On-the-fly Fashion Photograph Recommendation System with Robust Face Shape Features

Wan-Yu Chen, Jia-Lin Chen, and Liang-Gee Chen, *Fellow, IEEE*
DSP/IC Design Lab., Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan

***Abstract*-- In this paper, we demonstrate a novel method for retrieving fashion photographs with robust face shape features, a challenging problem due to the different face orientations, variant face size, and variation in lighting environment. In addition, we provide a novel dataset for labeling eight face shape categories, to enable future research on face shape classification. We also present experimental results through the proposed face shape features to reach 94% accuracy of face shape classification. Finally, we demonstrate an on-the-fly prototype application for the face-shape-dependent fashion photo recommendation system.**

I. INTRODUCTION

Pursuing the perception of beauty is a nature for human beings. Over the past centuries, artists and psychologists have been tried to explore the mystery of beauty [1]. Besides, fashion industry which constitutes an aesthetic economy has played an important role in global economy in decades [2]. Recently, due to the facility of the internet, international fashion economy is exploding around the world. The global apparel retail industry in 2011 reached a value of $1.1 trillion, growing to $1.3 trillion in 2016 [3].

Thanks to the development of internet shopping and fashion websites, more and more fashion photographs are uploaded for providing the consumers advice about how to enhance their apparel. However, it is difficult for consumers to retrieve the useful apparel examples from the enormous fashion photos. For example, the apparel which looks perfect for one person on a fashion photograph may not fit for another person who subscribes the photo. Nguyen et al [4] discovers how different modalities, i.e., face shape and dress collectively affect the human sense of beauty (or attractiveness). That is, the attractiveness of a specific fashion item, such as accessory and clothing is corresponded to a certain face shape. Nguyen et al [4] uses statistical tags from Amazon mechanical turkers to demonstrate the relationship between face shape and fashion items jointly. However, it needs a lot of effort and is time-consuming to tag the face shape and dress to derive the relationship. In this paper, we proposed a novel framework to recommend the fashion photos according to the consumers' face shape automatically. Therefore, the consumers can derive their apparel recommendation with similar face shape from the fashion photo database. Fig. 1 illustrates the proposed framework for fashion photograph retrieval based on similar face shape matching techniques. The main contributions of this work can be summarized as follows.

1. We propose a novel framework where we can recommend fashion photographs based on the robust face shape features of the user, which is time-efficient for the users to derive useful apparel suggestions.
2. We collect a 2D face shape dataset with 140 different celebrities for training the eight face shape classification model. The eight face shape categories are commonly adopted in fashion applications [5].
3. We construct a fashion dataset from top fashion website Chictopia.com with 1105 images and 788 images for women and men respectively. Each image is tagged with aesthetic votes from the website peers. The proposed fashion photograph recommendation is made according to the aesthetic votes after similar face shape matching.

This paper is organized as follows. The proposed method is described in Sec. II. In Sec. III, experimental results are discussed. The conclusion is summarized in Sec. IV.

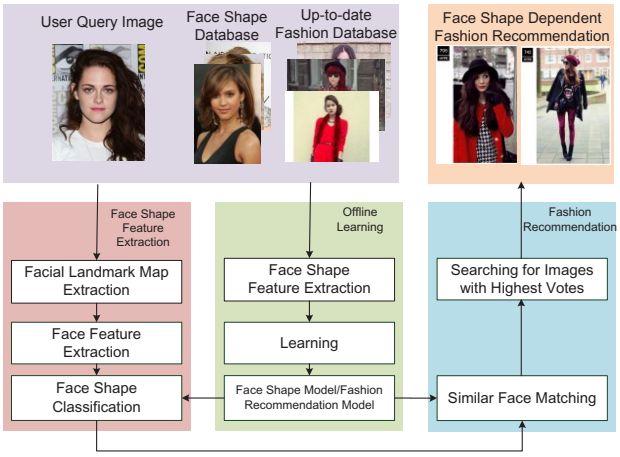

Fig. 1 System Block diagram

II. PROPOSED METHOD

The algorithm flow is shown in Fig.1. The proposed fashion photograph recommendation system is separated into two parts, off-line learning and on-the-fly best fashion recommendation. The off-line learning contains face shape feature extraction, learning and model construction of face shape and fashion recommendation. The face shape model is built based on the relationship between image database of human face shape and the corresponding face shape labels. We construct the learning models of different face shape types based on the extracted feature vectors. After constructing the learning models, the best fashion recommendation function is enabled. Feature vectors of the user's face are processed through the procedure of the feature extraction. The following face shape metric computing unit is adopted to calculate the feature distance between the feature sets of query image and up-to-date fashion database. After that, we sort the distance metrics of face shape features to find the most similar face shapes. Finally, we retrieve the face shape dependent fashion images with the highest aesthetic votes.

A. Facial Landmark Map Extraction

Instead of using low-level features, like Gabor or LBP, to compute the face shape features [6], we compute mid-level face shape features based on the facial landmark map result. In this paper, we adopt the robust face landmark extraction method to support uncontrolled face conditions in the real-world [7]. Zhu et al proposed the landmark extraction method based on mixtures of trees to capture topological changes due to viewpoint. In this way, the face landmark extraction performs well under the environment of variant face viewpoints and sizes.

After facial landmark map extraction, face is then transformed into a set of landmarks that represent distinguishable points such as jaw, eye pupils, lips etc. With the facial landmark extraction algorithm, total 68 landmarks are identified in a certain order, as shown in Fig. 2. We use landmark positions p52 – p60 and p61-p68 to enclose the face shape.

B. Robust Shape Feature Extraction

We calculate 13 mid-level face shape features based on the 68 landmark positions. By observing the face shape dataset in Fig. 3, we can find several face shape characteristics to classify 8 face shape categories visually. For example, the shape and width of jaw are usually represented

978-1-4799-1291-9/14/$31.00 ©2014 IEEE

for heart, round, rectangle, inverted triangle and pear shape. Besides, long face, oval and round face are commonly distinguished by the face length ratio. After the qualitative analysis, we implement the jaw shape characteristic by maximum transition angles, average slope, jaw max angle location and width ratio. Besides, we also utilize the ratio of face length and width to distinguish long, oval and round face shapes. The length ratio can be obtained from the ratio of face length and width as in Eq.1. Similarly, maximum transition angle (MTA) formula is described in Eq.2 where MTA is computed using adjacent vectors p_i , p_{i+1} and $p_{i-1}$of face shape landmarks.

$$\text{Length\_ratio} = \frac{||\frac{p_1+p_4}{2} - p_{52}||}{||p_{60} - p_{68}||} \quad (1)$$

$$\text{MTA} = \max_{i \in face\ shape} \text{acos}((p_{i-1} - p_i) * (p_{i+1} - p_i)) \quad (2)$$

C. Face Shape Classification

Face shape classification result are predicted from the face shape learning model. In this paper, LIBSVM with Gaussian kernel is adopted for learning model [8]. We pre-train the SVM model using the eight-category face shape database. We optimize SVM model of each category by regression of the LIBSVM cost and gamma parameters. Thus the most accurate model is derived after parameter regression.

D. Similar Face Matching

We generate the similar face matching by calculating and sorting L1 distance metric using Eq.3. And then we retrieve the best fashion photos according to the smallest face feature distance and highest aesthetic votes.

$$d_{L1} = \sum_{i=1}^{13} |Query\_image\_feature_i - Database\_image\_feature_i| \quad (3)$$

III. EXPERIMENTAL RESULTS AND ANALYSIS

Two experiments are conducted to evaluate classification accuracy and computation time of the proposed system. We collected two dataset for training the face shape model and fashion photograph recommendation model. The eight face shape dataset in Fig.3 consists of 140 celebrity photos collected from fashion website with a resolution of 500*500. And the proposed fashion recommendation dataset is constructed by with a resolution of 400*600. To ascertain our classification accuracy and compare with previous Gabor-LBP feature based method [6], we adopt the large scale LFW dataset [6]–13,233 face images with 3 face shape attribute classifier outputs. The experiments are performed on a PC with Intel Core I7 3.4G CPU and 4GB DDR RAM equipped.

Table I shows the system accuracy difference between the proposed face feature extraction method and one used in the previous study method [6]. With the proposed feature extraction mechanism, the classification accuracy is improved to 94.00% from 75.23%. The test image set is composed of 1000 images chosen from the LFW dataset with 500 highest score and 500 lowest score of face shape attribute value as positive and negative data, respectively. Eight face shape dataset is shown in Fig. 3. With the proposed SVM parameter optimization, classification accuracy of eight face shape dataset is further improved to 84.628% (stdev (standard deviation) = 1.59) from 81.11% (stdev = 3.48).

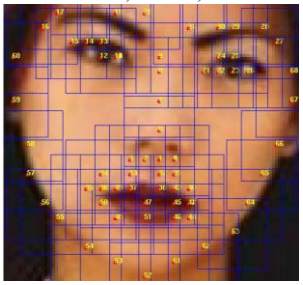

Fig. 2 Facial Landmark Map

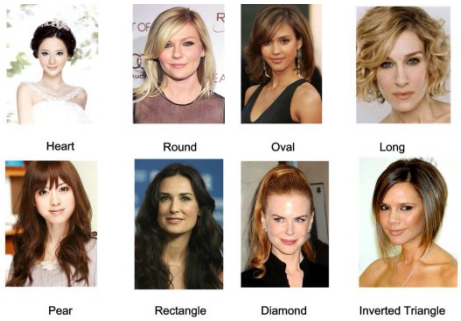

Fig. 3 Eight Face Shape Dataset.

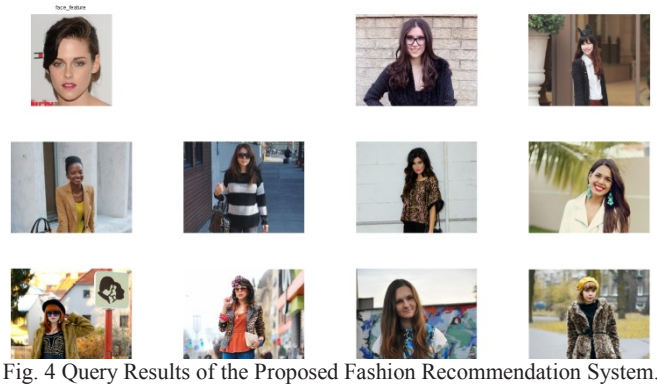

Fig. 4 Query Results of the Proposed Fashion Recommendation System.

TABLE I LFW DATASET ACCURACY COMPARISON

Face Shape Category	ICCV'09 [6] (%)	This Work (%)
Square	81.19	93.68
Round	74.33	93.67
Oval	70.16	94.67

The second experiment evaluates the computation time of the proposed system. The facial landmark map generation consumes 2.07s and feature extraction consumes 0.02s. Feature matching and fashion photograph retrieval consume 0.94s jointly. This experimental result demonstrates that our system provides on-the-fly fashion photograph retrieval while maintaining high retrieval quality as shown in Fig. 4.

IV. CONCLUSION

We present a robust face shape feature extraction method targeted at fashion photograph recommendation applications. The proposed face shape classification accelerates the apparel retrieval procedure. We adopted 13 mid-level face shape features to enhance the accuracy of face shape classification. Besides, regression of SVM model is utilized to achieve accuracy optimization. Results show the identical accuracy under different face shape categories and report decent performance.

REFERENCES

[1] E. Berscheid et al, "What is beautiful is good," in *Journal of Applied Social Psychology*, 1972.

[2] Joanne Entwistle, "The aesthetic economy: The production of value in the field of fashion modelling," in *Journal of Consumer Culture*, 2002.

[3] The Neverblue Fashion Performance Network, "International fashion statistics," in *http://www.fashionperformancenetwork.com/*, 2011.

[4] Bingbing Ni et al, "Sense beauty via face, dressing, and/or voice," in *Proc. of the 20th ACMMM*. IEEE, 2012, pp. 239–248.

[5] Yanqiu Xu et al, "Measurement and classification of shanghai female face shape based on 3d image feature," in *IEEE CISP*. 2010.

[6] N. Kumar et al, "Attribute and Simile Classifiers for Face Verification," in IEEE ICCV, Oct, 2009.

[7] X. Zhu et al, "Face detection, pose estimation, and landmark localization in the wild," in IEEE *CVPR.*, 2012.

[8] C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2011.