Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2015, 7(3):575-578

Review Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Research on facial features detection using RGBD camera

Qingxiang Wang

School of Information, Qilu University of Technology, Jinan, China

---

ABSTRACT

Facial features detection, which is a very active research direction in computer vision. It is widely used in many computer vision applications, such as animation, face recognition, expression analysis and transfer. RGBD camera is a new input device which could provide depth and RGB image. Some new methods were proposed based on the RGBD camera. In this paper, we review the works of facial features detection using RGBD camera in recent years.

Key words: Facial Features Detection; RGBD; Depth

INTRODUCTION

Facial features detection, which is a very active research direction in computer vision. It is widely used in many computer vision applications, such as animation, face recognition, expression analysis and transfer. It also provides input for many applications.

Some famous works on facial features detection are presented. Active appearance models and their improvements [1-3], Active Shape Models [4], 3DMM [5] and other detect method [6]. The methods are based on the 2D image and work well on the test images, but they are still suffered on the drawbacks of complex backgrounds, illumination changes and head pose changes.

Facial features detection in the wild under conditions of complex backgrounds, illumination changes and pose changes is difficult if only relying on 2D image. In recent years, more and more studies tend to use the 3D model and depth information. RGBD camera, like kinect, can provide depth information and RGB image information at the same time, which provides a new tool for better facial feature detection in the wild.

In this paper, we proposed a simple review of the newest methods of the facial features detection using RGBD camera. Most of them are based on Kinect. The Microsoft's Kinect SDK also provides a tool for facial features detection from 1.5 which is not include in this paper.

RGBD CAMERA

RGBD camera is a camera with RGB and depth which can provide an extra dimension for image process. The depth part is a kind of distance measuring equipment. It can get the distance from the object to the camera that makes the 3D information acquirement of the scene reality. The common methods include flight time (Time of Flight, abbreviated TOF) technology, the structure light, Binocular Stereo Vision, etc., and most of them use infrared. Some of them are Mesa Imaging SwissRanger SR4000/SR4500、PMD Technologies CamCube 2.0, Primesence Sensor, WAVI Xtion, LeapMotion and Kinect.

Since it was born, the depth camera has acquired a large number of applications, such as 3D reconstruction, orientation analysis, user interface and tracking and so on. In all of the cameras, kinect has got the most attention. Kinect 1.0 uses infrared speckle projection on the different distance to measure the depth of change and Kinect one (2.0) uses the structure light. Because of the low price and good robustness, it is widely used on animation [7-8], calibration, gesture recognition [9], human-computer interaction, visual and navigation, virtual reality, graphics, image processing, 3D reconstruction of scene and character, virtual fitting system, game system, etc.

FACIAL FEATURES DETECTION USING RGBD CAMERA

The facial features detection using RGBD camera can be divided into two methods, one directly extract the features from the depth map and the other combines the depth map and the RGB map to acquire the features.

1. Feature detection from the depth

Due to the use of infrared signal, the influence of illumination is very small. So it is well to process complex and poor lighting conditions. The depth of the image contains facial space information which can be directly used to extract the feature point's locations.

1.1 Deformation model matching method

Such methods use the existing 3D deformable model to match the depth of image and find the closest deformation coefficient the transformation matrix to the depth image with and then get the location of the feature points.

Qin Cai and Zhengyou Zhang presented a regularized maximum likelihood deformable model fitting (DMF) algorithm [10]. They proposed a likelihood estimation method based on iterative closest point algorithm, which find out the corresponding relationship between the depth map and the deformation model, work out the rotation and translation vector, and then use it to initialize and solve the model of deformation, get the coefficient and detect the feature points. The equation they used is

$$R(p_k + A_k r + b_k s) + t = g_k + x_k \tag{1}$$

Where p_k is the basic head models, A_k and B_k are the coefficient of vertex k of the constructed 3D facial models, A_k is the motion corresponds to the facial movements which is similar to FACS, B_k is the stable facial construction corresponding to the facial pattern, like eyebrow spacing motion corresponds to the facial movements, t is the translation, g_k is the depth map, x_k is the noise. See in Fig 1.

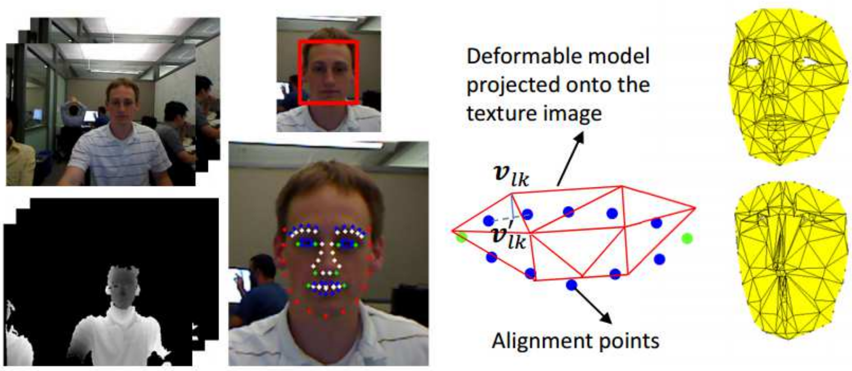

Fig. 1 DMF method [10]

M, Breidt[11] use 3DMM combined with principal component analysis method (3DMM-PCA) to match the TOF camera data and the facial motion unit model of the high precision scanner data, which converts the depth image of the low precision to facial model of high precision.

The state of the art methods use the FACS to define the 3D face models. Weise build the base facial expression library and personalized specific facial expression templates, analyze the depth flow to get the 3D head information and decompose it into the combination of the base facial expressions and then use the coefficient of combination to the 3D animation [12]. Chen Cao establishes the database of 3D facial expressions with kinect [13] and improves the Weise's method by regression to deal with 2D video [14] and without a user-specific training [15].

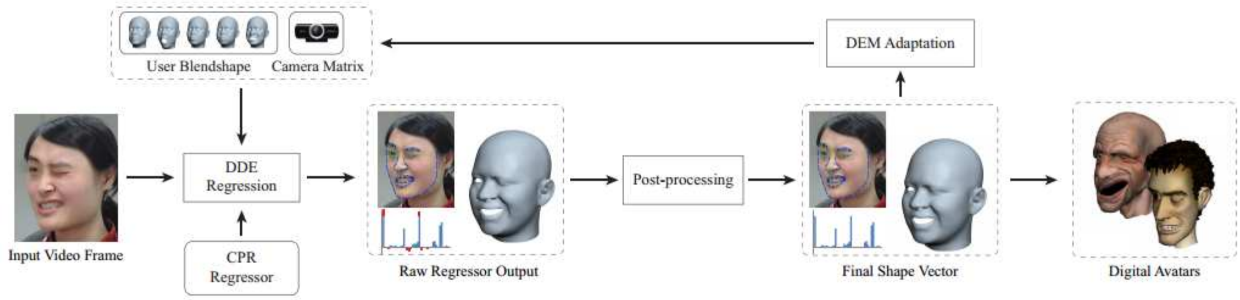

Fig. 2 Displaced dynamic expression regression method [15]

1.2 Statistics method

Such methods used a large number of examples, the depth image was defined for each part (such as a rectangular block) in the descriptor, adopt the unified framework or operator to his description, using statistical methods to classify, and then using the probability method to determine the new instance. Gabriele Fanelli presented random forests method to determine RGBD camera head orientation [8], and on this basis, put forward in the 3D face to determine the feature points location methods [16].

2. Feature detection from the combination of depth and RGB

Tadas Baltrusaitis etc. proposed a CLM - Z method [17] which used the RGB and depth of local template at the same time. They used support vector machine (SVM) to deal with local template image block, and the improved scheme matching shift method. The method could get good results under the bad light condition.

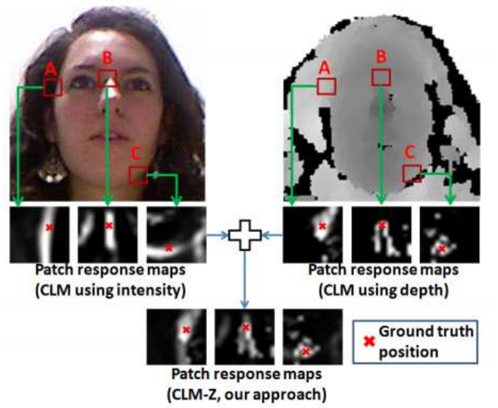

Fig.3 CLM-Z method [17]

Fei Yang, etc. also used RGB image and depth image at the same time [18]. They initialized the ASM algorithm which only use RGB image before with the depth image, and constrained the iteration process with depth that make it have higher accuracy and speed. Colin Bellmore used RGB image and depth to optimize ASM [19], and used that for face expression recognition.

CONCLUSION

The recent typical works from 2010 for facial features detection are mentioned in this review. From the works we can divide those into two orientations, one only use depth information and the other use the depth and RGB at the same time. The first type methods can work in the head pose change and poor light conditions and most of them are based on 3D deformable or blend model. But lack of the accuracy of the kinect, they still could not deal with the face detail such as wrinkle. FACS is effective method for blend model and is used in most of the works. The second type methods use the depth and RGB. They process the depth and RGB image separately and combine the result. In the future, the depth and RGB image may be used collaboratively.

Acknowledgments
This work was supported by Jinan star science and technology plan (No. 201406004 and No.20120104).

REFERENCES

[1] Cootes Timothy F., Edwards Gareth J., Taylor Christopher J. *European Conference on Computer Vision*, **1998**, 484-498.
[2] Xiao Jing, Baker Simon, Matthews Iainet al. *IEEE conference on Computer Vision and Pattern Recognition*, **2004**, 535-542.
[3] Georgios Tzimiropoulos, Maja Panti. *International Conference on Computer Vision*, **2013**, 593-600
[4] Cootes T., Cooper D., Taylor C.et al. *Computer Vision and Image Understanding*, **1995**, 61(1), 38-59.
[5] Blanz Volker, Vetter Thomas. *Annual Conference on Computer Graphics and Interactive Techniques*, **1999**, 187-194.
[6] Matthias Dantone, Juergen Gall, Gabriele Fanelli, Luc Van Gool. *IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, **2012**, 2578-2585.
[7] Sofien Bouaziz, Yangang Wang, Mark Pauly. *ACM SIGGRAPH*, **2013**, 1-10.
[8] Fanelli Gabriele, Weise Thibaut, Gall Juergenet al. *DAGM-Symposium*, **2011**, 101-110.
[9] Shotton Jamie, Sharp Toby, *Communications of the ACM*, **2013**, 56(1),116-124.
[10] Qin Cai, David Gallup, Cha Zhang et al. *European Conference on Computer Vision*, **2010**, 229-242.
[11] M Breidt, H Biilthoff H., C Curio. *IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG* **2011***)*, **2011**, 713-719.
[12] Weise Thibaut, Bouaziz Sofien, Li Hao, et al. *ACM Transactions on Graphics (Proceedings SIGGRAPH)*, **2011**, 30(4), 76-79.
[13] Chen Cao, Yanlin Weng, Shun Zhou, Yiying Tong, Zhou Kun. *IEEE Transactions on Visualization and Computer Graphics*, **2013**, 20(3), 413-425.
[14] Chen Cao, Yanlin Weng, Stephen Lin, Kun Zhou, *ACM Transactions on Graphics (Proceedings SIGGRAPH)* , **2013**, 32(4), 41
[15] Chen Cao, Qiming Hou, Kun Zhou. *ACM Transactions on Graphics (Proceedings SIGGRAPH)*, **2014**
[16] Fanelli Gabriele, Dantone Matthias, Gall Juergen. *International Journal of Computer Vision*, **2013**, 101(3), 437-458.
[17] Baltrusaitis T., Robinson P., Morency L. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, **2012**, 2610-2617.
[18] Yang Fei, Huang Junzhou, Yu Xianget al. *IEEE International Conference on Image Processing (ICIP* **2012***)*, **2012**, 561-564.
[19] Bellmore Colin. Facial Feature Point Fitting with Combined Color and Depth Information for Interactive Displays. Kate Gleason College of Engineering, **2012**.