Research and Application of Human-Computer Interaction System Based on Gesture Recognition Technology

Zhenxiang Huang, Bo Peng, and Juan Wu

College of Information and Electrical Engineering,
China Agricultural University, Beijing 100083, China
{zhenxianghuang,pengbo_cau,juanwu_cau}@126.com

Abstract. With the gradual improvement of the computer performance and the use of computer more deeply in many fields, mouse and keyboard, the traditional human-computer interactive way, show more and more limitations. In recent years,gesture recognition interaction based on machine vision is used more and more widely due to its simple,nature,intuitive and non-contact advantages, which is becoming the research hotspot in the world. This paper mainly studies the main idea about DTW,HMM,ANN and SVM methods used in gesture recognition.This paper also expounds the basic model and future applications of gesture interaction system based on machine vision and research meaning of its application in agriculture.

Keywords: gesture recognition, DTW, HMM, ANN, SVM, agriculture.

1 Introduction

With the rapid development of information technology and gradual increase of computer's performance, human-computer interaction has become an important part in people's daily life when the use of computer is deepening in various fields.

The traditional human-computer interactive way,such as mouse and keyboard,shows more and more limitations,especially in emerging application fields of virtual reality, augmented reality and pervasive computing.The fact promotes the research on human-computer interaction technology develop towards the direction of people-oriented, free and direct manipulation.Gestures, playing an important role in the expression of a specific intent in people's daily life, are adopted as a new human-computer interaction way in recent years. The gesture interaction based on machine vision has become the research focus of human-computer interaction due to its simple, natural, intuitive and non-invasive advantages.Gesture recognition is a technology which identifies a variety of gestures according to certain rules by computer , instructs the computer to translate into corresponding control commands or semanteme with the goal of computer operating or information exchange[1].

D. Li and Y. Chen (Eds.): CCTA 2012, Part I, IFIP AICT 392, pp. 210–215, 2013.

2 Gesture Interaction System Model Based on Machine Vision

The gesture interaction system based on machine vision input gestures through the acquisition module, in multocular case,cameras are distributed in accordance with a certain relationship in front of users[2],in monocular case,it require the plane where the camera is and the plane of the user's hand movements should be in the same level[3].

After reading the gesture images, interaction system detects gestures from the data stream on the basis of complexion and motion information[4],then separates the gesture signal from the video signal and launch trajectory tracking.The CamShift algorithm[5] and Kalman filter algorithm are commonly used.In the gesture analysis stage, the analysis process includes feature extraction and model parameter estimation.In the identification stage, gestures are classified by the model parameters and generated description as required.At last,system drive the specific applications according to generated gesture description.

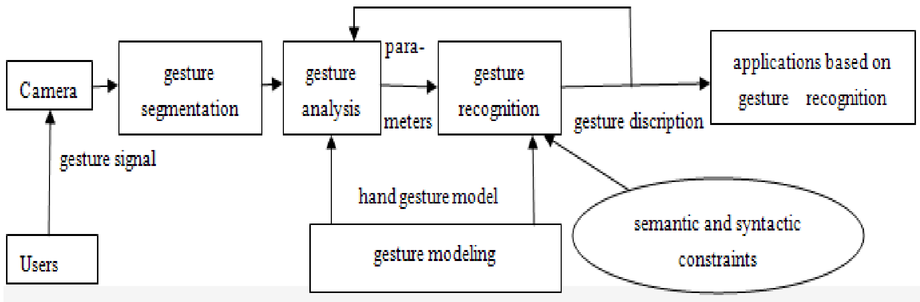

Fig. 1. Basic gesture interaction system model based on machine vision

3 Main Algorithms on Dynamic Gesture Recognition

Trajectories matching method and state space modeling method are the main dynamic gesture recognition algorithms.The most representative trajectories matching algorithm is Dynamic Time Warping(DTW) algorithm, and the typical state space modeling algorithms include Hidden Markov Model(HMM), Artificial Neural Network(ANN) and Support Vector Machine(SVM).

3.1 DTW Algorithm

Human activities can be modeled by structures of high-dimensional temporal trajectories,gesture recognition can then be performed by measuring the similarity or the distance between the input gesture trajectories and existing gesture trajectories template from the training sample set[6]. The Dynamic Time Warping(DTW) algorithm is designed to exploit some observations about the likely solution to make the comparison between sequences more efficient. DTW algorithm assumes that the endpoints of two modes are aligned accurately,then it translates the matching problem into the problem how to use dynamic programming techniques to find the optimal path through the limited grid efficiently.The sequences are warped non-linearly in the time dimension to determine measure of their similarity independent of certain non-linear variations in the time dimension. The advantages of DTW method are simple in concept, efficient and allow sufficient flexibility between the conceptual model and reference model[7].

3.2 HMM Algorithm

The state space modeling method is a algorithm that models for spatiotemporal characteristics of the gesture trajectories,it treats the trajectories as a series of transfer between the states. The specific approach is to learn the state transition parameters from training samples at first,then appraise the trajectories with different gesture models to achieve results. The typical algorithms include HMM,ANN and SVM.

A Hidden Markov Model(HMM) is a double stochastic process model which is represented with parameters and used to describe probability statistical properties of stochastic processes.It includes the random process of the state transition and the random process of observation symbols output. It evolves from the Markov Chain.Markov Chain is a mathematical system that undergoes transitions from one state to another, between a finite or countable number of possible states. It is a random process characterized as memoryless: the next state depends only on the current state and not on the sequence of events that preceded it. HMM can be regarded as a general form without the constraints of the Markov Chain[8]. As a widely used statistical method,HMM that general topology structure has good ability to describe the spatiotemporal variation of the hand signals, but the result is uncertain due to HMM has more than one migration curve for a given output result,so it can't determine the state matrix of input set by output results. In addition, HMM can't defined migration rules accurately, for example, the number of states and the number and types of migration.It depends on priori knowledge and try to determine these rules. In fact, the migration mapping result of Markov Process isn't satisfactory and the topology structure of HMM is general. These all cause the model too complicated in analysis of hand signals and large calculation of training and recognition.

3.3 ANN Algorithm

An Artificial Neural Network(ANN) is a technology that imitates the structure and/or functional aspects of biological neural networks by advanced engineering technologies. Its purpose is to enable the robot to perceive,learn and reasoning like human brain. ANN consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. BP Neural Network Model is used most widely in recent years. It is a multilayer feedforward neural networks of one-way transmission, have unilaminar or multilayer hidden layer nodes except input nodes and output nodes, no coupling among nodes of the same layer, input signal pass through all layer nodes in turn from input layer nodes to output layer nodes. Output results of each layer nodes only affect output results of the next layer nodes.BP Neural Network Algorithm has characteristics of self-organizing and self-learning for information processing, has strong ability of anti-noise, model spread and processing incomplete models.Its main defect is low convergence rate and unavoidable overfitting phenomenon.

3.4 SVM Algorithm

A Support Vector Machine(SVM) is a machine learning method based on Statistical Learning Theory(SLT), primarily research how to get the best results from small sample.In other word, it seeks a optimal balanced scheme between complexity of model and learning ability on the basis of limited sample information[9]. It solves practical problems well, such as small sample, nonlinearity, over learning, high-dimension pattern recognition and local minima, by the method of structural risk minimization principle and the kernel function. Due to its success in dealing with the relationship between the dimensions and computed strength and enhancing the computational efficiency greatly,SVM plays an important part in real-time dynamic gesture recognition.

4 Gesture Recognition Interaction Application in Agriculture

As a new human-computer interactive way that different from traditional interface, the characteristics of gesture recognition interaction like intuitiveness and naturalness, detemines its good application prospects in many fields,for example,in agriculture.

In the virtual agricultural field,operating objects in virtual environment by gestures can simulate the plant shape that crops may reach maximum yields accurately, imitate crop row spacing of intercropping and companion planting intuitively. These can provide useful reference for planning cultivated land reasonably. By the simulation of virtual agriculture,learning of agricultural technology will be visualization, interaction will be realistic,then the agricultural practitioners will understand and master agricultural knowledge better.

In the field of agricultural production, gesture recognition interaction will lower the use barrier of computer,which will help farmers to use intelligent systems for agricultural production conveniently.It also can control agricultural robot in long distance for collection or fishing in extreme environments,such as poor weather conditions[10].

Gesture recognition also can be used for interaction of information display system in public places including agricultural production site where mouse and keyboard aren’t suitable to be used[11]. In addition, gesture recognition interaction can be used for distance education system and video conference system to facilitate agricultural production and spread last agricultural information and technologys[12].

5 Gesture Recognition Interaction Application Prospects

With the help of gesture recognition technology, remote control of household appliances such as TV, DVD, stereo can be achieved and provide convenience for people's life[13][14],game operation will also be more natural and game playing experience will be enhanced.

Gesture recognition technology can be used for sign language translation system[15],then deaf-mute people can communicate with computer and normal human by it,which will improve their education level and provide convenience for two-side communicate.

With the development of gesture recognition technology, mouse and keyboard being replaced in the near future will not be a dream only, which will reduce the cost of human-computer interaction and e-waste, be consistent with the historical trend of green low-carbon.

6 Conclusion

As an important development direction of new generation human-computer interaction system, hand gesture recognition has raised widespread concern of researchers and have achieved certain results. Gesture recognition technology combined with agriculture informationization is good for the promotion of intelligent agricultural production equipment, rapid popularization of agricultural technology and fine control of agricultural production. These will significantly reduce production costs, increase productivity and make an important contribution to world food security.

References

1. Shneiderman, B.: Direct manipulation:A Step Beyond Programming Languages. IEEE Computer 16, 57–69 (1983)
2. Olsen Jr., D.R.: Where Will We Be Ten Years from Now. In: 10th annual ACM Symposium on User Interface Software and Technology, pp. 115–118. ACM, New York (2007)
3. Pavlovic, V.L., Sharma, R., Huang, T.S.: Visual Interpretation of Hand Gestures for Human-Computer Interaction:A Review. IEEE Transactions on Pattern Analysis and Machine Intelligence 19, 677–695 (1997)
4. Howe, L.W., Farrah, W., Ali, C.: Comparison of Hand Segmentation Methodologies for Hand Gesture Recognition. In: International Symposium on Information Technology, Kuala Lumpur, Malaysia, vol. 2, pp. 1–7 (2008)
5. Mahmoudi, F., Parviz, M.: Visual Hand Tracking Algorithms. In: International Conference on Geometric Modeling and Imaging, London, England, pp. 228–232 (2006)
6. Psarrou, A., Gong, S., Walter, M.: Recognition of human gestures and behaviour based on motion trajectories. Image Vision Comput. 20, 349–358 (2002)
7. Ren, H.B., Zhu, Y.X., Xu, G.: Vision Based Recognition of Hand Gestures:A Survey. Acta Electronica Sinica 2, 118–121 (2000)

8. Charniak, E.: Statistical Language Learning. MIT Press, Cambridge (1993)
9. Cristianini, N., Shawe-Taylor, J.: An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press, England (2000)
10. Wang, W.T., Li, S.Q.: Research and Implementation of the Object-Oriented Virtual-Hand Technology. Computer Engineering & Science 2, 45–47 (2005)
11. Chen, Q., Malric, F., Zhang, Y., Abid, M., Cordeiro, A., Petriu, E.M., Georganas, N.D.: Interacting with Digital Signage Using Hand Gestures. In: Kamel, M., Campilho, A. (eds.) ICIAR 2009. LNCS, vol. 5627, pp. 347–358. Springer, Heidelberg (2009)
12. Mitra, S., Acharya, T.: Gesture Recognition:A Survey. IEEE Transactions on Systems, Man and Cybernetics-part C: Applications and Reviews 37, 311–324 (2007)
13. William, T.F., Craig, D.W.: Television Control by Hand Gestures. In: IEEE International Workshop on Automation Face and Gesture Recognition, Zurich (1995)
14. Chang, J.S., Kim, S.H., Kim, H.J.: Vision-Based Interface for Integrated Home Entertainment System. In: Marques, J.S., Pérez de la Blanca, N., Pina, P. (eds.) IbPRIA 2005. LNCS, vol. 3522, pp. 176–183. Springer, Heidelberg (2005)
15. Yang, M.H., Ahuja, N., Tabb, M.: Extraction of 2D motion trajectories and its application to hand gesture recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 24, 1061–1074 (2002)