Journal of Zhejiang University SCIENCE A
ISSN 1009-3095 (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Mean shift texture surface detection based on WT and COM feature image selection*

HAN Yan-fang†, SHI Peng-fei

(*Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, Shanghai 200030, China*)

†E-mail: hyf@sjtu.edu.cn

Received Aug. 5, 2005; revision accepted Dec. 21, 2005

Abstract: Mean shift is a widely used clustering algorithm in image segmentation. However, the segmenting results are not so good as expected when dealing with the texture surface due to the influence of the textures. Therefore, an approach based on wavelet transform (WT), co-occurrence matrix (COM) and mean shift is proposed in this paper. First, WT and COM are employed to extract the optimal resolution approximation of the original image as feature image. Then, mean shift is successfully used to obtain better detection results. Finally, experiments are done to show this approach is effective.

Key words: Mean shift, Wavelet transform (WT), Co-occurrence matrix (COM), Texture defect detection
doi:10.1631/jzus.2006.A0969 **Document code:** A **CLC number:** TN911.73

INTRODUCTION

Visual inspection plays a vital role in assuring the quality of industrial products. Texture surface detection aims at detecting defects such as cracks, stains, broken points, etc. on texture surfaces. Due to the repetitive changes of texture gray values and structures, traditional detection methods based on intensity or edge detection are invalid, which makes texture surface detection one of the most intriguing problems during the past decades.

Generally, most of previous approaches to texture surface detection are based on clustering techniques, with texture feature extraction and texture classification being two major objectives. The texture feature extraction techniques range from the mainly statistics-based (Goldman and Cohen, 2004), model-based (Stan *et al*., 2002), Fourier-based (Chan and Pang, 2000), Gabor filtering (Ahmadian and Mostafa, 2003) to the latest wavelet approaches (Arivazhagan and Ganesan, 2003). While in terms of texture classification techniques, the proximity based classifiers (Huang *et al*., 2003) (such as Bayes, Euclidean distance, Mahalanobis distance, *K*-Means) and learning based classifiers (Park *et al*., 2004) (such as genetic, neural network) are successful and widely used.

However, on the one hand, among the texture feature extraction techniques, the statistical and model-based methods are, mainly, based on spatial domain processing and the features are extracted only in one single scale. Besides, the computational efficiency is a specific problem requiring extensive research. In terms of Fourier transform, which deals with the image in frequency domain, it is successful for detecting global and macro defects, however, unsatisfactory for local and micro defect detection. Then, starting from the late 1980s, due to the theoretical impact of the works of Daubechies (1988), who has provided the discrete wavelet transform (DWT) and Mallat (1989), who has established connection between WT and multi-resolution theory, wavelet method, which is a successful multi-resolution analysis tool in frequency domain, has received considerable attention in image processing.

* Project (No. 035115039) supported by the Scientific Committee of Shanghai, China

While on the other hand, the texture clustering techniques, proximity based and learning based classifiers, have unavoidable drawbacks when dealing with the vast image data, e.g., the proximity-based methods tend to be computationally expensive and the definition of a meaningful stopping criterion for the fusion (or division) of the data is not straightforward. Often, the learning based classifiers need to be trained by the non-defect features, which is a troublesome procedure and usually time consuming, therefore, limits its real-time applications. In addition, it is sometimes necessary to decide the number of clusters using prior knowledge so that over-segmentation can be avoided, which often makes it neither robust nor efficient.

Assuming the feature space of an image undertaking a kind of probability density distribution, mean shift, which was proposed by Fukunaga and Hostetler (1975) and performed by iteratively finding the modes of the density distribution, is robust and does not require prior knowledge of the number of clusters. Hence it is widely used in image segmentation. (Georgescu *et al*., 2003; Cheng, 1995; Singh and Ahuja, 2002; Yang and Liu, 2001). However, in this paper, when detecting the texture surface, some textures are wrongly detected as stains. Therefore, we propose an approach of extracting the approximation of the original image as feature image based on WT and COM so as to reduce the influence of the high frequency textures. Then regarding the feature space as certain empirical probability density function, mean shift is applied in stain detection.

The rest of this paper is organized as follows: in Section 2, mean shift is briefly introduced. Then, our approach based on WT and COM is elaborated in Section 3. After that, in Section 4 experiments are done to detect the stains on texture images by mean shift and our approach respectively, and comparison is made. Finally, concluding remarks are given in Section 5.

MEAN SHIFT

Mean shift (Abrantes and Marques, 2004), which is a non-linear kernel method proposed for clustering analysis, is an iterative technique. It tries to obtain the modes of the probability density function of the feature space, using a nonparametric estimate of the density function. And the number of clusters is obtained automatically by finding the centers of the densest regions in the space (the modes).

Assume that each data point in the feature space $\{x_i\}\in\mathbb{R}^d$, i=1, …, n. Then, the multivariate kernel density estimate obtained with kernel function $K(x)$ and window radius h, computed in the point x is defined as:

$$\hat{f}(x)=\frac{1}{nh^d}\sum_{i=1}^{n}K\left(\frac{x-x_i}{h}\right), \tag{1}$$

where the kernel function $K(x)$ can be different types such as Gaussian kernel, unit kernel, Epanechikov kernel and so on. But the Epanechikov kernel is proved to be the optimum kernel yielding minimum mean integrated square error (MISE) as described in (Comaniciu and Meer, 1999).

$$K_{\mathrm{E}}(x)=\begin{cases}\frac{1}{2}c_d^{-1}(d+2)\left(1-\|x\|^2\right), & \text{if } \|x\|^2<1,\\ 0, & \text{otherwise,}\end{cases} \tag{2}$$

where c_d is the volume of the unit d-dimensional sphere.

As we know, the dense regions in the feature space often correspond with the local maxima of the density function. And to find the local maxima, gradient of the density function is a useful tool. Thus, the gradient of the kernel density estimate can be computed as follows:

$$\hat{\nabla}f(x)\equiv\nabla\hat{f}(x)=\frac{1}{nh^d}\sum_{i=1}^{n}\nabla K\left(\frac{x-x_i}{h}\right). \tag{3}$$

The sample mean at x is

$$m(x)=\sum_{i=1}^{n}\frac{K(x_i-x)x_i}{K(x_i-x)}. \tag{4}$$

And the difference $\boldsymbol{M}_h(x)=m(x)-x$ is called mean shift. It has been proved in (Comaniciu and Meer, 2002) that:

$$\boldsymbol{M}_h(x)=C\frac{\hat{\nabla}f(x)}{\hat{f}(x)}, \tag{5}$$

where, C is a positive constant and $\boldsymbol{M}_h(x)$ is the mean shift value at the point x.

The repeated movement of data points to the sample means is called the mean shift algorithm. In each iteration of the algorithm, $x \leftarrow m(x)$ is performed for all $\{x_i\} \in \mathbb{R}^d$ simultaneously until $m(x)=x$.

TEXTURE DEFECT DETECTION BASED ON OUR APPROACH

Applying mean shift directly on the original image, stains can be detected with smaller kernel window size, however, with some textures being mistaken as defects too. While with the increase of kernel window, the influence of textures can be reduced but some stains cannot be detected. Fortunately, it is well-known that textures are high frequency parts while the stained background are low frequency contents in an image. So, if we can select only the low frequency background for mean shift detection, this problem may be settled. Therefore, combined with COM, which is a useful statistical tool in texture analysis, WT, which is a successful multi-resolution analysis tool in frequency domain, is used to select one approximation image with certain resolution as the feature image for mean shift detection.

As elaborated in many literatures, WT is defined as the inner product of a signal (image) with a family of real orthogonal basis functions, $\psi_{a,b}(x)$ with the computation formula as follows:

$$W_{a,b} = < f, \psi_{a,b} > = \int_{-\infty}^{+\infty} f(t)\psi_{a,b}(x)\mathrm{d}t, \tag{6}$$

where a is the dilating factor and b is the shifting factor.

The results of WT are wavelet coefficients W, which are a function of scale and position. Multiplying each coefficient by the appropriately scaled and shifted wavelet yields the constituent wavelets of the original signal:

$$f(t) = \sum_{a,b} W_{a,b}\psi_{a,b}(x). \tag{7}$$

Multi-scale wavelet transform can be implemented as a pyramid or tree structure. The decomposition process can be iterated, with successive approximations being decomposed in turn, so that one signal is broken down into many lower resolution components. Here, our objective is to select the optimal resolution approximation, in which the textures are removed, as the feature image for post-detection.

As is well known, COM (Latif-Amet *et al.*, 2000), which is a square matrix whose elements correspond to the relative occurrence frequency of pairs of gray level values of pixels separated by a certain distance in a given direction, well represents the texture distribution. If an intensity image is entirely flat (i.e., no texture contained), the resulting COM will be completely diagonal. As the image texture increases (i.e., as the local pixel intensity variations increase), the off-diagonal values of the COM become larger. Therefore, based on this notion, COM can be used to analyze the texture distribution for approximation selecting.

Mathematically, two COM features such as energy and local homogeneity are used for feature image selecting. Assuming C is the COM of image I, whose size is $M \times N$ with the gray level G, then,

$$C(i,j) = \frac{Num(P(x,y)=i \cap P(x+\Delta m, y+\Delta n)=j)}{\sum\limits_{m,n=1}^{M,N} \text{pairs}}$$

$$i=1, 2, \ldots, G, \tag{8}$$

where $P(x, y)$ represents the pixel value at (x, y), $Num(\cdot)$ is the operator for computing total occurrence number and the denominator is all possible pairs in the direction of $(\Delta m, \Delta n)$ in original image. The two COM features such as Energy and Local homogeneity can be computed as follows:

$$\text{Energy} = \sum_{i,j=1}^{N} C^2(i,j), \tag{9}$$

$$\text{Local homogeneity} = \sum_{i,j=1}^{N} \frac{1}{1+(i-j)^2} C(i,j). \tag{10}$$

These are two commonly used features and are analyzed by their variance to select the feature image from the approximation sub-images at variant decomposition levels. Let VE represent the variance of Energy and VL for that of Local homogeneity, then, the selected level (SL) can be represented as follows:

$$SL=\{m | VE_m = \max(VE_i) \cap VL_m = \max(VL_i),\ i \in [1,N]\}. \tag{11}$$

The approximation sub-image at this selected level is taken as the feature image.

Fig.1 gives an example of how the feature image is selected and its comparison with the original image.

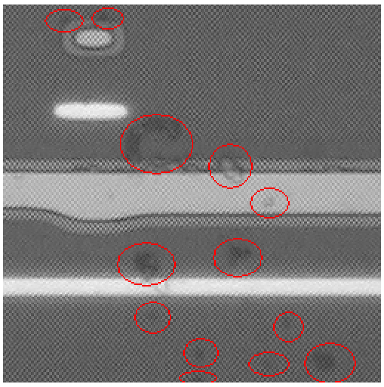
By WT and COM features analysis, the feature image is obtained. And the comparison of original and feature image is given in Fig.1. Fig.1a is the original texture image with Fig.1b as its COM. Fig.1c shows the variance of COM features at different decomposition levels. And Fig.1d is the feature image with its COM illustrated in Fig.1e.

Comparing the COMs of the original and feature image, it can be noted that the off-diagonal values of the original COM is large, which means the original image is with textures, while the COM of the feature image is completely diagonal, which means the feature image is flat.

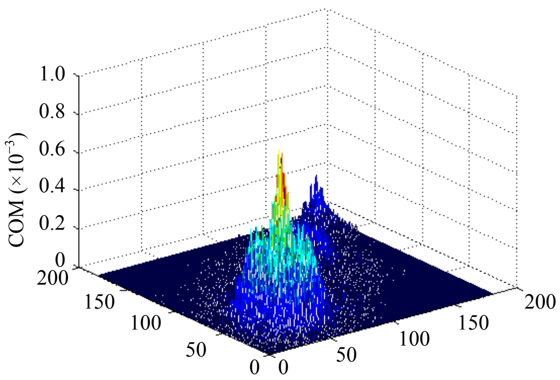
Based on the above analysis, a new approach is proposed based on WT, COM and mean shift for defect detection. Particularly, the main procedure is summarized as follows:

Step 1: Multi-scale WT of the original texture image;

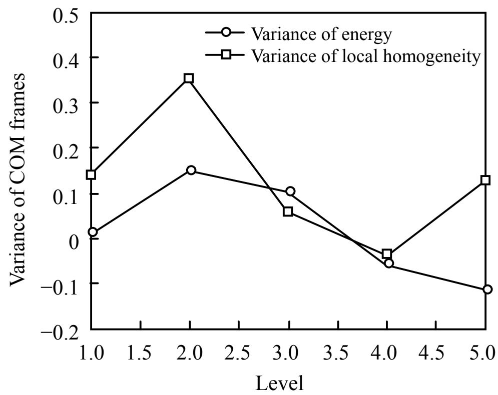
Step 2: Compute the COM and COM features of the decomposed sub-images following Eqs.(9) and (10);



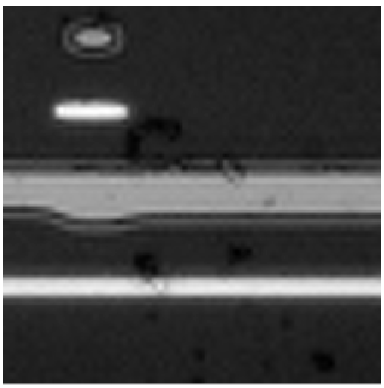
(a)



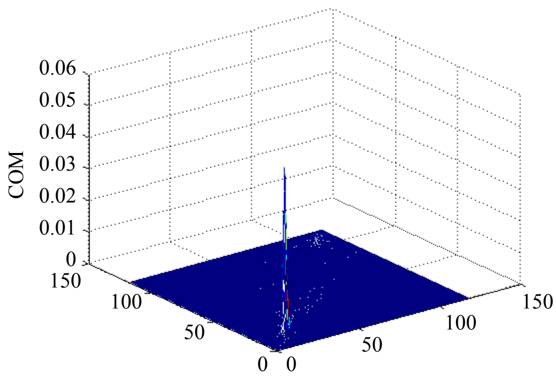
(b)



(c)



(d)



(e)

Fig.1 Comparison of original image and feature image. (a) Original texture image with stains marked in circles; (b) COM of original image; (c) COM feature variance of approximation sub-images at different levels; (d) Feature image at level 2; (e) COM of feature image

Step 3: Analyze the COM features of the approximation sub-images by their variance at different levels and select the optimal approximation as feature image;

Step 4: Compute the mean shift vector $\boldsymbol{M}_h$ on the feature image;

Step 5: In each iteration of the Mean Shift algorithm, $x \leftarrow \boldsymbol{M}_h(x)+x$ is performed for all $\{x_i\} \in \mathbb{R}^d$ simultaneously until $\boldsymbol{M}_h(x) \rightarrow 0$;

Step 6: Find the modes as cluster centers and merge or eliminate the smaller clusters.

It should be pointed out that there are many good wavelet bases proposed in previous references. However, in our experiments, we try db4 wavelet for WT and it is found that the detection results are good.

EXPERIMENTAL RESULTS

Based on the approach proposed above, experiments are done to detect the stains on texture surfaces. One original image is shown in Fig.1a, with the stains marked in circles. And the detecting results by mean shift and our approach are illustrated in Figs.2a and 2b respectively. From Fig.2a, it can be noted that some stain points on the texture image cannot be detected while some textures are wrongly detected as defects, which are marked with squares, while in Fig.2b, the stains are correctly detected and the detection is free from the influence of textures.

More experiments were done with the testing results given in Fig.3 to show the effects of our approach. In each image group of Fig.3, the original image (with the image size of 256×256), approximation image, mean shift detection results and results based on our approach are given in order. Where, the stains are marked with circles in the original image and the wrongly detected textures are marked in squares in the mean shift detection results.

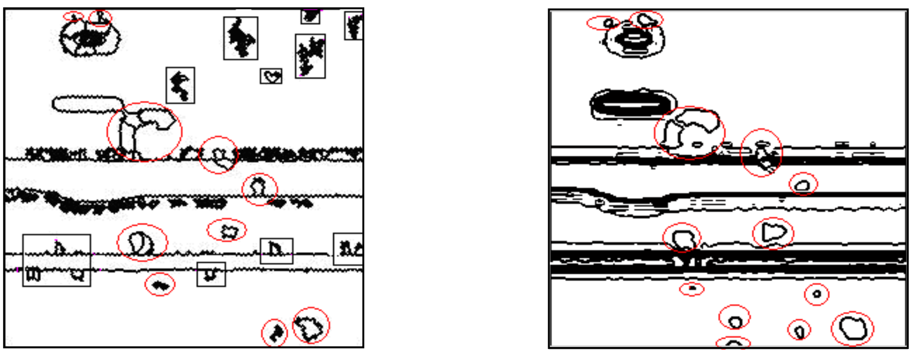
Table 1 gives the selected levels (*SL*) and corresponding COM feature variance in feature image obtaining. Where, the marked grids in the table represent the max COM feature variance.

All experiments were done under the environment of Pentium 2.4 G, Win XP, Matlab 6.5. From the above results, the error-detection rate and lost-detection rate for mean shift and our approach are computed and listed in Table 2.

From the above statistical numbers, the error-detection rate and lost-detection rate for mean shift is much larger than our approach, which proves that our approach is effective.

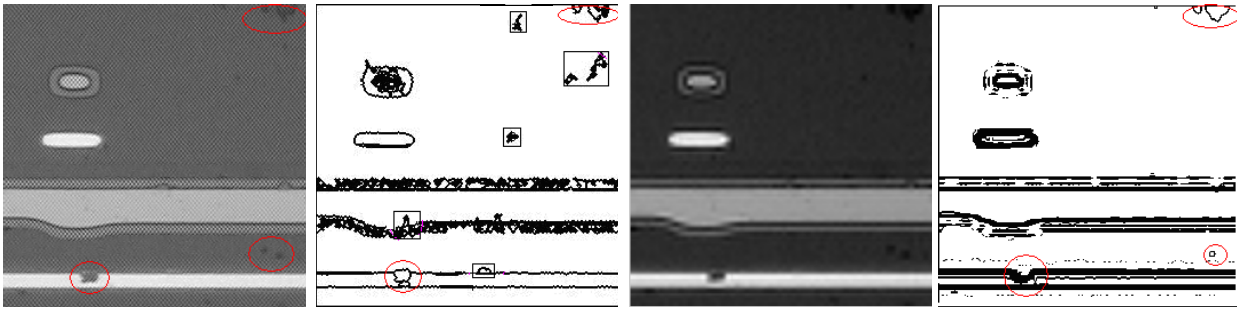
CONCLUSION

Although mean shift is an efficient iterative clustering approach in image segmentation, in the above experiments, its detection result is greatly influenced by the textures. Some stain points cannot be correctly detected and many textures are wrongly detected as defects. According to this problem, we start from the idea of extracting non-texture approximation of the original image with multi-level WT and COM for mean shift detection. Thus the textures are removed from the stained background. Experiments and comparison showed that the error detection rate and lost-detection rate for mean shift are much greater than our approach. Therefore, the performance and effect of our approach are proved.

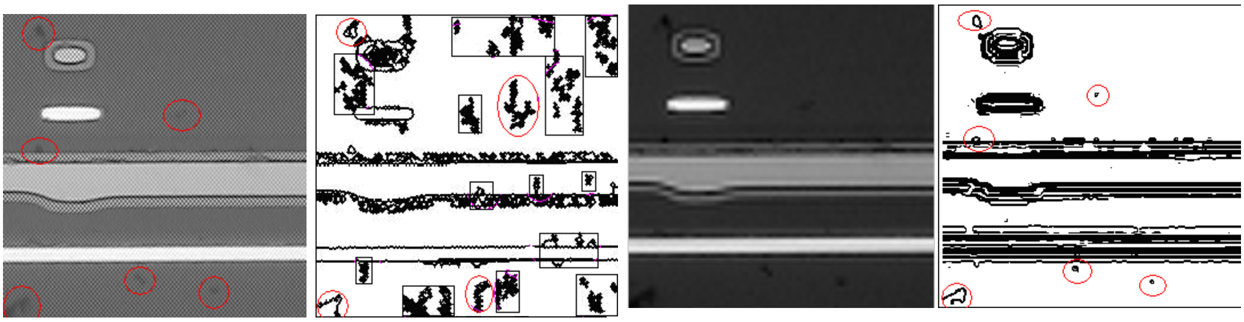


(a) (b)

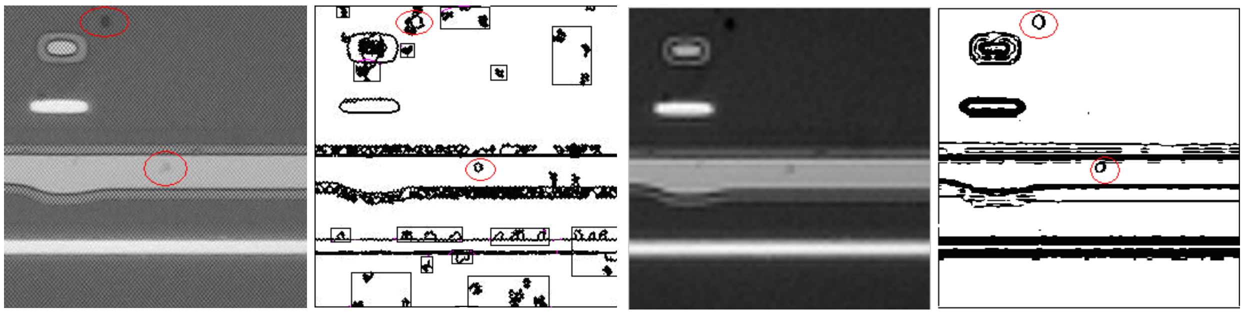
Fig.2 Comparison of detection results for Fig.1a. Stains are marked with circles and wrongly detected areas are marked with squares. (a) Detection result of mean shift; (b) Detection result of our approach



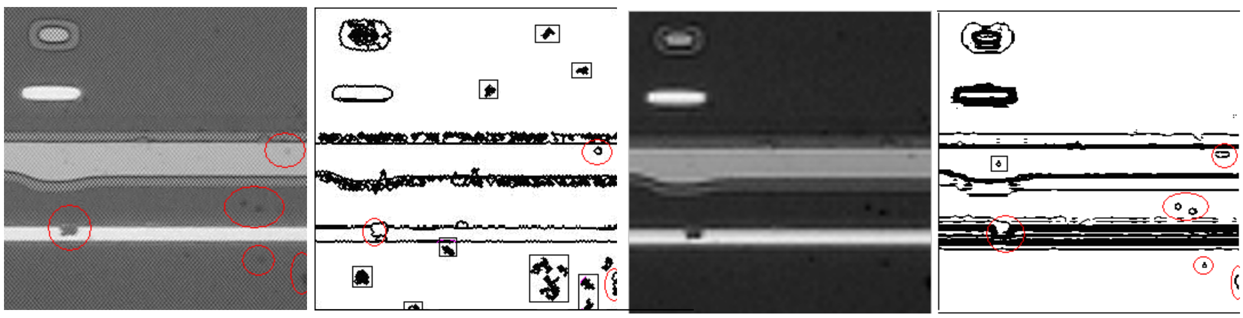
(a)



(b)



(c)



(d)

Fig.3 Comparison of detection results, from (a) to (d), images in the first column are original images, the second column mean shift detection results, the third approximation images and the last detection results based on our approach, with stains marked in circles and wrongly detected textures in squares

Table 1 Selected levels and corresponding COM feature variance in feature images

Original image	COM feature	Level 1	Level 2	Level 3	Level 4	Level 5	*SL*
Fig.1a	*VE*	0.0089074	0.14624	0.099638	−0.059688	−0.115220	2
	VL	0.1383500	0.35151	0.057630	−0.040992	0.126150	
Fig.3a	*VE*	0.0095838	0.18541	0.152130	−0.053312	−0.044617	2
	VL	0.1350800	0.37181	0.068173	−0.051990	0.015789	
Fig.3b	*VE*	0.0072970	0.24950	0.040034	−0.084794	−0.066177	2
	VL	0.1174700	0.39452	0.012750	−0.052299	0.023991	
Fig.3c	*VE*	0.0093941	0.18312	0.114470	−0.101600	−0.087365	2
	VL	0.1399100	0.36320	0.040728	−0.063235	0.002672	
Fig.3d	*VE*	0.0093319	0.16816	0.129320	−0.083590	−0.124380	2
	VL	0.1334700	0.36546	0.060543	−0.047540	0.107140	

Table 2 Error-detection rate and lost-detection rate of mean shift and our approach

Result images		Fig.2a	Fig.3a	Fig.3b	Fig.3c	Fig.3d	Total number	Detection rate (%)
Stain numbers on original image		13	3	6	2	5	29	–
Mean shift	Error numbers	10	5	13	14	8	50	172.41
	Lost numbers	3	1	2	0	2	8	27.59
Our approach	Error numbers	0	0	0	0	1	1	3.44
	Lost numbers	0	1	0	0	0	1	3.44

References

Abrantes, A.J., Marques, J.S., 2004. The Mean Shift Algorithm and the Unified Framework. Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04), **1**:244-247.

Ahmadian, A., Mostafa, A., 2003. An Efficient Texture Classification Algorithm Using GABOR Wavelet. Proceedings of the 25th Annual International Conference of the IEEE EMBS. Cancun, Mexico, (17-21):930-933.

Arivazhagan, S., Ganesan, L., 2003. Texture segmentation using wavelet transform. *Pattern Recognition Letters*, **24(**16**)**:3197-3203. [doi:10.1016/j.patrec.2003.08.005]

Chan, C.H., Pang, K.H., 2000. Fabric defect detection by Fourier analysis. *IEEE Trans. on Industry Applications*, **36**(5):1267-1276. [doi:10.1109/28.871274]

Cheng, Y.Z., 1995. Mean shift, mode seeking, and clustering. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **17**(8):790-799. [doi:10.1109/34.400568]

Comaniciu, D., Meer, P., 1999. Mean Shift Analysis and Applications. Proceedings of the 7th IEEE International Conference on Computer Vision, **2**:1197-1203.

Comaniciu, D., Meer, P., 2002. Mean shift: a robust approach toward feature space analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **24**(5):603-619. [doi:10.1109/34.1000236]

Daubechies, I., 1988. Orthonormal bases of compactly supported wavelets. *Commun. Pure Appl. Math*, **41**:909-996.

Fukunaga, K., Hostetler, L.D., 1975. The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Trans. Info. Theory*, **21**(1):32-40. [doi:10.1109/TIT.1975.1055330]

Georgescu, B., Shimshoni, I., Meer, P., 2003. Mean Shift Based Clustering in High Dimensions: A Texture Classification Example. Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV'03), **1**:456-463.

Goldman, A., Cohen, I., 2004. Anomaly detection based on an iterative local statistics approach. *Signal Processing*, **84**(7):1225-1229. [doi:10.1016/j.sigpro.2004.04.004]

Huang, Y., Chan, K.L., Zhang, Z.H., 2003. Texture classification by multi-model feature integration using Bayesian networks. *Pattern Recognition Letters*, **24**(1-3):393-401. [doi:10.1016/S0167-8655(02)00263-5]

Latif-Amet, A., Ertuzun, A., Ercil, A., 2000. An efficient method for texture defect detection: sub-band domain co-occurrence matrices. *Image Vision and Computing*, **18**(6-7): 543-553. [doi:10.1016/S0262-8856(99)00062-1]

Mallat, S.G., 1989. Theory for multi-resolution signal decomposition: the wavelet representation. *IEEE Trans. PAMI*, **11**:674-693.

Park, S.B., Lee, J.W., Kim, S.K., 2004. Content-based image classification using a neural network. *Pattern Recognition Letters*, **25**(3):287-300. [doi:10.1016/j.patrec.2003.10.015]

Singh, M.K., Ahuja, N., 2002. Mean-Shift Segmentation with Wavelet-Based Bandwidth Selection. Proceedings of the 6th IEEE Workshop on Applications of Computer Vision (WACV'02), (3-4):43-47.

Stan, S., Palubinskas, G., Datcu, M., 2002. Bayesian selection of the neighborhood order for Gauss-Markov texture models. *Pattern Recognition Letters*, **23**(10):1229-1238. [doi:10.1016/S0167-8655(02)00070-3]

Yang, X.Y., Liu, J., 2001. Unsupervised texture segmentation with one-step mean shift and boundary Markov random fields. *Pattern Recognition Letters*, **22**(10):1073-1081. [doi:10.1016/S0167-8655(01)00057-5]