

A Survey of Major Techniques of Clothing in Vision^{*}

Dongyan Huang^a, Li Liu^{a,b,*}, Xiaodong Fu^a, Lijun Liu^a, Qingsong Huang^a

^a*Yunnan Provincial Key Laboratory of Computer Technology Application, Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China*

^b*National Engineering Research Center of Digital Life, State-Province Joint Laboratory of Digital Home Interactive Applications, School of Data and Computer Science, Sun Yat-sen University Guangzhou 510006, China*

Abstract

Clothing in vision plays an important role in many applications of digital clothing, virtual fitting, fashion design and manufacturing. The research of clothing in vision is full of technical challenges, such as background detection and features extraction. In this paper, we introduce the latest remarkable works of clothing in vision area, including clothing segmentation, clothing recognition and clothing retrieval. There are many remarkable advances on the basis of the predecessors' achievements in the three fields. Firstly, we provide an overview of the techniques for clothing segmentation and clothing recognition. Then, we introduce and analyze the representative works on clothing retrieval. Finally, we forecast broader application prospect and future research work of clothing in vision.

Keywords: Clothing in Vision; Clothing Segmentation; Clothing Recognition; Clothing Retrieval

1 Introduction

With a rapidly increasing use of the phones and computers, there is much digital image information, among which lots of picture information is applied to the commercial website. As the inevitable trend, retrieval has moved beyond text to other modalities like images. Currently, there are thousands of e-commerce platforms in the website and lots of people around the world would like to buy clothes on the web. Factually, Chinese total retail sales in the clothing market have reached 15.12 billion dollars in the first half year of 2014, and its apparel market is one of the “fastest growing markets in the world”. Driven by the huge profit potential in the fashion industry, intelligent fashion analysis based on techniques for clothing is receiving a great deal of attention in the multimedia and computer vision literature.

^{*}This research is supported by the National Natural Science Foundation of China (Nos. 61462051, 61462056, 81360230, 71161015), the Applied Fundamental Research Project of Yunnan Province (No. 2014FB133), the Applied Fundamental Research Key Project of Yunnan Province (No. 2014FA028), the Science Research Funded Project of Kunming University of Science and Technology (No. KKSJ201403119), and the Open Research Program Key Lab of Computer Technologies Application of Yunnan Province.

^{*}Corresponding author.

Email address: kmust_mary@163.com (Li Liu).

There were several review articles about clothing in vision, but these articles all placed extra emphasis on the previous works. They lack the introduction of the latest remarkable developments. The main idea of this paper is aimed at the survey of major techniques of clothing in vision recently, including clothing segmentation, clothing recognition, and clothing retrieval. We analyze the worldwide outstanding works and summarize the representative ones of them.

The rest of this paper is organized as follows. Section 2 gives fundamental knowledge for clothing segmentation techniques. Section 3 introduces the principles and details of several major techniques on clothing recognition. Section 4 presents the comparisons of the latest clothing retrieval system. In Section 5 we conclude and discuss the future work.

2 Clothing Segmentation

There were many segmentation methods on the basis of the facial detection [1] or skin detection [2]. We expatiate several representative works with novel methods in feature extraction and clothing classification.

Weber et al. [3] proposed a segmentation approach which was able to separately segment a person's upper and lower clothing regions, taking the person's pose into account. They focus on person retrieval in multimedia data, specifically in TV series. They also combine the individual masks to obtain a full segmentation of the person's upper and lower clothing. The approach can deal with unconstrained poses and occlusions of persons. However, the limitation of the masks' quality influences the segmentation performance, which needs to be further improved. The true positive rate of the segmentation is not satisfied. It may be increased by taking facial features into account. Furthermore, it is not adaptable to the whole-length clothing.

Yang et al. [4] developed an integrated system for clothing co-parsing. They proposed a phase of inference, which was referred as "image co-segmentation". The phase iterates to extract consistent regions on images and jointly refines the regions over all images by employing the exemplar-SVM technique. One contribution of this work is that they set up a high-resolution street fashion photos dataset with annotations. The method can deal with more complex backgrounds and clothing styles than some state-of-art methods. Moreover, it can suppress the image clutters and generate coherent regions effectively. However, it would fail in the following scenarios: the background is extremely disordered, different clothing garment items share similar appearance, ambiguous patterns exist within a clogging garment item, and the illumination condition is poor.

Freire-Obregon et al. [5] introduced a new clothing segmentation method based on the application of the GrabCut technique over a trixel ("Triangular Superpixels") mesh. They adopted the trixel to simplify the image data into perceptually meaningful atomic regions. And they proposed a new method to generate trixels. They used the distance maps from unthresholded magnitudes and the delaunay triangulation to create the trixels topological map (TriToM). The TriToM was used as the input graph of the GrabCut algorithm. The clothing segmentation method makes use of the face, the local head context and the clothing information of the upper torso to improve the performance. And it is more fast and efficient than the traditional GrabCut by using the trixels. Nevertheless, the method would be invalid to the image without facial and head context information. And it can not be used for generic clothing segmentation. Table 1 shows the advantages