ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 20, April 2016

A Preliminary Survey on Recognizing Clothing Styles of People for Intelligent Image Understanding

Devendran V
Associate Professor,
Dept of Computer Applications
PSNA College of Engg & Tech,
Dindigul, Tamilnadu, India
svdevendran@gmail.com

Nagarajan A
Assistant Professor
Dept of Computer Science and Engineering
Alagappa University
Karaikudi, Tamilnadu, India
nagarajan_mca2003@yahoo.co.in

Abstract: Intelligent image understanding have been an important research for a decade, as it has itself a long list of applications including robotic vision. This paper highlights the importance of human clothing as the one of the key to the intelligent image understanding. Clothing styles makes a human to be viewed differently in shapes and in their contour. We could able to derive many attributes from the styles of clothing from the humans, like their status, their occupation, nativity, culture and etc. Also, we could extract useful information from the people those who are in uniforms such as Departments they belong, cadre and ranks. In this line of research, now-a-days, clothing recognition is getting the attention of our researchers. So, this paper does a preliminary research, recalls most of the interesting researches of clothing styles of human and proposes future scope in this direction.

Keywords: Intelligent Image Understanding, Clothing Recognition,

I. INTRODUCTION

"Looking at people" is currently one of the most active domains of research in computer vision. People detection is a particularly challenging image understanding problem. As a result of variations in clothing and posture, the appearance of people may vary enormously from image to image. Identifying the people in photographs wearing same kind of dresses is the primary objective of work irrespective of the kind of the dresses people are wearing. Dress codes are very important key to understand the common group of people from unknown images such as Doctors, Civil Policemen, Traffic Policemen, Nurses, Dance Performers, Military and Navy People, Air hosters, Waiters in star hotels, People in Jail, People attending political meetings, People participating in our Republic Day Parade, People of Rajasthan, People of Kerala, People of Tamilnadu and etc., are the best examples the way how we recognize the common people in images as shown in Fig. 1.

This paper is organized as follows. Section II gives the detailed literature survey. Section III shows the related works and applications. Section IV depicts the future scope for research. Section V concludes with a conclusion.

II. LITERATURE SURVEY

Initially, a paper [2] demonstrates a content-based retrieval strategy that can tell whether there are naked people present in an image or not. The approach combines color and texture properties to obtain an effective mask for skin regions. The scope of the paper [1] is limited to work on whole-body or hand motion; it does not include work on human faces The ability to recognize humans and their activities by vision is key for a machine to interact intelligently and effortlessly with a human-inhabited environment. In general, the processing framework [5] of visual surveillance in dynamic scenes includes the human identification from multiple cameras. In paper [7], authors present a concept of matching clothes detection is attempted, which is a challenging task for many blind people. This system can handle clothes in deficient color without any pattern, as well as clothing with multiple colors and complex patterns to aid both blind and color deficient people. In paper [6], authors investigate the human occupation prediction problem by modeling the appearances of human clothing as well as surrounding context. The human clothing, regarding its complex details and variant appearances, is described via part-based modeling on the automatically aligned patches of human body parts. The Paper [9] mainly target the cloth simulation application where cloth (highly deformable)

All Rights Reserved © 2016 IJARTET

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 20, April 2016

collides with deforming skin of a moving human model. Authors introduced an inclusion-field technique for fast detection of collisions between a highly deformable object and another object with limited deformations. Most of existing techniques to address the problem of summarizing a video sequences. This has been the goal of a quickly evolving research area known as video summarization [8]. Recent advances in technology have increased the availability of video data, creating a strong requirement for efficient systems to manage those materials. The paper[10] devised a multimedia event detection method from a large set of consumer-generated web videos taken in unconstrained environments. The proposed deformation model [9] provides intuitive control over the three parameters independently, while producing aesthetically pleasing deformations of both the garment and the human body. It presents a framework for learning a three-layered model of human shape, pose and garment deformation. In paper[15], authors tackle the clothing-parsing problem using a retrieval-based approach. The approach combines parsing from: pre-trained global clothing models, local clothing models learned on the fly from retrieved examples, and transferred parse-masks from retrieved examples. In paper[14], authors discussed 'Person re-identification problem across non-overlapping camera views is a rather challenging task. Due to the difficulties in obtaining identifiable faces, clothing appearance becomes the main cue for identification purposes. So, authors present a comprehensive study on clothing attributes assisted person re-identification. Paper [11] describes an easy-to-use system to estimate the shape of a human body and his/her clothes. Hence, clothing recognition has been given the high importance since long.

III. RELATED WORK AND APPLICATIONS

A. Clothing retrieval

There is a growing interest in clothing recognition problems, perhaps due to the huge potential for impact on e- commerce clothing applications (annual revenue for on- line shopping totals over $200 Billion dollars annually). Automatic clothing recognition methods could enable natural and semantic image search for users of online fashion shops. This is reflected in the increasing number of papers related to clothing recognition for retrieval or recommendation applications. Most notably, the work of Liu et al. [16] proposes a street-to-shop application, which tries to match pictures of clothing taken in the real world to clothing images in online shopping sites. In their approach, the authors consider a mapping between street and shopping images with a sparsely coded transfer matrix so that the difference between these two distributions does not affect the quality of retrieval. Kalantidis et al. [17] take a similar cross-scenario retrieval approach, where they utilize clothing parsing to explicitly represent each item and look at a similar problem, but with a focus on efficiency for mobile scenarios. In addition to applications directly focused on clothing retrieval, clothing appearance similarity has been used for applications whose goal is to find the same person in similar poses across image or video collections [14]. As the interest in clothing-related applications grows, alongside these projects, there have been concerted efforts to create fashion-related datasets [15].

B. Attribute recognition

Attributes of clothing are a natural way to describe its visual characteristics. For example, a user of an online shopping site might be looking for a "blue striped shirt" or a "red spectator pump". In general attributes relate to visual characteristics of objects such as color, pattern, or shape. Attributes for clothing have been explored in several recent papers [6][14]. In general the attribute analysis is built upon detection and localization of items or parts of items in a picture. The clothing attribute recognition focused on recognizing clothing composites on upper-body detections. Since attributes usually do not exist in isolation, Chen et al. considers co-occurrence between attributes during prediction using conditional random fields. Some work has been done in this area, using fine-grained attribute detection or using human-in-the-loop approaches to interactively whittle down search results to what the user is looking for or to build user specific models during search.

C. Clothing and person identification

Another important application of clothing recognition is the identification of people by their clothing. For example in temporal image collections, e.g. many pictures from an event, a person will be wearing the same clothes throughout the event. Therefore, clothing recognition is a strong cue to identity and has been used to recognize people in personal photo collections, repeated shots, or in surveillance scenarios [3]. The clothing we wear is also a strong cue for predicting other characteristics about ourselves, such as social status, occupation, wealth, or occasion, to name a few. In this direction, there has been recent work on clothing recognition for occupation recognition [6], fashion style recognition [15], or social tribe prediction in group-photos. In the inverse direction, Liu et al. propose a system to recommend outfits (sets of clothing items) according to the occasion or event [9].

All Rights Reserved © 2016 IJARTET

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 20, April 2016

D. Clothing parsing

Clothing parsing is a relatively new computer vision task, but one that is important for enabling the above applications and for developing useful clothing representations. Early work on clothing representation modeled clothing as a grammar of sketch templates. Other work took a subspace approach to describe clothing deformations, or deformable spatial priors. These approaches focus mainly on how to model shape deformations for clothing recognition. The clothing-parsing problem was first formulated as an MAP estimation of super pixel labels using conditional random fields. The main insight of this method was the use of body pose estimation for clothing parsing. Dong et al. [18] later propose clothing parsing as an inference problem over parselets, a basis group of image regions that constitute clothing items.









All Rights Reserved © 2016 IJARTET

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 20, April 2016

Fig 1. Randomly Downloaded Images from Google

Liu et al. also propose a method to eliminate pixel-level supervision in learning using image-level color tags.

E. Semantic segmentation

Clothing parsing is directly related to the well-studied image-parsing problem, where the goal is to assign a semantic object label to each pixel in an image. Most related to our paper are non-parametric methods for semantic segmentation, which have demonstrated state-of-the-art performance on the image-parsing problem. Our approach shares the same non-parametric design for clothing parsing, but can additionally take advantage of pose estimates during parsing, and we do so in all parts.

F. Pose estimation

Effective clothing parsing strongly depends on accurate human body pose localization. Therefore, we take advantage of recent advances in pose estimation. Some previous clothing recognition work has used face detection to first find a person's head and torso and use this to bootstrap localization of given clothing items. Also some approaches to the pose estimation problem itself have taken advantage of image segmentation for improving performance. In this paper, we show empirical evidence that semantic clothing segmentation is beneficial to improving pose estimation.

G. Style Finder: Recognition and Retrieval

With the rapid proliferation of smartphones and tablet computers, search has moved beyond text to other modalities like images and voice. For many applications like Fashion, visual search offers a compelling interface that can capture stylistic visual elements beyond color and pattern that cannot be as easily described using text. However, extracting and matching such attributes remains an extremely challenging task due to high variability and deformability of clothing items. Here, the paper gives a fine-grained learning model and multimedia retrieval framework to address this problem. First, an attribute vocabulary is constructed using human annotations obtained on a novel fine-grained clothing dataset. This vocabulary is then used to train a fine-grained visual recognition system for clothing styles. Authors report benchmark recognition and retrieval results on Women's Fashion Coat Dataset and illustrate potential mobile applications for attribute-based multimedia retrieval of clothing items and image annotation.

H. Event Detection

Most studies for event detection have been aimed at identifying events in professionally produced videos such as sports and movies, or in surveillance videos [3]. These studies used event-specific methods, which rely heavily on the spatial-temporal structures of the target events. By Assfalg et al. [19], for example, three types of highlights in soccer games such as penalty kick were modelled by three-state HMMs, using constant camera motions and the location of players as features. Li et al. [20] detected dialog events in movie videos by shot clustering and audio analysis. Adam et al. [21] modelled the stream of people in surveillance videos from optical flows to detect unusual events such as running in the mall. While these methods can be applied when the target events are specified and the spatial-temporal structures of the events are always stable, it is difficult to apply them to general events appearing in consumer-generated videos due to two major reasons. First is that, general events widely vary and the definitions of them are not always clear. Second is that consumer-generated videos do not have stable spatial-temporal characteristics since they are taken by amateurs from different points of view and often include unsettled camera motions or haphazard editing.

All Rights Reserved © 2016 IJARTET

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 20, April 2016

I. Predicting Occupation of Human

Authors propose to recognize human occupation via two properties of human-related images, human clothing and scene contexts. Predicting human occupations in photos has great application potentials in intelligent services and systems. How- ever, using traditional classification methods cannot reliably distinguish different occupations due to the complex relations between occupations and the low-level image features. In this paper, we investigate the human occupation prediction problem by modeling the appearances of human clothing as well as surrounding context. The human clothing, regarding its complex details and variant appearances, is described via part-based modeling on the automatically aligned patches of human body parts. The image patches are represented with semantic-level patterns such as clothes and haircut styles using methods based on sparse coding towards informative and noise-tolerant capacities. This description of human clothing is proved to be more effective than traditional methods. Different kinds of surrounding context are also investigated as a complementarity of human clothing features in the cases that the background in- formation is available. The preliminary study shows the human occupation is reason- ably predictable using the proposed clothing features and possible context.

J. Clothes Simulation On Virtual Actors

Clothes are animated as any kind of deformable objects. They differ only by how they are constructed, that is, by assembling 2D panels. The clothing design and simulation process is divided in to two parts: Firstly, the 2D panel design process, which uses 2D drawing software to design garment models as flat fabric panels. Seams are defined around the borders of the panels. Secondly, the 3D simulation process, which uses mechanical simulation to assemble the garment panels in the context where the animation will take place and then simulates the animated scene and characters. The scene may contain several objects, static or animated, that will interact with the garments through collision. In particular, the actor to be dressed is typically an animated object. Finally, we record the animated garments frame by frame as animations that can be used as input data for subsequent computations. This allows incremental garment design for complex clothing.

IV. FUTURE SCOPE FOR RESEARCH

Worldwide, the problem of Clothing Identification is being given the high importance in the sense it is having huge scope of research in variety of application as follows.

- Visual surveillances
- Activity recognition
- Tracking of Human
- Looking-at-people in Specific
- Culture recognition
- Intelligent image recognition
- Intelligent searching through dress codes and etc.

Let the future research will be solving many issues to make it automatic.

V. CONCLUSION

This paper describes the importance of clothing recognition and tries to show that it is the vital part in the direction of an intelligent image understanding. So far, the authors concentrated much on object detection, segmentation, object recognition and so on, by looking at things individually in different dimensions in images and videos. This paper gives summary of research papers and their scope of the research done so far on clothing styles. Clothing styles makes a human to be viewed differently in shapes and their contour. We could able to derive many attributes from the styles of clothing from the humans, like their status, their occupation, nativity, culture and etc. Also, we could extract useful information from the people those are in uniforms such as departments they belong, cadre and ranks. In this line of research, this paper takes an initiative with the preliminary study on clothing recognition, which is getting the attention of our researchers towards intelligent image understanding.

REFERENCES

[1.] M.P. Gascuel et al., “A Modeling System for Complex Deformable Bodies Suited to Animation and Collision Processing,” J.*Visualization and Computer Animation,* Vol. *2,* 1991.

[2.] M. Fleck, D. Forsyth, and C. Bregler. “Finding naked people”, In Proc. of ECCV, Cambridge, England, April 1996.

[3.] D.M.Gavrila, “The Visual Analysis of Human Movement: A Survey”, Computer Vision and Image Understanding, Vol. 73, No. 1, pp. 82-98, 1999.

[4.] Liang Zhao, “Dressed Human Modeling, Detection, and Parts Localization”, Doctorate Thesis, The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, July 26, 2001.

[5.] Weiming Hu, Tieniu Tan, Fellow, IEEE, Liang Wang, and Steve Maybank, “A Survey on Visual Surveillance of Object Motion and Behaviors“, IEEE Transactions On Systems, Man, And Cybernetics - Part C: Applications And Reviews, Vol. 34, No. 3, August 2004.

[6.] Zheng Song; Meng Wang; Xian-Sheng Hua; Shuicheng Yan, "Predicting occupation via human clothing and contexts", IEEE International Conference

All Rights Reserved © 2016 IJARTET

on Computer Vision (ICCV), pp.1084 - 1091, Nov. 2011.

[7.] Shuai Yuan, Ying Li Tian, and Aries Arditi, "Clothing Matching for Visually Impaired Persons", Technology and Disability, Vol. 23, No. 2, pp. 75-85, 2011.

[8.] Jurandy Almeida, Neucimar J. Leite, Ricardo da S. Torres, "VISON: VIdeo Summarization for ONline applications", Pattern Recognition Letters, Vol. 33, No. 4, pp. 397–409, 2012.

[9.] Asma A. ElBadrawy, Elsayed E. Hemayed, Magda B. Fayek, "Rapid collision detection for deformable objects using inclusion-fields applied to cloth simulation", Journal of Advanced Research, Vol. 3, No. 3, pp. 245–252, 2012.

[10.] Yusuke Kamishima, Nakamasa Inoue and Koichi Shinoda, "Event detection in consumer videos using GMM supervectors and SVMs", EURASIP Journal on Image and Video Processing, SpringerOpenAccess, pp. 1-13, 2013.

[11.] Neophytou A, Hilton A, "A Layered Model of Human Body and Garment Deformation", 2nd International Conference on 3D Vision (3DV), vol.1, pp.171 - 178, Dec. 8-11, 2014.

[12.] Fahim A., Mutahira N. Naseem, "Facebook Spam and Spam Filter Using Artificial Neural Networks", International Scholarly and Scientific Research & Innovation, Vol. 9, No. 1, pp. 220 – 223, 2015.

[13.] Ming Zeng, Liujuan Cao, Huailin Dong, Kunhui Lin, Meihong Wang, Jing Tong, "Estimation of human body shape and cloth field in front of a kinect", Neuro computing, Vol. 151, No. 2, pp. 626–631, 2015.

[14.] Annan Li, Member IEEE, Luoqi Liu, Kang Wang, Si Liu, and Shuicheng Yan, Senior Member IEEE, "Clothing Attributes Assisted Person Reidentification", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 25, No. 5, pp. 869 - 878, May 2015.

[15.] Kota Yamaguchi, Member IEEE, M. HadiKiapour, Student Member, IEEE, Luis E. Ortiz, and Tamara L. Berg, Member, IEEE, "Retrieving Similar Styles to Parse Clothing", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 37, No. 5, pp. 1028 – 1040, 2015.

[16.] S. Liu, Z. Song, G. Liu, C. Xu, H. Lu and S. Yan, "Street-to-shop: Cross-scenario clothing retrieval via parts alignment and auxiliary set," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, pp. 3330-3337, 2012.

[17.] Yannis Kalantidis, Lyndon Kennedy, Li-Jia Li, "Getting the look: clothing recognition and segmentation for automatic product suggestions in everyday photos", Proceedings of the 3rd ACM conference on International conference on multimedia retrieval (ICMR '13), pp. 105-112, 2013.

[18.] Jian Dong, Qiang Chen, Wei Xia, Zhongyang Huang, Shuicheng Yan, "A Deformable Mixture Parsing Model with Parselets", The IEEE International Conference on Computer Vision (ICCV), pp. 3408-3415, 2013.

[19.] J Assfalg, M Bertini, A Del Bimbo, W Nunziati, P Pala, "Soccer highlights detection and recognition using HMMs", in IEEE International Conference on Multimedia and Expo, (IEEE, Lausanne, 26–29 August 2002), pp. 825–828, 2002.

[20.] Ying Li, Shrikanth Narayanan, C.-C. Jay Kuo, "Content-Based Movie Analysis and Indexing Based on AudioVisual Cues", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 14, No. 8, August 2004.

[21.] Adam A, Rivlin E, Shimshoni I, Reinitz D, "Robust real-time unusual event detection using multiple fixed-location monitors", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 30(3), pp. 555-560, 2008.

All Rights Reserved © 2016 IJARTET