Feature Extraction for Effective Content-Based Cloth Image Retrieval in E-Commerce

Lingli Li(✉) and JinBao Li

Heilongjiang University, Harbin, China
Lilingli_grace@163.com, jbli@hlju.edu.cn

Abstract. Cloth image retrieval in E-Commerce is a challenging task. In this paper, we propose an effective approach to solve this problem. Our work chooses three features for retrieval: (1) description (2) category (3) color features. It can handle clothes with multiple colors, complex background, and model disturbances. To evaluate the proposed method, we collect a set of women cloth images from Amazon.com. Results reported here demonstrate the robustness and effectiveness of our retrieval method.

Keywords: Information retrieval · E-Commerce · Content-based image retrieval

1 Introduction

In recent years, with the rise of E-Commerce and the perfection of logistics services, online shopping has gradually replaced the traditional shopping way, and becomes a new fashion. Throughout famous shopping websites, such as http://www.like.com and http://www.amazon.com/, they all provide the users product retrieval service. However, most of them provide keyword-based or category-based interface for search. Such search interfaces are efficient for the products which could be identified with the name or the category, such as books and digital products. However, they are not suitable for clothes, since the major features of clothes are visual ones and difficult to express in words. For example, it is hard to describe some clothes designed with complex patterns and multiple colors by any keywords or category. Thus, visual information should be taken into consideration for clothes search in E-Commerce and Clothes REtrieval Based on Image(CREBI) is in demand.

CREBI brings following technical challenges.

(1) The first challenge is how to describe the visual features of a cloth image. Many works have been devoted to representing local features like sift, surf, and etc. A study from Nanjing Univ. of Aeronaut. & Astronaut. [6] analyzes the effects of different local features on product image retrieval. It was tested on 12 kinds of product, among which glasses and bikes have best performance. They only extracted the features from certain regions that best describe the product. However, unlike these products mentioned above, cloths images are more complicated since customers are willing to look at global visual features to make decision. Therefore, local features are insufficient to meet the needs in CREBI.

W. Che et al. (Eds.): ICYCSEE 2016, Part I, CCIS 623, pp. 374–383, 2016.
DOI: 10.1007/978-981-10-2053-7_33

(2) Second, which global feature could have a better performance applied into cloth image retrieval? Global features include texture, color and shape, etc. Many researchers have involved studies about texture and shape and found it is easy to extract these features from cloth images [4, 7], but the extraction lacks precision since E-Commerce cloth is displayed in variant shape and its soft property greatly affects the extraction. Also, it is hard to speed up the matching process[8, 9] and has a bad user experience of real-time search which is critical for E-Commerce sites.
(3) Cloth image is complicated and its background is likely to be mistakenly considered as part of the cloth during the feature extraction. Various light conditions, complex patterns, and model's skin color also increase the difficulty of feature identifications.

According to these challenges, the crucial step for effective CREBI is to extract useful features effectively. Thus in this paper, we focus on the extraction of features that are suitable for CREBI in E-Commerce. We have following contributions.

(1) With the consideration of both effective and efficiency issues, we choose Categories, Description, and Color as the features for CREBI in E-Commerce. Experimental results demonstrate that such features are sufficient for retrieving the most relevance results.
(2) As to extract the color feature from the images in E-Commerce that may contain models and other confusing noise such as mottled gray background, we develop two methods. The former attempts to remove the color same as human skin to erase the disturbance of models. The latter removes the background of clothes by threshold-based filtering strategy. Experimental results show that these two methods could increase the quality of retrieval significantly.
(3) For efficient and effective CREBI, to reduce the size of candidates and accelerate the retrieval, we apply a hierarchical filtering method to retrieve similar cloth images by color feature, text description and category.

The paper is organized as follows. Section 2 introduces the framework of our system. Section 3 focuses on the reasons of the selected features. Section 4 highlights the process of extracting features. Experiment results are summarized in Sect. 5. Section 6 concludes the work.

2 Framework

We implement a system to search similar product image according to search inputs which could be the arbitrary combination of image, keywords and category. To balance the efficiency and effectiveness, we select three features: color, text descriptions, and categories. The details of feature selection will be discussed in Sect. 3.

Our approach first extracts an 11-dimensional color vector from each image in dataset. As illustrated by Fig. 1, we first upload a query cloth image as input. Then, the color vector V for the input image is extracted. We retrieve top-K images with the color

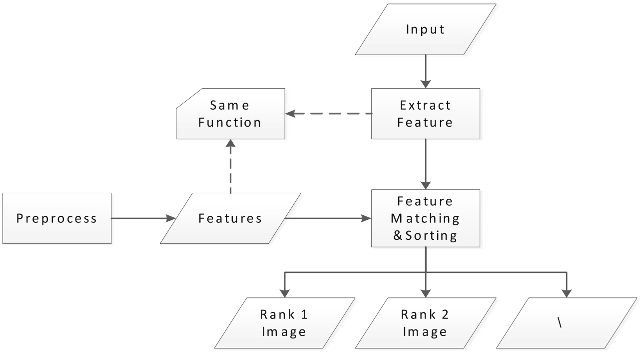

Fig. 1. Flow chart of image retrieval system

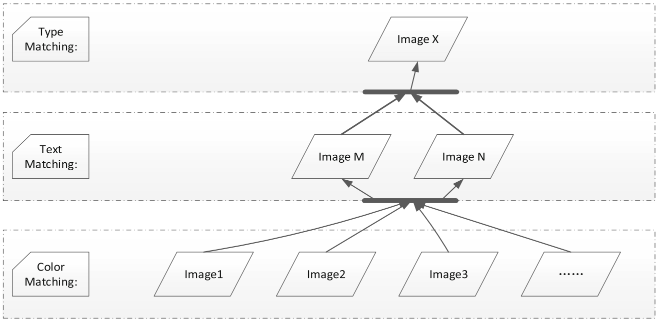

Fig. 2. Three layers of feature matching.

vectors the nearest to V in term of Euclidean distances as in Eq. 1, and could quickly output the orders of corresponding images.

$$\mathrm{dist}(\vec{u}, \vec{v}) = \sqrt[2]{\sum\nolimits_{i=1}^{n} \left(u^{(i)} - v^{(i)}\right)^2} \tag{1}$$

Meanwhile, in order to improve the precision of searching results, we also support query by text description and category, and we have made some betterment based on that. Therefore, every retrieval result of our system is filtered in the order of the efficiency of the filtering, that is, text description, categories, and color features, as shown in Fig. 2. In practice, if we first search by words and categories, it could help us narrow the search scope which may greatly accelerate the search algorithm.

3 Feature Selection

In this section, we will discuss the motivations of selected features, that is description text, category and color feature.

First, we use global features instead of local ones. Most local features are applied into the detection of objects with some nice properties like scale invariance, rotation invariance [10–12] and etc. However, they have a bad performance in cloth images of E-Commerce for the reason that their detection could be easily affected by some unique properties of cloth image. Issues, like variant light source and cloth model, will make negative contributions to local feature detection and matching. As shown in Fig. 3, surf feature [11] matching on E-Commerce images may mislead the matching. It detection mainly concentrates on the matching of model's head and leg in cloth image, and fails to match the clothes. However, most of E-Commerce images are displayed with models wearing cloth product, so local features could not satisfy our need.

It is known that global features involve texture, shape, and color. Among them, we choose color feature as the global visual feature, due to the following two reasons.

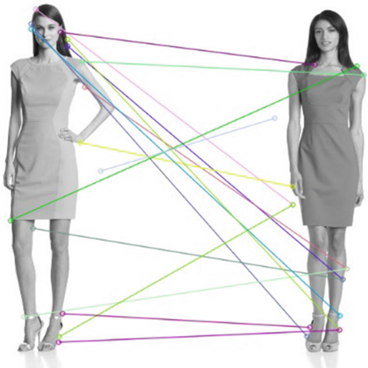

Fig. 3. SURF extraction and matching.

Fig. 4. Image background mixed with color

(1) Most cloth images contain both cloth and model since online shop owners ask models to try on their clothes in order to provide an overall effect for customers and motivate their purchases. In this case, the cloth image may be very diverse and have different outlooks. The model's body and pose will affect the extraction of shape feature and texture feature in a great extent.

(2) Since clothes are deformable, shadows and wrinkles may be confused as part of the texture pattern. Arbitrary shooting distance and viewing direction when taking

the cloth pictures present a variant shape of cloth. Some cloth pictures are taken as full-length portrait while some are half-length. The issues above may cause many matching errors for texture and shape. In contrast, color is invariance with the position and angle of the cloth in the image.

From above discussion, for cloth images, color feature is relative stable and the cloth shape has little influence on its color. As a result, we use color vector to represent the color feature. This vector contains 11 dimensions, each entry of which is the percent of the pixels with corresponding color. Besides 8 colors in [3], we divide yellow into orange, light orange and yellow, and blue into blue and cyan.

However, color is required but only color may lead to a low accuracy. Text descriptions and category are also in great need of specifying a cloth image. They, together, could help us improve the search precision. And the category is particularly important for us to describe our purchase requirements. Despite the fact that we could filter the clothes into a certain color group with their color feature, without category, we still could not search out what we want exactly, like a purple dress or a purple coat.

Since the extraction techniques for text description and category are well-studied [1, 2]. The effective extraction of color features is the centric issue of CREBI. In the next section, we will discuss how to extract color features effectively.

4 Feature Extraction

For a given image, we use a general color extraction approach [3]. Firstly, for RGB space, we extract black and white color since they are difficult to extract from HIS space. According to our observation, we set the rule of black and white color extraction as follows. For a color (R, G, and B) satisfying:

$$\begin{aligned} & |R-B| \leq 3 \;\&\&\; |R-G| \leq 3 \;\&\&\; |G-B| \leq 3 \\ & \text{where } \; R, G, B \epsilon [0, 255] \end{aligned} \tag{2}$$

It is treated as white or black. While all the values of (R, G, and B) $\in (215, 255]$ denote white, others are black.

Then we extractor other colors in HIS space. We detect colors based on saturation S, luminance I, and hue H [3]. Since the rest pixels are not white or black, we only employ the hue information for the other 9 entries in the color vector, Eq. 3 [5] shows the transferring method and θ. dotes the angle of Hue value.

$$\theta = \cos^{-1}\left(\frac{\mathrm{R-G+(R-B)}}{\sqrt{(R-G)^2 + (R-B) + (G-B)}}\right)$$

$$\mathrm{H} = \begin{cases} \theta & G \geq B \\ 2\pi - \theta & G < B \end{cases} \tag{3}$$

Algorithm 1. Color Percentage Computation

Input: *color percentage after previous step*

Output: *final color percentage*

For percentage (white)

IF percentage (white) > threshold (t) **THEN**

$$\text{Percentage(white)} = \frac{white - area * t}{area * (1 - t)}$$

ELSE

$$\text{Percentage(white)} = 0$$

FORother colors*c***DO**

$$\text{Percentage}(c) = \frac{c}{area * (1 - t)}$$

return Percentage

As shown in Fig. 5, hue is displayed as a 360 °color wheel. We define the color "red" between 0–9 ° and 345 °–360 °, "orange" in the range of 9 °–42 °, "light orange" between 42 °–62 °, "yellow" between 62 °–75 °, "green" between 75 °–160 °, "cyan" between 160 °–210 °, "blue" between 210 °–280 °, "purple" between 280 °–315 °, "pink" between 315 °–345 °.

Intuitively, the extraction of color is trivial. However, the complexity in the images in E-Commerce brings following difficulties in this step.

(1) Models disturbance
(2) Confusing background noise

Fig. 5. Hue representation in HIS color (Color figure online)

Therefore, in this section, we focus on the solution of these difficulties. After our approaches are applied, the color features could be extracted directly with the methods in description, category, and color feature.

4.1 Remove Skin Color

Intuitively, clothing manufactures will not produce clothes that have the same color as people's skin. Hence, we figure out the color spectrum of people's skin and get rid of colors in this scope (Eq. 4 and 5). We define the following scopes as skin color [5].

$$R \in [215, 240] \,\&\&\, G \in [160, 190] \,\&\&\, B \in [120, 180] \,\&\&\, G - B \geq 12 \tag{4}$$

$$R \in [100, 200] \,\&\&\, G \in [55, 160] \,\&\&\, B \in [0, 145] \,\&\&\, G - B \geq 12 \,\&\&\, R - B \geq 30 \tag{5}$$

To sum up, a certain dimension of the color vector space satisfies equation as follows.

$$\text{Percentage} = \text{color}/(\text{area} - \text{Backgrond\_White} - \text{skin}) \tag{6}$$

Here *Percentage* is how much a certain color accounts for in an image; *area* reflects image's pixels; *Backgroud_White* represents the pixels of white background; *skin* is the area of people's skin.

4.2 Remove Background Noise

As showed in Fig. 4, these images are bright and colorful, but their background is not pure white and mixed with some color noises. In order to solve this, we select 500 representative images for a statistical analysis. And we finally find out that this kind of "bad" image accounts for average 60 % (the threshold) pixels in an image of all the research samples. Therefore, we revise the process step above as following algorithm.

Fig. 7. Results without category.

Fig. 6. .

Fig. 8. Results with category.

Fig. 9. Sample image for model

5 Experimental Results

5.1 Databases

The robustness and effectiveness of our approach are evaluated on 21598 pieces of women cloth image from Amazon.com. Categories cover seven kinds: Wedding dresses, Wear to Work, Special Occasion dresses, Jackets & Coats, Casual, Cardigans Sweaters, and Active Shirts & Tees. We explore on colors including white, yellow, red, purple, black, and even multiple color. Some of the clothes are dressed by models, while others are not. Most of them have a pure white background, while others may be represented as mottled gray. The dimension of each image is around 385*500 pixels. Since for a large image set, the result set may be very large. Thus a user cares precision more than recall. Therefore, we only use precision to test the quality of search results.

5.2 The Effectiveness of Selected Features

We return 15 results from each retrieval tasks according to color matching. After testing for about 500 times, our method achieves 93 % accuracy rate. Some successful results of CREBI are shown in Figs. 7 and 8. Figure 6 is an example of our input image. Comparing the retrieval outputs in Figs. 7 and 8, it presents a list of images from the same retrieval with category search and without respectively. This show the selection of category could help to improve the quality of search result.

5.3 The Effectiveness of Feature Extraction

Model Disturbance. In our test, we evaluated model's skin color's influence on the proposed method. The testing results of cloth image (showed in Fig. 9) are summarized in Table 1. After our revision, the percentage of yellow in the image greatly reduces. In this case, model's disturbance is minimized.

Table 1. Comparison before and after skin color removal.

Colors	Before remove skin	After remove skin
White	0.00	0.00
Black	51.17	74.32
Yellow	43.71	18.52
Red	3.82	5.26
Blue	0.03	0.04
Green	0.14	0.21
Purple	0.12	0.17
Pink	0.69	1.00
Cyan	0.34	0.49

Fig. 10. .

Fig. 11. Before removing noises.

Fig. 12. After removing noises.

5.4 Background Noise

After removing the background noises, we could see an obvious improvement displayed in Fig. 11, and Fig. 12 when we upload a query image showed in Fig. 10.

5.5 Summary

In summary, we have following conclusions from the experimental results.

(1) Text description, category, and color features are suitable for CREBI in E-Commerce.
(2) Our image processing methods could improve the accuracy of CREBI in E-Commerce significantly.

6 Conclusions and Future Work

Content-based image retrieval for cloth plays an important role in E-Commerce. However, current image-retrieval techniques are not suitable for the clothes images on E-Commerce web sites. To solve this problem, we propose a Clothes REtrieval Based on Image (CREBI) method in this paper. We select suitable features for CREBI and design image processing strategies to increase the accuracy for the CREBI. Experimental results demonstrate that proposed method is suitable for CREBI. Further work include scale the methods to large image set and involve new features for more effective CREBI.

Acknowledgment. This paper was partially supported by National Sci-Tech Support Plan 2015BAH10F01 and NSFC grant U1509216,61472099,61133002. The Scientific Research Foundation for the Returned Overseas Chinese Scholars of Heilongjiang Provience LC2016026.

References

1. Tapeh, A.G., Rahgozar, M.: An ontology-based semantic extraction approach for B2C ecommerce. Int. Arab J. Inf. Technol. **8**(2), 163–170 (2011)
2. Ding, Y., Korotkiy, M., Omelayenko, B., et al.: Goldenbullet: automated classification of product data in e-commerce. In: Proceedings of the 5th International Conference on Business Information Systems (2002)
3. Yuan, S., Tian, Y.L., Arditi, A.: Clothing matching for visually impaired persons. Technol. Disabil. **23**(2), 75–85 (2011)
4. Guo, Z., Zhang, L., Zhang, D.: Rotation invariant texture classification using LBP variance (LBPV) with global matching. Pattern Recogn. **43**(3), 706–719 (2010)
5. Vezhnevets, V., Sazonov, V., Andreeva, A.: A survey on pixel-based skin color detection techniques. Proc Graphicon **3**, 85–92 (2003)
6. Zou, G., Ma, R., Ding, J., et al.: Research on product image retrieval with local features. In: 2010 Chinese Conference on Pattern Recognition (CCPR), pp. 1–5. IEEE (2010)
7. Tseng, C.H., Hung, S.S., Tsay, J.J., et al.: An efficient garment visual search based on shape context. WSEAS Trans. Comput. **8**(7), 1195–1204 (2009)
8. Wu, Q., Yu, Y.: Feature matching and deformation for texture synthesis. ACM Trans. Graph. (TOG) **23**(3), 364–367 (2004)
9. Gong, M., Yang.Y.H.,: Near real-time reliable stereo matching using programmable graphics hardware. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, pp. 1:924–1:931. IEEE (2005)
10. Ke, Y., Sukthankar, R.: PCA-SIFT: a more distinctive representation for local image descriptors. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2004, vol. 2, pp. II-506–II-513. IEEE (2004)
11. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: speeded up robust features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006, Part I. LNCS, vol. 3951, pp. 404–417. Springer, Heidelberg (2006)
12. Juan, L., Gwun, O.: A comparison of sift, pca-sift and surf. Int. J. Image Process.(IJIP) **3**(4), 143–152 (2009)