International Journal of u- and e- Service, Science and Technology
Vol.9, No. 7 (2016), pp.91-100
http://dx.doi.org/10.14257/ijunesst.2016.9.7.09

Study of Copyright Protection for Merchandise Pictures in E-Commerce

Liyi Zhang[1] and Chang Liu[2*]

[1,2]School of information management, Wuhan University
[2]liuchang0310@163.com*

Abstract

For solving the picture misappropriation problem in e-commerce, this paper proposes a content-based copyright detection method. In the method, a picture is smoothed by bilateral filter to ensure the integrity of its key points, and the shape feature and texture feature are extracted by Hu moment and Gabor filter respectively. Finally copyright-infringing pictures are detected by calculating the Euclidean distance and we use fuzzy heuristics to measure similarity between the query and the database images. The test dataset is crawled from the largest e-commerce website (Taobao.com). The copyright-infringing pictures can be detected by the method and the average accuracy reaches 91%, achieved the desired effect.

Keywords: *e-commerce; picture misappropriation; copyright detection; fuzzy inference; picture features*

1. Introduction

Presently, China's e-commerce is in a state of rapid growth. By the end of 2013, its gross merchandise volume (GMV) had reached US$300 billion, and China has surpassed the U.S. as the world's largest online retail market. However, counterfeit products and false advertising hamper the further growth of e-commerce. Some merchants have misappropriated and abused merchandise pictures belonging to other merchants or websites in advertising their own counterfeit products. This paper proposes a content-based copyright detection method for pictures. The method would allow e-commerce websites to detect illegal pictures used by merchants in real time.

Currently, digital watermarks are the most commonly method used for copyright protection of pictures. However, many deficiencies of digital watermarks have been noted. First, watermarks must be embedded prior to publication, and copyright protection or detection is impossible for unmarked works. Second, watermarks are vulnerable to hacking. Once cracked, a watermark system no longer offers any protection. For these reasons, content-based image protection has also been studied. The visual features of an image, such as shape, color, and texture, are directly related to its content [1]. Point feature is used for duplicate picture detection and sub-graph retrieval [2]. A smart CBIR application based on color and texture features is proposed in paper [3], but it is not effective against deformed (*e.g.* rotated) images. The implicit knowledge is extracted from the image data by carrying-out clustering and set the rules based on the relevance feedback given by experts in order to refine the results by improving clusters [4]. The growth of e-commerce has given new applications for theories and technologies of computer visuals, including that of picture retrieval. A recommendation system named the Mobile Merchandise Evaluation Service Platform (MMESP), which allows using product pictures for real-time product identification [5]. But the method is unfeasible for products

[*] Corresponding Author

ISSN: 2005-4246 IJUNESST

that do not have text markings, or ones that are difficult to describe using text, such as furniture.

2. Problems in Images Protection

For protecting the interests of merchants, digital watermarks are used in Taobao image copyright protection. Merchants should upload the images to apply for certification, and use the certificated images to show their products. But it still exists some limitations. First, the requirement to the image is quite high. There can be no stitching and text in the picture, and the picture should not be too simple and have no visual decoration. Currently, it can only be used in dress pictures. Second, it is vulnerable to hacking. Once cracked, the system no longer offers any protection.

In their misappropriation of pictures, merchants would often manipulate the original pictures in various ways, such as translation, rotation, cutting and pasting, insertion of text, and color changes. SIFT is capable of matching pictures that have undergone translation, rotation and other deformations. Shape features is invariant to location, orientation and translation, and can reflect the characteristics of image effectively. Texture is measured by the relative brightness but will still be affected by the image quality. We present a copyright detection approach, which is based on texture, shape and SIFT features.

3. Methodology

3.1. SIFT Feature Extraction Based on Bilateral Filter

Fundamentally, SIFT involves finding extreme points in the scale space. Traditional SIFT applications have used Gaussian filter for picture smoothing. In general, SIFT tends to select points in the corners and high-contrast regions, and disregard key points from the edges. Bilateral filter is a non-linear filter that achieves edge preservation and noise reduction [6-7].

We used bilateral filter for picture smoothing and generation of scale space. A bilateral filter uses pixel locations and brightness gradients to control the weight values. Points around regions of great brightness variations are given low weights for edge preservation, while regions of small brightness variations are given normal Gaussian weight values for smoothing.

Let $f(x,y)$ be the input picture, and $g(x,y)$ be the output picture after smoothing. The smoothing process of bilateral filter can be expressed as:

$$g(x,y)=\frac{\sum_{n=-w}^{w}\sum_{m=-w}^{w} f(x+m,y+n)exp(-\frac{m^2+n^2}{2\sigma_1^2})exp(-\frac{(f(x,y)-f(x+m,y+n))^2}{2\sigma_2^2})}{\sum_{n=-w}^{w}\sum_{m=-w}^{w} exp(-\frac{m^2+n^2}{2\sigma_1^2})exp(-\frac{(f(x,y)-f(x+m,y+n))^2}{2\sigma_2^2})} \quad (1)$$

Where σ_1 is the parameter that controls Gaussian shape in space and σ_2 is the parameter to control the effects of brightness changing? When σ_1 increases, the filter will cause a stronger blurring effect. When σ_2 decreases, the edges will be preserved. The purpose of using bilateral filter is to protect the edges, and ensure more key points along edges can be detected.

For each picture, its gradient magnitude $M(x,y)$ and orientation $O(x,y)$ can be expressed as:

$$M(x,y)=\sqrt{(\frac{\partial g(x,y)}{\partial x})^2+(\frac{\partial g(x,y)}{\partial y})^2} \quad (2)$$

$$O(x,y)=arctan(\frac{\partial g(x,y)}{\partial y}/\frac{\partial g(x,y)}{\partial x}) \quad (3)$$

Figure 1 shows the picture scale spaces generated after smoothing by bilateral and Gaussian filters, with the left using bilateral filter and the right using Gaussian filter.

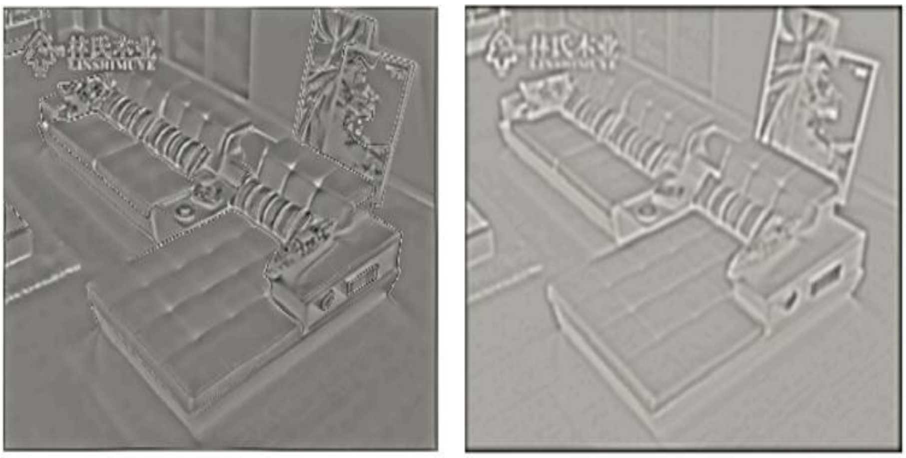

Figure 1. Comparison of Picture Scale Spaces Generated by Bilateral Filter and Gaussian Filter

The edges are better preserved in the left. Several obvious extremes can be seen in the picture processed with Difference of Bilateral, which are not found in the picture processed with Difference of Gaussian, as these values have been filtered out by the Gaussian filter.

Figure 2 shows the number of key points generated after smoothing by bilateral and Gaussian filters.

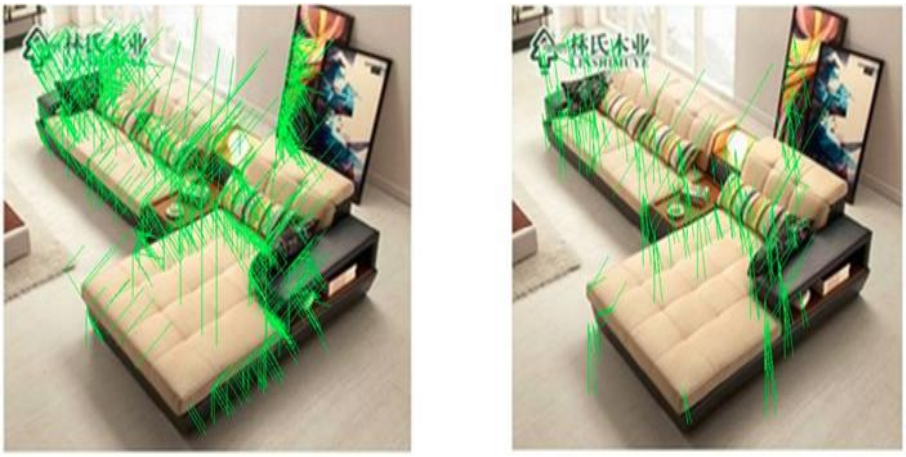

Figure 2. Comparison of Number of Key Points Generated by Bilateral Filter and Gaussian Filter

More key points are visible in the left, generated by the bilateral filter, than the right, generated by the Gaussian filter.

Figure 3. shows a comparison of key point matching with pictures smoothed by Gaussian and bilateral filters.

Figure 3. Comparison of Matching Results Using Bilateral Filter and Gaussian Filter

There are notably more matching key points from using the bilateral filter than the Gaussian filter. This is due to the edge protection effect of bilateral filters, which increased the amount of extracted information, and thus increased the number of matching features.

3.2. Extraction of Shape Feature

Hu moment which is invariant to location, orientation and translation is widely used in the area of image classification. The moment of order $(a+b)$ of an picture $I(x,y)$ is defined as $m_{a,b}=\sum_x\sum_y x^a y^b I(x,y)$, where $a,b=0,1,2$ .... The corresponding central moment is defined as $u_{a,b}=\sum_x\sum_y (x-\bar{x})^a (y-\bar{y})^b I(x,y)$, where $\bar{x}={m_{10}}/{m_{00}}$, $\bar{y}={m_{01}}/{m_{00}}$, which are referred to as region center. The scale invariant moments $\eta_{a,b}$ where$(a+b)\geq 2$can be defined as $\eta_{a,b}={\mu_{a,b}}/{\mu_{0,0}^{\gamma}}$, where $\gamma=\left[\frac{a+b}{2}\right]+1$. Then seven shape features $\emptyset_1—\emptyset_7$ which are scaling, rotation and translation invariants can be extracted. [8]

3.3. Extraction of Texture Features

The texture feature of images is commonly represented by co-occurrence matrix. Gabor filter is the most effective and widely used texture analysis method. Gabor wavelet function is sensitive to image edges, provides excellent direction and scale selectiveness, and is also insensitive to illumination changes, and tolerant to a certain degree of image rotation or deformation.

When a picture is processed by Gabor filter, the input is a convolution of picture $I(x,y)$and Gabor function $g(x,y)$. After applying Gabor filter on scale and orientation, an array can be obtained [9]:

$$E(m,n)=\sum_x\sum_y |G_{mn}(x,y)| \quad (4)$$
$$m=0,1,\ldots,M-1,\ n=0,1,\ldots,N-1$$

The magnitudes represent the energy content at different scale and orientation of image. Texture-based image retrieval allows find pictures or regions with similar textures. The texture feature of regions are expressed using mean μ_{mn} and standard deviation σ_{mn}:

$$\mu_{mn}=E(m,n)/(P\times Q) \quad (5)$$

$$\sigma_{mn} = \sqrt{\sum_x \sum_y (|G_{mn}(x,y)| - \mu_{mn})^2 / (P \times Q)} \quad (6)$$

Where M represents the scale; N represents the orientation; P and Q represent the height and width of the input image. The resulting mean $\mu_{m,n}$ μ_{mn} and standard deviation $\sigma_{m,n}$ σ_{mn} constitute two feature vectors respectively, which are then combined into a single feature vector as the texture descriptor.

3.4. Similarity Detection

Before the feature vectors are combined, they are normalized to reduce the effects of different feature dimensions and variances of the feature components.

$$F_D = [\omega_s \times \frac{f_s}{N_s \cdot \delta_s \mu_s}, \omega_t \times \frac{f_t}{N_t \cdot \delta_t \mu_t}] \quad (7)$$

Where and are the dimensions of shape and texture feature vectors; δ_s, δ_t and μ_s, μ_t are the average values and standard deviations of shape and texture; ω_s and ω_t are weights of shape and texture; $0 \leq \omega_s, \omega_t \leq 1$ and $\omega_s + \omega_t = 1$.

Once the feature vector is extracted, a similarity detection method can be used to compare the input image against images in the database. We chose Euclidean distance for our similarity detection:

$$d(\boldsymbol{F}^Q, \boldsymbol{F}^T) = \sqrt{\sum_{i=0}^{n-1} (F_i^Q - F_i^T)^2} \quad (8)$$

Where n is the dimensions of the feature vectors, $\boldsymbol{F}^Q$, $\boldsymbol{F}^T$ are feature vectors of the input picture and an picture from database.

For detecting the illegal images or the images similar to the original images, fuzzy inference method is used to detect the similarity. Shape features and texture features are two visual features of images that have been retrieved using Moment Invariants and Gabor wavelet. The first priority is given to the shape features, as shape of an image is not easily affected by external factors, and also it is invariant to the rotation, translation and orientation. The second priority is given to the texture features. The performance will be improved by defining these criteria along with the fuzzy rules. The Mamdani fuzzy inference method is used to perform fuzzy rules in our proposed approach [10].

We introduce a set of fuzzy rules to process the results achieved by applying the two algorithms discussed above. (1) We define a number of inputs. There are two inputs which are shape distance, and texture distance between query image and database images in our approach. (2) The membership functions for two types of input have been defined. The types of fuzzy set that identified each input as low, medium and high. Three types of output fuzzy sets have been declared such as high similar, medium similar and low similar. (3) A fuzzy rule can be defined as a conditional statement. Fuzzy rules applied using logical operator. To process the Mamdani fuzzy inference method. We take the crisp inputs and fuzzy them to determine the degree to which these inputs belong to each of the appropriate fuzzy set. The AND fuzzy operator is applied to get one number that represents the result of antecedent of rules. (4) The above fuzzy rules are used for data aggregation. (5)The aggregate output fuzzy set should transform to a single crisp number.

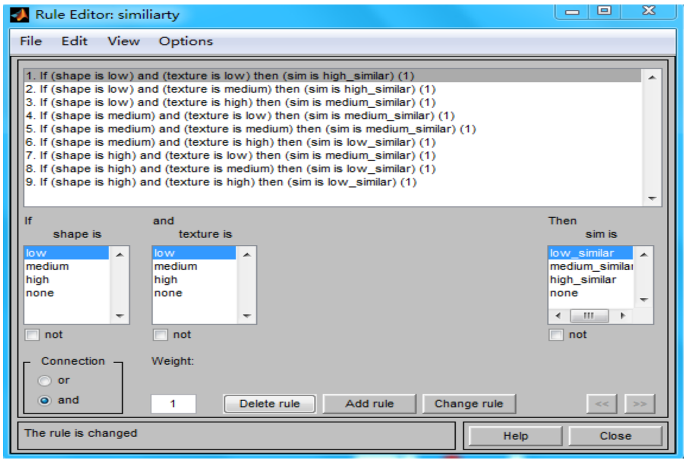

Figure 5. A Set of Fuzzy Rules Applied to Priorities Results

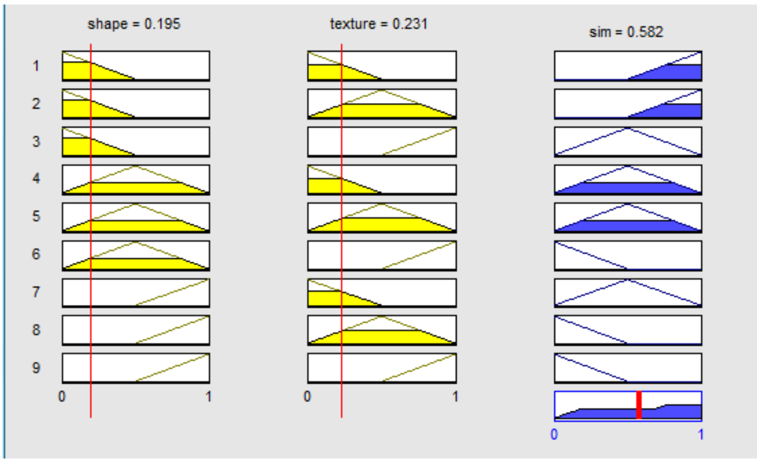

Figure 6. Fuzzy Rules Representation

4. Results

For verifying the feasibility of the method, we created a database using pictures crawled from Taobao. We searched for the key word “furniture” on Taobao and stored the 1000 pictures from the first 20 pages in our test database. As the method is based on content-based image retrieval systems, it would be compared to the methods proposed by literature [4] and literature [9].

In the follow 3 figures, the first picture is the query picture which the database pictures are compared against. Figure 7 shows the content-based detection method is capable of identifying pictures that are modified based on the query picture. The results of Figure 7 are a notable improvement on Figures 8 and 9, which are closer to each other. In e-commerce sites, for differ from the original pictures, people will make some changes to color using some technical measures. Therefore, the accuracy is not good enough when detect the copyright using color features. The measures proposed by ElAlami and Wang X Y *et al*. used color and texture features to retrieve pictures, for the above-mentioned reasons, the results are not good enough. Considering the practical situations in e-commerce site, we use texture, shape and SIFT features to detect the similarity, and the result is better than the results using the measures proposed by literature [4] and literature [9]. For a more direct comparison, the results can be described using precision and recall.

Figure 7. Results of our Method

Figure 8. Results of Method Proposed by Literature [4]

Figure 9. Results of Method Proposed by Literature [9]

The most commonly used indicators for evaluation of image retrieval systems are precision and recall. Precision represents the ratio of retrieved pictures relevant to the query, and recall represents the ratio of relevant pictures that have been retrieved. They can be defined as [9]:

$$precision = \frac{a}{a+b} \quad (9)$$

$$recall = \frac{a}{a+c} \tag{10}$$

Where a is the number of relevant pictures that retrieved, b is the number of irrelevant pictures retrieved, and c is the number of relevant pictures that have not been retrieved. The precision and recall ratios are then combined into a single parameter F-score:

$$Fscore = \frac{2 \times precion \times recall}{precision + recall} \tag{11}$$

The performances of three retrieval methods are calculated for comparison below.

Table 1. Comparison of Our Method Versus Two Existing Methods

	Precision	Recall	F-score
Our method	0.91	0.63	0.74
Method proposed by literature [4]	0.72	0.33	0.45
Method proposed by literature [9]	0.75	0.39	0.51

It is obvious that the method introduced in this study performed significantly better than the two existing methods. The other two methods had similar performances, with the latter slightly better than the former.

5. Conclusion

In view of the picture misappropriation problem in e-commerce, this paper proposes a content-based copyright detection method. Copyright-infringing pictures can be detected by the method. We conducted tests using pictures crawled from Taobao, China's largest e-commerce website, which resulted in an average accuracy of 91%. A comparison of results using precision and recall as indicators for evaluation shows that the method had significantly better performance than existing methods, and greatly improved the results of detection.

More work should be done in order to improve the efficiency of detection and make the method more practical for databases with massive amounts of pictures. In addition, it is a common practice in the industry for multiple licensees to share the use of pictures provided by one distributor. Removing this type of pictures from detection results will further improve the accuracy.

References

[1] K. Shkurkol and X. Qi, “A Radial Basis Function And Semantic Learning Space Based Composite Learning Approach To Image Retrieval”, IEEE International Conference on Acoustics, Speech and Signal Processing, Honolulu, USA, **(2007)** April 15-20.

[2] D. G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision, vol. 60, no. 2, **(2004)**, pp. 91-110.

[3] C. H. Lin, R. T. Chen and Y. K. Chan, “A smart content-based image retrieval system based on color and texture feature”, Image and Vision Computing, vol. 27, no. 11, **(2009)**, pp. 658-665.

[4] M. E. Elalami, “A novel image retrieval model based on the most relevant features,” Knowledge-Based Systems, vol. 24, no. 1, **(2011)**, pp. 23-32.

[5] C. C. Lo, T. H. Kuo, H. Y. Kung, H. T. Kao, C. H. Chen, C. I. Wu and D. Y. Cheng, “Mobile merchandise evaluation service using novel information retrieval and image recognition technology,” Computer Communications, vol. 34, no. 2, **(2011)**, pp. 120-128.

[6] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” Sixth International Conference on Computer Vision, Bombay, India, **(1998)** January 04-07.

[7] M. Tanaka and M. Okutomi, “Latent common origin of bilateral filter and non-local means filter,” IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, San Jose, USA, **(2010)** January 17.

[8] M. K. Hu, “Visual Pattern Recognition by Moment Invariants”, IRE Transactions on Information Theory, vol. 8, no. 2, **(1962)**, pp. 179-187.

[9] X. Y. Wang, H. Y. Yang and D. M. Li, "A new content-based image retrieval technique using color and texture information", Computers & Electrical Engineering, vol. 39, no. 3, **(2013)**, pp. 746-761.
[10] D. R. Keshwani, D. D. Jones, G. E. Meyer and R. M. Brand, "Rule-based Mamdani-type fuzzy modeling of skin permeability", Applied Soft Computing Journal, vol. 8, no. 1, **(2008)**, pp. 285-294.

Authors

Liyi Zhang, is a professor in Wuhan University, China. His research interests include theory and application of electronic commerce, ubiquitous network integration and information system integration, and big data analysis.

Chang Liu, he is now studying the PHD of Management Science and Engineering in Wuhan University, and the research interest is electronic commerce.