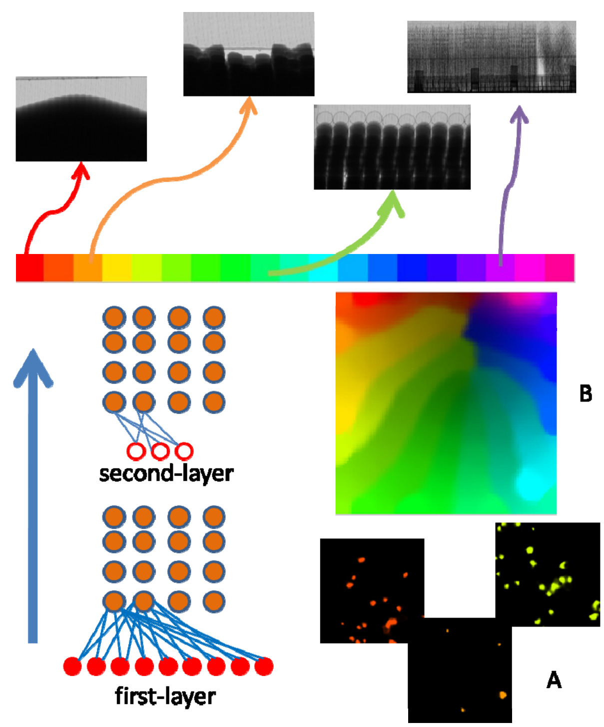


Figure 9: 2-layers SOM feature profile. The first-layer SOM is linked to feature data directly. By unsupervised learning, Data is maped into different nodes regions (see A). After filling different hue for node regions mapped by feature data with different class label, the second-layer SOM is built on RGB data. Node regions representing the same class are mapped into one single region (see B). The higher the saturation, the stronger the feature discriminality.

Figure 9 B is the joint shape and texture feature profile generated by 2-layers SOM model. The higher the saturation, the stronger the feature discriminality. We can see the boundary between feature data with different class label is quite clear. This result may imply that classification is welcome.

V. EXPERIMENTS

Data set: we collected 818 cargo X-Ray images, which cover most import and export cargo categories.

Task: according to sections of HSCODE established by WCO (World Customs Organization) [7], we try to classify cargo into 22 categories, which are 1) live animals & animal products, 2) vegetable products, 3) animal or vegetable fats and oils and their cleavage products, 4) prepared foodstuffs & beverages & spirits & vinegar & tobacco & manufactured tobacco substitutes, 5) mineral products, 6) products of chemical of allied industries, soap & organic surface-active agents & washing preparations & lubricating preparations, 7) plastics and articles thereof & rubber and articles thereof, 8) raw hides and skins & leather & fur skins and articles thereof, 9) wood and articles of wood & cork and articles of cork & manufactures of straw, 10) pulp of wood or other fibrous cellulosic material, 11) textiles and textile articles, 12) footwear & head gear & umbrellas walking-sticks & seat-sticks, 13) articles of stone & plaster & cement & asbestos & mica or similar material, 14) natural of cultured pearls, 15) base metals and articles of base metal, 16) machinery and mechanical appliances & electrical equipment & parts thereof, 17) vehicles & aircraft & vessels and associated transport equipment, 18) optical photographic & cinematic graphic, 19) arms and ammunition & parts and accessories, 20) miscellaneous manufactured articles, 21) works of art, and 22) special trading products, respectively.

Due to specificity of some cargos, such as works of art and special trading products, the data set doesn't include these categories. In experiment, we only consider 19 cargo categories. We use 661 of 818 images as training set, and the rest images as test set. All experiments are implemented by SVM classifier by 10-folder cross-validation.

In the rest of this section, we describe: (1) how to preprocess X-Ray image, (2) extract container region, (3) get rid of non-classical cargo samples, (4) encode superpixel and normalization, and (5) give some contrast experiment classification results.

A. Image Preprocess

The physical size of X-Ray image is usually larger than 1000 * 2000. It's not fitted to process further. We resize X-Ray image to 0.15 of physical size. In order to get rid of noise influence, smoothing image by Gaussian kernel with size 3 * 3. After preprocessing, cargo X-Ray image size is about 150 * 300.

B. Extracting Container Region

In general, container only takes up a small region in image. In order to extract the feature of cargo region in container and prepare training samples for SVM classifier, we have to extract container region. Some details are shown in Algorithm 3.

Algorithm 3 Extracting Container Region

Input: X-Ray image
Output: container region image
1: binarizing X-Ray image by OSTU [25]
2: projecting binarization image to x-axis
3: finding left and right bounds of container from projecting histogram(see figure 10 bottom)
4: extracting local binarization image by left and right bounds
5: projecting local binaryization image to y-axis.
6: finding up bound(see figure 10 right)
7: set the bottom bound by experience.
8: **output** container region image by left, right, up and bottom bounds.

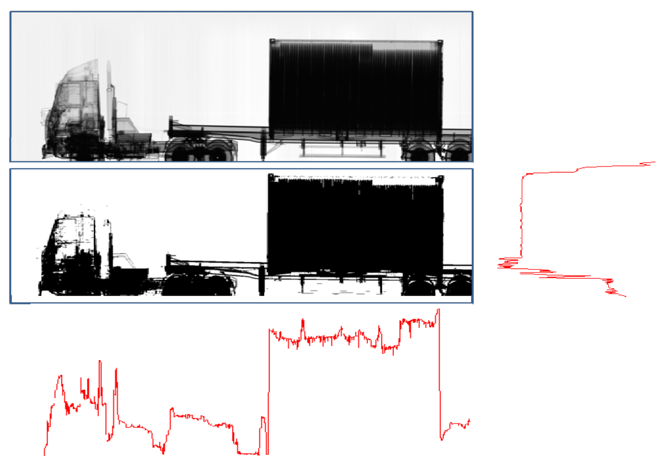
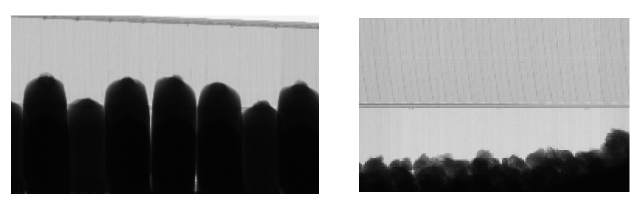


Figure 10: extracting container region from X-Ray image

C. Getting Rid of Non-Classical Cargo Samples



classical cargo sample　　non-classical cargo sample

Figure 11: classical and non-classical cargo sample

We believe cargo quantity in container should be larger than half of container space. Small amounts of cargo couldn't reflect the regular stacking mode, which is just the key part of describing cargo information in X-Ray image. By experiments, we find that after getting rid of those non-classical training samples, the classification performance could improve 0.02 averagely. Classical and non-classical cargo instances are shown in figure 11.

D. Superpixel Encoding and Normalization

We use SRM to build superpixel image. Under complexity 0.2, SRM could find proper superpixel. In order to compare the classification performance between joint shape and texture feature proposed in section 4 and dense sampling BOW, we encode every superpixel by two different ways.

For edge based sampling mode, LM texture descriptor size is 10 * 10 and Gaussian distribution using around edges is mean zero and variance 10. For dense sampling BOW, LM filter size is 10 * 10 and sampling step is 5. Normalize superpixel feature by Whitening transform [26].

E. Classification Results

All experiments are complemented by SVM classifier from libsvm [27].

In table 1, we give top hit 1, 3, 5 ratio for 157 test samples and the corresponding average ratio by encoding superpixel with joint shape and texture feature.

TABLE I. CLASSFICATION PERFORMANCE

h\t	1	3	5
01	0.33(1 of 3)	0.67(2 of 3)	0.67(2of 3)
02	0.17(1 of 6)	0.33(2 of 6)	0.83(5 of 6)
03	0.0(0 of 1)	1.0(1 of 1)	1.0(1 of 1)
04	0.72(13 of 18)	0.94(17 of 18)	1.0(18 of 18)
05	0.14(1 of 7)	0.43(3 of 7)	0.86(6 of 7)
06	0.39(12 of 31)	0.87(27 of 31)	0.94(29 of 31)
07	0.33(5 of 15)	0.8(12 of 15)	1.0(15 of 15)
08	0.8(4 of 5)	0.8(4 of 5)	1.0(5 of 5)
09	0.55(5 of 9)	0.78(7 of 9)	0.89(8 of 9)
10	0.58(7 of 12)	0.67(8 of 12)	0.83(10 of 12)
11	0.67(8 of 12)	0.83(10 of 12)	0.83(10 of 12)
12	1.0(3 of 3)	1(3 of 3)	1.0(3 of 3)
13	0.33(2 of 6)	0.67(4 of 6)	0.67(4 of 6)
15	0.92(12 of 13)	1.0(13 of 13)	1.0(13 of 13)
16	0.6(3 of 5)	0.8(4 of 5)	0.8(4 of 5)
17	0(0 of 3)	0.33(1 of 3)	0.67(2 of 3)
18	0.5(1 of 2)	1.0(2 of2)	1.0(2 of 2)
20	0.5(2 of 4)	0.75(3 of 4)	0.75(3 of 4)
21	0(0 of 2)	0(0 of 2)	0(0 of 2)
ave	0.51	0.78	0.89

top hit 1,3,5 ratio by encoding superpixel with joint shape and texture feature. HSCODE of every class is listed at the first column.

In table 2, we give top hit 1, 3, 5 ratio for 157 test samples and the corresponding average ratio by encoding superpixel with dense sampling BOW.

Compared with dense sampling BOW, classification performance brought by joint shape and texture feature is improved 4%, 11%, 9% at top hit 1, 3, 5 ratio respectively.

However, even for joint shape and texture feature, top hit 1 ratio just achieves 51%. Large intra-class discrimination and inter-class similarity is the key factor of poor classification performance (Figure 1). Although sea food and milk powder belong to HSCODE 01, the stacking mode and interior structure information is completely different. The same problem also occurs in other cargo categories, such as HSCODE 02, 10 and 15. For tomato sauce (HSCODE 04) and lube oil (HSCODE 05), the similarity of image information could almost deceive human eyes. Potato (HSCODE 02), rubber (HSCODE 07) and cotton (HSCODE 11) also face this problem. This is why cargo X-Ray image classification is such difficult.

TABLE II. CLASSFICATION PERFORMANCE

h\t	1	3	5
01	0(0 of 3)	0(0 of 3)	0.67(2 of 3)
02	0.33(2 of 6)	0.33(2 of 6)	0.67(4 of 6)
03	1.0(1 of 1)	1.0(1 of 1)	1.0(1 of 1)
04	0.39(7 of 18)	0.72(13 of 18)	0.89(16 of 18)
05	0.28(2 of 7)	0.57(4 of 7)	0.85(6 of 7)
06	0.45(14 of 31)	0.65(20 of 31)	0.84(26 of 31)
07	0.27(4 of 15)	0.67(10 of 15)	0.87(13 of 15)
08	0.8(4 of 5)	0.8(4 of 5)	0.8(4 of 5)
09	0.67(6 of 9)	0.78(7 of 9)	0.78(7 of 9)
10	0.58(7 of 12)	0.58(7 of 12)	0.75(9 of 12)
11	0.58(7 of 12)	0.67(8 of 12)	0.75(9 of 12)
12	0.67(2 of 3)	0.67(2 of 3)	0(3 of 3)
13	0.67(4 of 6)	0.67(4 of 6)	0.67(4 of 6)
15	0.69(9 of 13)	0.84(11 of 13)	0.84(11 of 13)
16	0.2(1 of 5)	0.6(3 of 5)	0.8(4 of 5)
17	0(0 of 3)	0.67(2 of 3)	0(2 of 3)
18	0.5(1 of 2)	0.5(1 of 2)	0.5(1 of 2)
20	0.75(3 of 4)	0.75(3 of 4)	1.0(4 of 4)
21	0(0 of 2)	0(0 of 2)	0(0 of 2)
ave	0.47	0.67	0.8

top hit 1,3,5 ratio by encoding superpixel with dense sampling BOW. HSCODE of every class is listed at the first column.

In table 3, we give average top hit 1, 3, 5 ratio of different sampling method with SIFT [28] feature. Although SIFT feature couldn't describe region gray information, with the help of edge based sampling, it also gives suboptimum performance. Therefore, besides gray and contrast information in X-Ray image, cargo stacking mode is the second key factor to improve the classification performance of cargo X-Ray image.

TABLE III. CLASSFICATION PERFORMANCE

	edge based sampling	dense sampling
top hit 1	0.49	0.46
top hit 3	0.75	0.7
top hit 5	0.86	0.8

classification performance of edge based sampling BOW with SIFT feature and dense sampling BOW with SIFT

VI. CONCLUSION

In this paper, we propose joint shape and texture feature corresponding to the classification task of cargo X-Ray image. Depending on sampling around edges, the feature captures both cargo stacking mode and texture information in the superpixel. In our experiments, regardless of top 1, 3, 5 ratio, SVM classifier using joint shape and texture feature give better classification performance.

However, top hit 1 ratio 51% still couldn't satisfy the actual need. The high similarity between images in different cargo categories and the great variance of images in the same cargo category are the main reason.

Introduction of structure information may improve the existing feature performance. We believe that there exists a joint probability distribution for co-occurrence of different visual words, extracted from X-Ray images of the same cargo category. How to describe this distribution is the key of bringing structure information.

Splitting the high-dimensional feature space maybe a feasible solution. The classifier model is applied in every feature subspace.

Combing other information, such as multi-views and dual energy, is also an attractive selection. Multi-views provide the image content information under different viewpoints, and dual energy reveals material atomic number.

All of these three aspects are our next step to go.

ACKNOWLEDGMENT

This research was supported by Nuctech.

REFERENCES

[1] D.Mery. X-ray Testing by Computer Vision. In Proceedings of the 9th IEEE CVPR workshop on Perception Beyond the Visible Spectrum, 2013

[2] D.Mery. Automated X-ray Object Recognition Using an Efficient Search Algorithm in Multiple Views. In Proceedings of the 9th IEEE CVPR workshop on Perception Beyond the Visible Spectrum, 2013

[3] D.Merry, Riffo V, Zuccar I, Pieringer C. Automated X-Ray Object Recognition Using an Effieicent Search Algorithm in Multiple Views. Computer Vision and Pattern Recognization Workshops, 2013

[4] D.Mery. X-Ray Testing: the State of the Art.

[5] T Franzel, USchmidt, S Roth. Object Detection in Multi-View X-Ray Images. Springer 2012.

[6] Muhammet Bastan, Mohammad Reza Yousefi, Thomas M.Breuel. Visual Words on Baggage X-Ray Images. Computer analysis of images and patterns, 2011

[7] http://www.hscode.org/

[8] T. Kohonen, The Self-Organizing Map, Proceedings of the IEEE. 1990, pp. 1464-1480

[9] http://www.generation5.org/content/2004/aiSomPic.asp

[10] Redpath.D.B, Lebart.K. Boosting Feature Selection. Pattern Recognition and Data Mining. 2005, pp. 305-314

[11] Breiman L. Random forests. Machine Learning. 2001, pp. 5-32

[12] Varma.M, Zisserman.A. A Statistical Approach to Texture Classification from Single Images. International Journal of Computer Vision.2005, pp. 61-81

[13] Opelt.A, Pinz.A, Zisserman.A. A Boundary Fragment Model for Object Detection. Computer Vision–ECCV 2006. Springer Berlin Heidelberg, pp. 575-588

[14] G. Csurka, C. Bray, C. Dance, and L. Fan. Visual Categorization with Bags of Keypoints. Workshop on Statistical Learning in Computer Vision, ECCV, 2004, pp.1–22

[15] Shotton, J.Johnson, M.Cipolla. Semantic Texton Forests for Image Categorization and Segmentation. In Computer vision and pattern recognition, 2008, pp.1-8

[16] Moosmann, F., Triggs, B., & Jurie, F. Fast Discriminative Visual Codebooks Using Randomized Clustering Forests. In NIPS

[17] Niebles, J. C., Fei-Fei, L. A Hierarchical Model of Shape and Appearance for Human Action Classification. In Computer Vision and Pattern Recognition, 2007, pp.1-8

[18] Plath, N., Toussaint, M.Nakajima. Multi-Class Image Segmentation Using Conditional Random Fields and Global Classification. InProceedings of the 26th Annual International Conference on Machine Learning, pp. 817-824

[19] Domke J. Beating the Likelihood: Marginalization-Based Parameter Learning in Graphical Models, 2012.

[20] D.M.Blei, A.Ng, and M. I. Jordan. Latent Dirichlet Allocation. JMLR, 2003, pp.993-1002

[21] D. M. Blei and J. D. McAuliffe. Supervised Topic Models. In NIPS, 2007

[22] Cao L, Fei-Fei L. Spatially Coherent Latent Topic Model for Concurrent Segmentation and Classification of Objects and Scenes. Computer Vision, 2007, pp.1-8

[23] T. Leung and J. Malik. Representing and Recognizing the Visual Appearance of Materials Using Three-Dimensional Textons. International Journal of Computer Vision, 2001, pp.29-44

[24] Nock R, Nielsen F. Statistical Region Merging. Pattern Analysis and Machine Intelligence, IEEE Transactions on 2004, pp.1452-1458

[25] http://en.wikipedia.org/wiki/Otsu's_method

[26] http://en.wikipedia.org/wiki/Whitening_transformation

[27] http://www.csie.ntu.edu.tw/~cjlin/libsvm/

[28] Lowe D G. Distinctive Image Features from Scale-Invariant Keypoints. International journal of computer vision, 2004, pp.91-110