ATLANTIS PRESS

A HMM-Based Structured Model for Human Behavior

Weihua Wang[1, a]

[1] Key Laboratory of Machine Vision and Intelligent Information System, Chongqing University of Arts and Sciences, Yongchuan, Chongqing, 402160, China

[a]y2002ww@163.com

Abstract. A structured recognition method of human behavior based on HMM is proposed to build the human action in this paper, only one camera which does not need to be adjust should be used in this approach. The robustness of the human behavior recognition model which is established by the free view angle of the camera is improved. The human behavior outline feature is segmented in video frame, some algorithm is used to establish the structure rob posture model and then to be standardized, at the last, the standardized rob postures of the same behavior are used to build the human body hidden Markov model (HMM) which has the independent position, independent translation, independent scale, independent rotating and the advantages of the free view angle, this model is verified by experiment has higher recognition rate.

Introduction

Human behavior analysis and understanding is a research hotspot in the field of computer vision, it widely used in the video surveillance, medical image analysis, sports behavior analysis, crime analysis and other fields, its key technology is the behavior modeling and motion behavior capturing which is the precondition of the human behavior analysis and understanding. Using an effectively fixed model to describe the basic human action is the behavior modeling, and different actions can be distinguished by different behavior modeling parameters of the models. Behavior capturing is seeking the human behavior model parameters from the motion outline extracting from the motion video frames to model the behavior.

Human pose which describes people's posture in an image in video frame is static image, and behavior which made up a series of consecutive frame in video is a continuous motion, thus, behavior modeling is a comprehensive description of a sequence of the human pose models which described in a series of static posture. Before the human behavior modeling, the pose model must be built firstly, the pose model is descript by many methods which include stick graph model, 2D model and 3D model etc. [1]. The human pose was descript by points which represent the people's joints and line segments which represent the people's skeleton body in the stick graph model[2,3,4,5] . 2D human pose model use simple rectangles to descript the every parts of the people [1,3,6]. In order to represent the depth information of the people's pose in the video frame, 3D human pose model was proposed, it compose two departments: pose model which represent the relationship between peripheral of human body[7,8] and skeleton model which is the stick graph model. The problems of shield and collusion have been settled by the depth information in the 3D model which is fit for multi-view human action trace.

Human behavior model is the series of pose model sequence, lots of human behavior model were using to analysis and understanding the human action. A. Bouziane, Y.Chahir , M.Molina, F.Jouen [9] prosed a 3D Zernike moments based human behavior model which using Markov spatio-temporal random walks build the human behavior model and using it to analysis the human behavior, Lin Lin, YongSu Seo, Mitsuo Gen, Runwei Cheng[10] prosed a method which using color feature and distance feature to compute the similarity of the observation object to trace the human motion object, this trace was used to create the human behavior model to analysis and understanding the human action. Huang, Kong and Li [11] prosed a human recognition method based on Fourier combined

with HMM, using Fourier transform implements the invariance of image rotation, move and scale, then, improved HMM was used to extracts feature vector for classing the object. Panagiota Antonakaki, Dimitrios Kosmopoulos, Stavros J. Perantonis etc. [12]proposed a human behavior understanding method based on bottom-up multiple camera system, using two different standards (the time length and the behavior of the track) to divid the behaviors into normal and abnormal behavior, this system using a support vector machine (SVM) identifier to identify the abnormal behavior.

A structured recognition method of human behavior based on HMM is proposed to build the human action in this paper, only one camera which does not need to be adjust should be used in this approach. The robustness of the human behavior recognition model which is established by the free view angle of the camera is improved. The detail information is descript in the following.

Human Behavior Modeling

Human Pose Modeling. As we all know that the whole human body composed with head, torso, left upper and lower arm, right upper and lower arm, left thigh and leg, right thigh and leg etc.10 parts. Here, we integrated the advantages of stick graph and 2D rectangular model, using the simplified stick graph model to describe the human pose model, in this model, each part of the body is described by line segments and the each part of the connection is described by dots. This model is shown in Fig. 1.
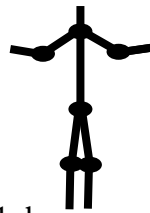

Fig. 1. human pose model

This model can be described by the following expression.

$$C(i, j) = \begin{cases} 0, i \quad relates \quad j \\ 1, i \quad independence \quad j \end{cases} \tag{1}$$

Different people have different height, thus, there is different length in the same parts in different people body, but, each person's proportion relationships between the parts are same almost. Fig. 2 shows the proportion relations of each part of the adults.
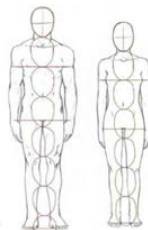

Fig. 2. the ratio of each part of the adults

While standing, the people's height is 7 times of the head, torso is 2.5 times of the head, upper limbs is 3 times of the head, lower limb is 4 times of the head.

Human Behavior Modeling. There are many human behaviors and tracking model, the present study on this domain is very many, it can be divided into three layers: rough description, middle tier description and detailed description [13, 14, 15]. There are two behavior recognition methods: static method and dynamic method. In the static method, to recognize the behavior, you must analyzing every frame image and extract the characteristic to create the behavior model firstly. In the dynamic method, the behavior model is created in the processing of the recognition. Here, a dynamic method based on HMM is used to created the human behavior model, it use Baum-Welch method, Forward Backward and Viterbi method to trance the human behavior in this method. The HMM can be described in Fig. 3. In Fig. 3, S_t is the hidden state, and it is the estimate of the human posture, O_t is

the observe state, and it is image sequence of the human video. The processing of human tracing and the human behavior model learing is shown in the following expression
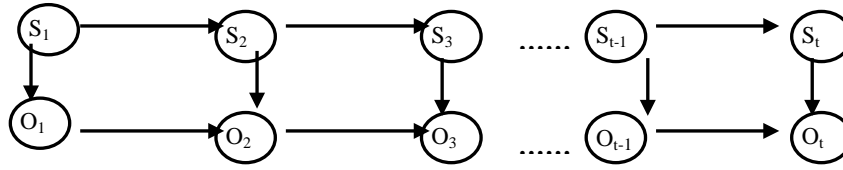

Fig. 3 HMM.

$$P(S_{1:T}, O_{1:T}) = \prod_{t} P(S_t \mid S_{t-1})P(O_t \mid S_t) \tag{2}$$

$$S_{1:T} = \{S_1, S_2, ..., S_t\}$$

$$O_{1:T} = \{O_1, O_2, ..., O_t\}$$

The training process is a probabilistic inference mode, formula 3 shows the method of getting the maximum probability in a video sequence images.

$$\begin{aligned}\hat{S}_{1:T} &= \arg_{S_{1:T}} \max P(S_{1:T} \mid O_{1:T}) \\ &= \arg_{S_{1:T}} \max P(S_{1:T}, O_{1:T}) \\ &= \arg_{S_{1:T}} \max \prod_{t} P(S_t \mid S_{t-1})P(O_t \mid S_t)\end{aligned} \tag{3}$$

S_t is a hidden state in video frame in the HMM, it is an two-dimensional vector which in the following.

$$S_t = \begin{bmatrix} P_t \\ A_t \end{bmatrix}$$

P_t is the position coordinates of the being traced object. is the appearance vector description of the being traced object. The assumption of the probability model is shown in the following.

$$P(S_t \mid S_{t-1}) \propto e^{-\|S_t - S_{t-1}\|^2} \tag{4}$$

$$P(O_t \mid S_t) = P(O_t(P_t) \mid A_t) \propto e^{-\|O_t(P_t) - A_t\|^2} \tag{5}$$

Formula 4 can be used to describe that the being traced object from one frame transfer into another using Brownian motion method. Formula 5 is our probability model.

Conclusions

A stick human behavior model based on HMM which can recognize the human action effectively was proposed, the modeling process is divided into three steps: firstly, extracting the posture from the video frame and getting rid of the shadow of the object to establish the pose model, secondly, standardizing the pose model, finally, using the normalized pose model to set up a dynamic HMM of the human behavior. Because of the independence of the position, translation and rotation of the normalized human behavior HMM, this method has the independence of the position, translation and rotation too.

Acknowledgments

References

[1] Li Haojie1, Lin Shouxun1, Zhang Yongdong .A Survey of Video Based Human Motion Capture .JOURNAL OF COMPUT ER-AIDED DESIGN & COMPUTER GRAPHICS, Vol.18, No.11,Nov, 2006.1645-1650

[2] ZHU ANG Yue-T ing,LIU Xiao-Ming,PAN Yun-He,YANG Jun. 3D Human Skeleton Reconstruction from Motion Image Sequence. JOURNAL OF COMPUT ER AIDED DESIGN AND COMPUTER GRAPHICS, Vo l. 12, No . 4,Apr . , 2000,245-250

[3] ZHUANG Yue-Ting, LIU Xiao-Ming, and PAN Yun-He. A VIDEO-BASED HUMAN SKELETON EXTRACTION TECHNIQUE FOR ANIMATION. JOURNA L OF COMPU TER RESEARCH & DEVELOPMENT, Vo l. 37, No . 4,498-506. Apr. 2000

[4] Zhou Jin, Liu Jiwei, Zhang Lei. Research on human skeleton extraction. Micro computer information,2006,Vol.22,No7-1,227-229

[5] LIYan-t ing, LUO Yu-p in, TANG Guang-rong. Activity recognition method of multiple view angles from monocular videos. Computer Applications, V o.l 26 No. 7, Ju ly 2006,1592-1594

[6] SU Bo2chao ,CHEN Gang ,CHE Ren2sheng.Kinetic people model learning of modified HMM. Optics and Precision Engineering. ol. 17,No. 6,Jun. 2009,1485-1495

[7] Deut scher J, Blake A, et al □ Art iculat ed body mot ion capture by annealed particle filtering [C] Proceedings of IEEE International Conference on Computer Vision and Pat tern Recognition, Hilt on Head, 2000, 2: 126- 133

[8] Gavrila D M, Davis L S□ 3-D mode-l based t racking of humans in action : a multi view approach [C] Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, San Francisco, 1996: 73- 80

[9] A. Bouziane, Y. Chahir, M. Molina, F. Jouen.Unified framework for human behaviour recognition: An approach using 3D Zernike moments. Neurocomputing, Volume 100, 16 January 2013, Pages 107-116

[10] Lin Lin, YongSu Seo, Mitsuo Gen, Runwei Cheng. Unusual human behavior recognition using evolutionary technique.Computers & Industrial Engineering, Volume 56, Issue 3, April 2009, Pages 1137-1153

[11] Huang Jing，Kong Lingfu，Li Haitao. Human Behavior Recognition Based on HMM. Computer Simulation, 2011,Vol.28.No.7,pp:245-248

[12] Panagiota Antonakaki, Dimitrios Kosmopoulos, Stavros J. Perantonis. cting abnormal human behaviour using multiple cameras. Signal Processing, Volume 89, Issue 9, September 2009, Pages 1723-1738

[13] ZHANG Li-jun, WU Xiao-juan, SHENG Zan, QI Lei.Behavior Recognition Method in Complex Environment Using HMM.Computer Engineering, Vol.34 No.7,212-214

[14] Hötter M, Thoma R. Image Segmentation Based on Object Oriented Mapping Parameter Estimation[J]. Signal Processing, 1998, 15(3): 315-334.

[15] Musmann H G, Hötter M, Ostermann J. Object-oriented Analysis-synthesis Coding of Moving Images[J]. Signal Processing: Image Communication, 1989, 1(2): 117-138.