19th International Computer Science and Engineering Conference (ICSEC)
Chiang Mai, Thailand, 23-26 November, 2015

Classification of Clothing with Weighted SURF and Local Binary Patterns

Wisarut Surakarin
Department of Computer Engineering,
Faculty of Engineering, Chulalongkorn University,
Bangkok, Thailand
Wisarut.Sur@student.chula.ac.th

Prabhas Chongstitvatana
Department of Computer Engineering,
Faculty of Engineering, Chulalongkorn University,
Bangkok, Thailand
prabhas@chula.ac.th

***Abstract*—This work proposes a recognition system for clothing classification by computer vision. The input is an image of the type of fashion catalog where the clothes are fully exposed with models showing their faces. For the preprocessing and features extraction the Bag of Features (BoF) is employed. There are four steps in the proposed classification method: (i) the cloth in an image is identified and located, then it is segmented by GrabCut; (ii) the area in the image of cloth is divided into three sub-windows (right side, middle, and left side); (iii) the feature extraction, Speed-Up Robust Features (SURF) and Local Binary Patterns (LBP) are applied to each sub-window to create a codebook; (iv) the classification is done by Support Vector Machine (SVM). Our dataset consists of total 1131 images out of which the training set is 991 images and the remainder is the testing set. We separate types into seven categories of clothing image which included, jacket, shirt, suit, sweater, t-shirt, polo-shirt and tank top. The result of the experiment illustrates that the proposed method can recognize types of clothing images accurately 73.57%.**

Keywords— Bag of Features, Local Binary Patterns, clothing recognition and clothing classification

I. INTRODUCTION

The appearance of clothing is very important in daily life to represent the lifestyle, culture and social status. Clothing gives us a lot of information. For example, we can recognize the people by the clothes they wear. Many contexts can be inferred from the type of clothes, for example, the people who wearing a suit shows that they were in situation of business. However, recognizing the type of clothes is a challenging problem in the area at computer vision.

There are many researchers develop methods for searching and content-based image retrieval. M. Mizuochi et al. [1] implemented clothing retrieval by using multiple image for search similar clothing image in online shopping market. We focus only image recognition based on Bag of Features.

Bag of Feature (BoF) [2] is inspired by the Bag of Words (BoW) model for text analysis. The Bag of Features represents image as a words and applied text analysis method to image classification. We introduce BoF to classify types of clothing. Firstly, we represent the key points in the image and using SURF to extract the key points. Secondly, all of key points are clustered by K-mean clustering. Finally, a codebook is created and using this codebook to classify types of clothing. The disadvantages are that SURF is not good at handling illumination/light changes. We use LBP feature to describe local textures and surface characteristics before extract the key points because of The LBP feature is robust under random noise and illumination/light changes.

We separate three sub-windows before detecting the key points in the image. In this proposed, we focus on upper body clothing to classify types of clothing into seven categories that included, jacket, shirt, suit, sweater, t-shirt, polo-shirt and tank top. The dataset is acquired from online e-commerce, social network, and Google image. The result shows that the proposed system can correctly classify the sample in average 73.57%.

We have organized the rest of this paper as follows. Section 2 describes related works. Section 3 proposes method of clothing recognition and classification. Section 4 reports the experiment and result of this work. The conclusion is presented in Section 5.

II. RELATED WORKS

Many researchers have proposed techniques for cloth search and retrieval. A. Nodari et al. [3] and Cushen and Nixon [4] proposed content-based image retrieval to query image form online shopping in mobile application. There are less work on clothing recognition and classification. S. Miura et al. [5] proposed dividing clothing image into five areas which include, shoes, bottom, top inner, top outer and search clothing image by same area detection. X. Yuan et al. [6] present integration of SIFT-LBP based on Bag of Features to achieve the superior performance on a accuracy score. S. Banerji et al. [7] improved classification performance for scene image by Bag of Features LBP descriptor. In this work, we do not focus on clothing image retrieval. We are improving classification performance of clothing by being more selective in the sub-area of the clothes combine with weighted SURF and LBP descriptor.

III. PROPOSED METHOD

The overall design method is shown in Figure 1. The following subsection describes the processing steps how to recognize the types of clothing in the image.

978-1-4673-7825-3/15/$31.00 ©2015 IEEE

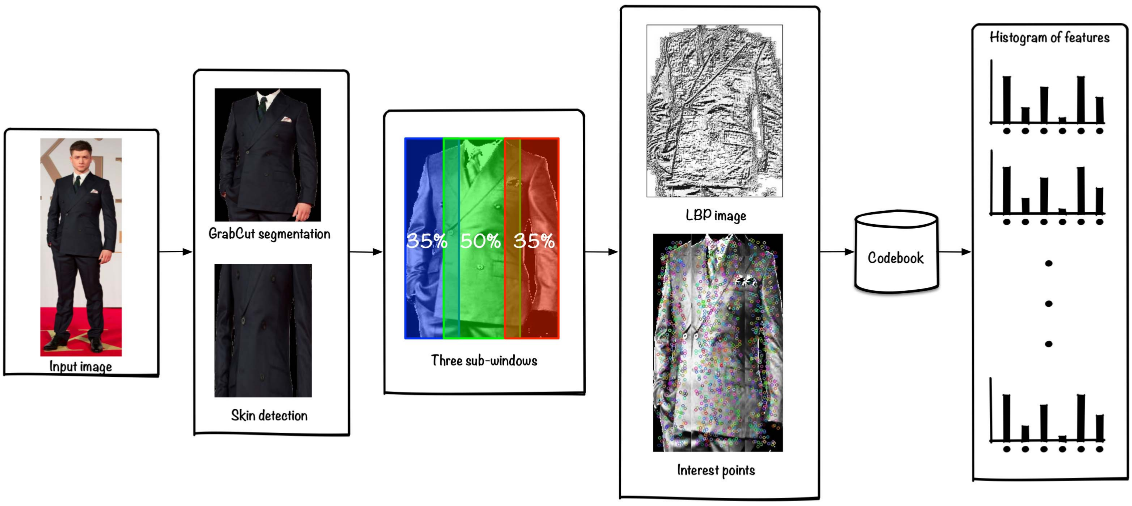


Fig 1. Overall propose method

A. Clothing Segmentation

We firstly detect facial area in the image by using face detection algorithm. Viola-Jones [8] propose face detection algorithm to determines the locations and the sizes of the human face in the image. We extract the upper body area based on the locations and the sizes of the detected face area by using human proportion [9] The human proportion can be measured all body parts by head size. After that, we crop the upper body area with regions of interesting (ROI) using bounding box as shown in Figure 2.

Fig. 2. Bounding box of upper body and face

We attempt to eliminate the non-clothing or background from the bounding box of the upper body by using GrabCut [10] segmentation. The GrabCut segmentation was popular algorithm to segment foreground and background. We set area of foreground in the bounding and set a few pixels area around boundary of an image as a background. Figure 3 shows the result of segmentation by using GrabCut.

The bounding box included clothing and human skin color. The skin is not good texture to extract the features. We attempt to eliminate the skin by using color segmentation. Albiol et al. [11] reported that every color space there exists an optimal skin detection. We change color space form RGB to HSV. Based on this work, we set a threshold method for skin pixel using range H = [0, 42], S = [32, 235], V = [60, 255]. Figure 3 shows the mask of skin pixel. If a pixel(x, y) is not skin color will be set to one, if not skin color will be set to zero.

(1) (2)

Fig. 3. Output image form (1) GrabCut detection and (2) skin detection

B. Feature extraction

Preprocess and normalize a bounding box in the image. After that, we use SURF and LBP feature based on Bag of Features to extract the features. The area is divided into three sub-windows to improving the performance.

1) Preprocessing and Normalize Image

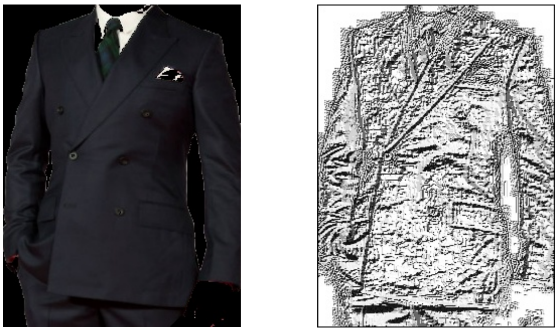
The bounding box was preprocessed and normalized the size of upper body to $size_{x,y}(256, 512)$ in all of the image dataset. We divided the bounding box of an image into three sub-windows which include right-side, middle-side and left-side. The areas of sub-windows are not equally important in all three sub-windows. We set the area of right-side and left-side to 35% of image sizes and middle-side set to 50% of image sizes.

2) Local Binary Patterns (LBP)

Local Binary Patterns (LBP) [12] is a powerful approach for extract local feature descriptors in the image. We use LBP to filter out of the noises where illumination/light changes. The

LBP operator was assign a label to every pixel of an image by threshold 3×3 neighborhood with the center value. The result of LBP code converts into binary number.

Figure 4, shows the output of LBP code. The input image is retrieved from Google image [13]



(1) (2)

Fig. 4. The output of LBP code

3) Speed-Up Robust Features (SURF)

Speed-Up Robust Features (SURF) [14] is inspired by the Scale-invariant feature transform (SIFT) descriptor. SURF is good at handling to changes in rotation, scaling and a few changes in viewpoints. We use SURF to detect the interest points as called *keypoints* for each sub-window from LBP image. In this work, we optimize weight of interest points in the middle-side more than three times of the right-side and the left-side. We believe that the middle-side has the distinction of types of clothing, for example, the middle-side of the shirt has a length of a zip which difference from polo shirt, T-shirt and sweater. Figure 5 shows the interest points of each sub-window to generate visual vocabulary as called *codebook*.

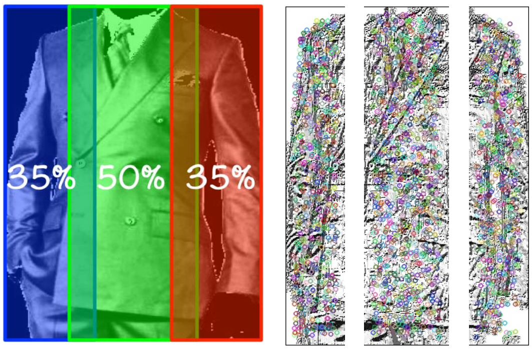


Fig. 5. The interest points of each sub-window

C. Clothing Classification

The Bag of Features (BoF) model [2] is used to classify the types of clothing with Support Vector Machine (SVM). In computer vision, the BoF was used to create visual vocabulary or codebook and represent the dictionary by K-mean clustering. We use 1,024 words for K-mean clustering. The output from cluster by K-mean is a histogram of features. The histogram of features was used to recognize and classify types of clothing by Support Vector Machine (SVM).

IV. Experiment and Result

First, we show the image dataset for recognition and classification. Next, the result for comparing with SIFT, SURF and SURF with sub-windows. An overview for each category of the results can be shown in Figure 6.

The image dataset includes upper body, full body and frontal face with clear background from Google image, online e-commerce and imageNet [15] Table 1 shows that our dataset consists of seven categories out of which training set are 991 images and test set are 140 images.

TABLE I. SEVEN CATEGORIES DATASET

	Training set	Test set
Jacket	140	20
Shirt	120	20
Suit	170	20
Sweater	135	20
T-shirt	121	20
Polo-shirt	160	20
Tank top	145	20

We selected Support Vector Machine (SVM) to train all of features with linear kernel. We achieve an accuracy score of 73.57% as shown in Table 3. The confusion matrix is shown in Table 2, which columns represent prediction class and rows represent actual class.

TABLE II. CONFUSION MATRIX OF OUR PROPOSED

	Jacket	Shirt	Suit	Sweater	T-Shirt	Polo-Shirt	Tank top
Jacket	18	2	0	1	0	0	1
Shirt	0	10	1	5	0	0	1
Suit	2	2	19	0	0	0	0
Sweater	0	6	0	12	0	2	2
T-shirt	0	0	0	2	18	3	0
Polo-shirt	0	0	0	0	2	11	1
Tank top	0	0	0	0	0	4	15

TABLE III. F-SCORE REPORT OF OUR PROPOSED

	Precision	Recall	F-score
Jacket	90.00%	81.82%	85.71%
Shirt	50.00%	58.82%	54.05%
Suit	95.00%	82.61%	88.37%
Sweater	60.00%	54.55%	57.14%
T-shirt	90.00%	78.26%	83.72%
Polo-shirt	55.00%	78.57%	64.71%
Tank top	75.00%	78.95%	76.92%
Overall	73.57%	73.37%	72.95%

Table 4, we compare four methods which include, SIFT feature, SURF feature, SURF with sub-windows and SURF-LBP with sub-window. The report of precision and recall, we represent by using F-score criterion.

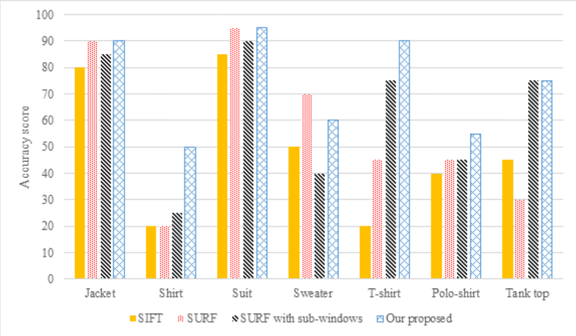


Fig. 6. Overview for each category

TABLE IV. THE REPORT OF COMPARE IN TERMS OF F-SCORE

	SIFT	SURF	SURF with sub-window	Our proposed
Precision	48.57%	57.86%	62.14%	73.57%
Recall	47.87%	59.47%	60.32%	73.37%
F-score	46.87%	55.81%	60.62%	72.95%

V. CONCLUSION

In this paper, we proposed recognition types of clothing by using a combination of SURF and LBP base on Bag of Features. The three sub-windows can optimize and improve the performance of interest point detection with height accuracy score. The experiment showed the proposed method achieves precision score 73.57%.

REFERENCES

[1] M. Mizuochi, A. Kanezaki, and T. Harada, "Clothing Retrieval Based on Local Similarity with Multiple Images," presented at the Proceedings of the ACM International Conference on Multimedia, Orlando, Florida, USA, 2014.

[2] S. O'Hara and B. A. Draper, "Introduction to the Bag of Features Paradigm for Image Classification and Retrieval," *Computing Research Repository (CoRR),* 2011.

[3] A. Nodari, M. Ghiringhelli, A. Zamberletti, M. Vanetti, S. Albertini, and I. Gallo, "A mobile visual search application for content based image retrieval in the fashion domain," in *Content-Based Multimedia Indexing (CBMI), 2012 10th International Workshop on*, 2012, pp. 1-6.

[4] G. A. Cushen and M. S. Nixon, "Mobile visual clothing search," in *Multimedia and Expo Workshops (ICMEW), 2013 IEEE International Conference on*, 2013, pp. 1-6.

[5] S. Miura, T. Yamasaki, and K. Aizawa, "SNAPPER: Fashion Coordinate Image Retrieval System," in *Signal-Image Technology & Internet-Based Systems (SITIS), 2013 International Conference on*, 2013, pp. 784-789.

[6] X. Yuan;, J. Yu;, Z. Qin;, and T. Wan, "A SIFT-LBP IMAGE RETRIEVAL MODEL BASED ON BAG-OF-FEATURES," *Proceedings of the International Conference on Image Processing (ICIP 2011),* pp. 1061-1064, 2011.

[7] S. Banerji, A. Sinha, and C. Liu, "A New Bag of Words LBP (BoWL) Descriptor for Scene Image Classification," in *Computer Analysis of Images and Patterns*. vol. 8047, R. Wilson, E. Hancock, A. Bors, and W. Smith, Eds., ed: Springer Berlin Heidelberg, 2013, pp. 490-497.

[8] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," *Int. J. Comput. Vision,* vol. 57, pp. 137-154, 2004.

[9] D. A. Jusko. (2015, July, 7). *Human Figure Drawing Proportions*. Available: http://www.realcolorwheel.com/human.htm

[10] C. Rother, V. Kolmogorov, and A. Blake, ""GrabCut": interactive foreground extraction using iterated graph cuts," presented at the ACM SIGGRAPH 2004 Papers, Los Angeles, California, 2004.

[11] A. Albiol, L. Torres, and E. J. Delp, "Optimum color spaces for skin detection," in *Image Processing, 2001. Proceedings. 2001 International Conference on*, 2001, pp. 122-124 vol.1.

[12] T. Ahonen, A. Hadid, and M. Pietikainen, "Face Description with Local Binary Patterns: Application to Face Recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 28, pp. 2037-2041, 2006.

[13] (2015, july, 27). *Kingsman: The Secret Service*. Available: https://www.facebook.com/KingsmanTheSecretServiceThailand

[14] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded Up Robust Features," in *Computer Vision – ECCV 2006*. vol. 3951, A. Leonardis, H. Bischof, and A. Pinz, Eds., ed: Springer Berlin Heidelberg, 2006, pp. 404-417.

[15] D. Jia, D. Wei, R. Socher, L. Li-Jia, L. Kai, and F.-F. Li, "ImageNet: A large-scale hierarchical image database," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, 2009, pp. 248-255.