Research on Dynamic Gesture Recognition Algorithm Based on Multi-touch Human-computer Interaction Devices in Classroom Teaching

Zhao Zhang[1, a], Xiuling He[2, b], Zengzhao Chen[2, c] and Ke Wu[2, d]

National Engineering Research Center for E-Learning, Central China Normal University, Wuhan, 430000, China

[a]nercelzz2015@mails.ccnu.edu.cn, [b]xlhe@mail.ccnu.edu.cn

[c]llczz@qq.com, [d]wuke@mail.ccnu.edu.cn

Keywords: Human-computer Interaction; Dynamic Gesture Recognition; Education Cloud Integration Platform of starC

Abstract. In this paper, we propose a dynamic gesture recognition approach based on multi-touch devices. This method contains three stages: dynamic gesture stroke collection, gesture feature extraction, gesture analysis and recognition. In order to recognize nine kinds of gestures in classroom interaction process, a point packet model with fault tolerance is proposed by applying convex of the gesture stroke, track length width ratio, geometric feature ratio and Freeman chain code information. Furthermore, this method combines the number of track points and fuzzy pattern recognition method, establishes the membership function of fuzzy set, and find out the fuzzy similarity relationship of gesture recognition. The experiments show that gesture recognition rate reaches 93% to recognize nine different dynamic gestures.

Introduction

Human-computer interaction refers to a mean of communication between human and computer [1], and it is also a signal and action information exchange between human and computers. The evolution of computer brings not only the computer performance, but also the changes of the way how we input, we are more likely to input through the touch screen than traditional input way like mouse or keyboard, thus as a natural and humanized way of human-computer interaction, dynamic gesture interaction is being applied in more areas [2].

With the rapid development of information technology, e-learning equipment and teaching tools software grow up gradually, and the traditional teaching model like chalk blackboard is gradually replaced by electronic whiteboard and other multimedia devices [3]. In the classroom teaching, some of frequent operations like clicks, drag and drop bring extra operating burden to teachers and students [4]. However, the current common gesture operating system can only handle some easy operations that are not suitable for classroom interaction. This paper mainly focuses on the human-computer interaction of classroom whiteboard tools, and find out five frequently operations (select, erase, function menu, add and delete pages and page turn), then design some dynamic gestures that can be applied in classroom. We propose a dynamic gesture recognition method based on number of track points and fuzzy pattern recognition of dynamic gestures are completed in the foundation of this. In the teaching process, the application of dynamic gestures ensures that the teachers can naturally and smoothly edit and modify the teaching contents, and improve teaching efficiency after getting rid of the inconvenience and restraint from the devices like keyboard and mouse.

As far as we can see, most interaction of the multi-touch devices still operating objects in a way of two-point-touch, and common gestures are two hands or two fingers opening, closing and rotating, which achieve the changes of picture and other screen object, and fingers sliding which can do actions like scrolling up and down. Although these gestures solve the problem in initial single-touch interaction, when applied in multi-touch handwriting scene, they are still very inconvenient for students and teachers as a matter of fact that when they need to selection, edit,

erase, flip the content, they still need to use the toolbar. As a result, these common interaction has not been improved.

From the 60s of last century, scientists at home and abroad have done a lot of research on multi-touch technology and gesture recognition. In 1990, a German company called Xerox achieved scaling and rotation gestures on a image with two fingers in touch screen [5]. In 1998, FIngerworks company made a set of IGesture gesture library [6], and produced the capacitive multi-touch screen, which was bought by Apple and lay a good foundation for iPhone. In 2008, in Microsoft win7 system, finger touch function, and multi-touch are supported. And the surface book launched later represents the coming of touch screen era.

Most of the static gesture recognition is based on the image. This paper focuses on the dynamic gesture recognition of the touch devices. Since there is not any specific or common algorithms for gesture recognition on multi-touch devices, it is generally necessary to recognize with a specific devices or fixed gesture based on a particular implementation system.

For example, the Microsoft's TouchMe Gesture system, provides 30 kinds of user-customizable gestures, including single-click, sliding, two-finger zoom and rotation which supports some common operations in win10 system. However, these gestures are not adapted to the teaching environment and cannot complete the operations such as erase, add or delete pages by gesture.

The professor Lingyun Xiang from NUDT [7] classifies the frequently used gestures in military command according to the number of gesture contacts and the position of the various contacts. They recognize gestures from the number of fingers, and the changing trend of the angle, distance and direction between gestures and the axis. However, this method can only support static gesture recognition.

Gestures researchers like Dexin Wang [8] divide multi-touch gestures into atomic gestures and combination gestures. The combination gesture is modeled by Petri net uesd logical temporal and spatial relations. After that, they according to the atom gesture detected by the BP network, the transfer of combination gesture Petri Net is triggered and thus recognized combination gesture. Because of this method is limited to process time series, it is not suitable for dynamic gesture recognition. Wang Xiying [9] model the dynamic gestures in five levels using HMM Hidden Markov model. This method needs a lot of training and repeated calculation, so it is not available in real-time gesture recognition in a classroom.

This paper designs a set of general multi-point dynamic gestures in harmony with classroom environment according to human-computer interaction, and solves the problem that the handwriting input and gestures input cannot be distinguished. And we applied the dynamic gesture recognition algorithm into whiteboard applications in starC, a cloud integration platform, thus improving the interactive experience in classroom.

Dynamic Gesture Recognition Algorithm

The recognition of dynamic gesture based on multi-touch contains three stages: dynamic gesture stroke collection, gesture feature extraction, gesture analysis and recognition. In the stage of dynamic gesture stroke collection, distinguish between handwriting input and gesture input through a point packet model. Preprocess the stroke of gesture to eliminate the influence of noise, and established the model of point packet. In the stage of feature extraction, track length width ratio, geometric feature ratio and Freeman chain code information. Through these characteristics response the change of the gestures. Gesture analysis and recognition stage, for different types of gestures, to establish the corresponding membership functions, and find out the fuzzy similarity relationship of fuzzy set, combined with chain code direction information, finally complete the identification of dynamic gestures. As depicted in Fig.1.
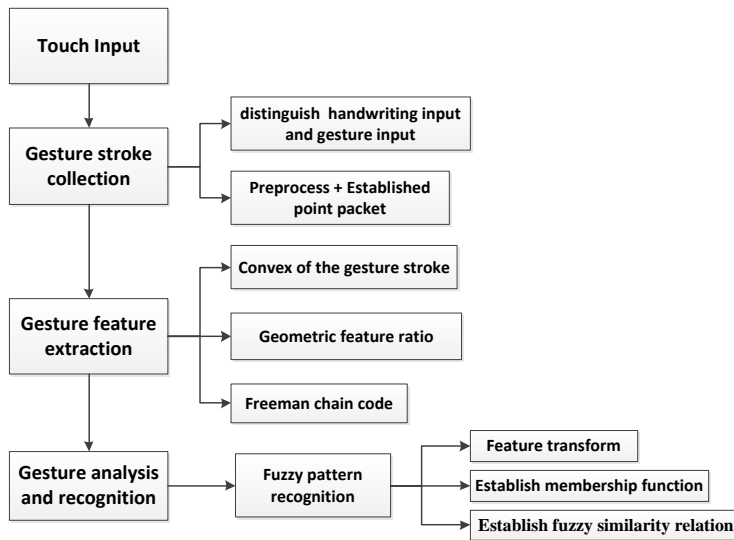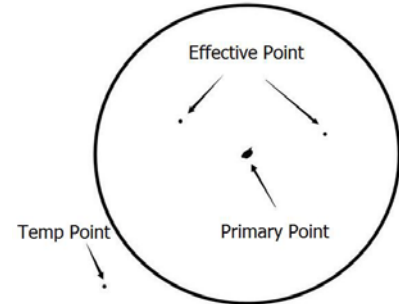
Fig.1. The flow chart of recognition algorithm   Fig.2. Point packet model

Dynamic gesture stroke collection. At present, the teaching environment of the interactive whiteboard in human-computer interaction, computer is very difficult to distinguish between gesture input and handwriting input. In this paper, through the design of a point packet model, by judging the relative relationship between the different touching points. Calculate the maximum Euclidean distance of different touching points, to effectively distinguish multi handwriting input and gesture input. The point packet model include the track information of the primary point, the number of effective points, the track information of the temporary point, the maximum Euclidean distance of points in the point packet. When the gesture operation is completed, or determine the touch operation is invalid, and the resource is released. As depicted in Fig.2.

$$V = \{P, E, T, D\} \qquad E = \{e_1, e_2, e_3 ... e_i\} \tag{1}$$

V is the collection of point packet. P is the track of primary point. E is the set of effective points. T is temporary point. D is the maximum Euclidean distance of points in the point packet. i is the number of effective point.

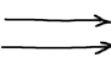When a new touch event occurs, first detected on the screen whether there is touching point packet, if there is no other touching point packet, then establish a point packet, the touching point is set as primary point. If there are other touching points and point packet, the touching point add to point packet as a temporary point. Next, calculate the maximum Euclidean distance between the touching point and the primary point in the point packet, if the distance greater than the threshold, this point set as invalid temporary noise, doesn't record data, if less than the threshold, the temporary point will set as the effective point. When all points are leave the screen, the track is collected.

Dynamic gesture definition. With the development of multi-touch technology. Multi-touch gesture with intuitive, fast and efficient advantage gradually become a trend in the future classroom [10]. However, there are still some problems in the application in the teaching process. Gesture is not suitable for the teaching environment, and the practical value is low, there is no uniform design standards. Different platforms and software are in accordance with their own understanding of gestures for gesture design, no unified theoretical basis [11]. So we find a set of recognized standards of behavior. Such as Windows, Android, IOS in general multi-touch gestures including single point gestures and multi-touch gestures, like click, double-click, press, drag, slide single touch gestures, multi-touch gestures have two finger pinch, expansion, rotation, press and click and drag etc. These gestures no need to learn and training, people easy to use.

In the teaching environment of multi-touch devices, the statistics of the most frequently used for five types of operation include selecting, erasing, function menu, adding and deleting pages, page turning. According to the teaching environment, the multi-touch environment is simple, intuitive and easy to use the principle of gesture design [12]. We design nine kinds of dynamic gesture. As depicted in Table.1.

Table.1. Dynamic gesture definition

GES	Command	illustrate	description
1	Chose		Double finger draw circle select elements
2	Erase		Palm grip into a fist, 4 points and above move on the screen
3	Function menu		Three point touch screen to keep 2 seconds or more
4	Page add		Double finger draw a triangle in the upper left corner of the screen
5	Page delete		Double finger draw a cross in the upper right corner of the screen
6	Next page		Double finger horizontal sliding to right
7	Previous page		Double finger horizontal sliding to left
8	First page		Double finger vertical sliding to up
9	Last page		Double finger vertical sliding to bottom

Dynamic gesture feature extraction. In dynamic gesture recognition, the selection of gesture feature has great influence on the accuracy of recognition results. The single feature is often affected by the environment and other factors. Considering the computational complexity, real-time recognition and feature invariance and many other factors, this paper first defines the 6 basic variables include: convex hull of gesture stroke, largest-area triangle that fits inside the convex hull, smallest area enclosing rectangle, the number of point in the packet, point packet for the screen coordinates. The first five basic variables have rotation invariance, scale scalability, translation invariance, suitable for gesture recognition system. In order to identify the gesture types, this paper selects seven characteristic parameters. Formula (2) - (8) are as follows.

Characteristic T0: The primary point and effective points in point packet

$$T_0 = M \tag{2}$$

Characteristic T1: The perimeter squared of the convex hull and its area ratio

$$T_1 = \frac{P_c^2}{A_c} \tag{3}$$

Characteristic T2: The largest triangle inside area and convex hull area ratio

$$T_2 = \frac{A_t}{A_c} \tag{4}$$

Characteristic T3: The length and width of enclosing rectangle ratio

$$T_3 = \frac{H_r}{W_r} \tag{5}$$

Characteristic T4: The convex hull area and enclosing rectangle area ratio

$$T_4 = \frac{A_c}{A_r} \tag{6}$$

Characteristic T5: Change of gesture position

$$T_5 = P_t - P_{t-1} \tag{7}$$

Characteristic T6: Gesture duration

$$T_6 = D_t - D_1 \tag{8}$$

M is the number of primary point and effective points in point packet. A is the area of the graph, P is the perimeter of the graph, Hr and Wr is height and width of enclosing rectangle, P_t and P_{t-1} is the current t and t-1 moments of the position. The complexity of the above 7 features is low, and the operation efficiency is high, which can guarantee the system real time feedback. T_0 as a prerequisite to judge the type of gesture, $T_1 T_2 T_3 T_4$ is proportional feature, which can satisfy the invariance of rotation, translation and scaling. Use Graham scanning method[13] to obtain the convex hull of gesture stroke.

Gesture analysis and recognition. First through the characteristics T1, gesture is divided into two points, three points, multi-point gesture.

For the situation that the number of primary point and effective points in point packet is two. the gesture stroke can regard as a line, circle, triangle, cross of four geometric patterns. For the geometric figure of the standard specification with reference to the MJ Fonseca [14]. Using fuzzy pattern recognition method, Select five characteristic parameters composition attribute set $T_x(T_1, T_2, T_3, T_4, T_5)$. For the gesture set G on the input set U, every $x \in U$, have one $\mu_G(x) \in [0,1]$ can represent the degree of x belong to G. $\mu_G(x)$ unique determination, called membership function of set G.

$$G = \sum_i \frac{\mu_G(x_i)}{x_i}, x_i \in U \tag{9}$$

The membership degree of a number of geometric patterns on the domain is got by the two element compare method, and the membership function of the intermediate type is established according to the principle of the maximum membership degree.

$$\mu_G(x) = \begin{cases} 0 & , & x \le a \\ \dfrac{x-a}{b-a} & , & a < x \le b \\ 1 & , & b < x \le c \\ \dfrac{d-x}{d-c} & , & c < x \le d \\ 0 & , & x > d \end{cases} \tag{10}$$

As depicted in Fig.3. A fuzzy set is defined by four values. After the selection of the characteristic for each shape, we compute these four values. Values b and c are the confidence interval of the degree of membership function, and a and d to the minimum and maximum. After we have identified the outliers in each distribution. The elimination of outliers minimizes confusion between different shape families. By collecting a large number of gestures, the membership functions of fuzzy sets of different shapes are obtained.
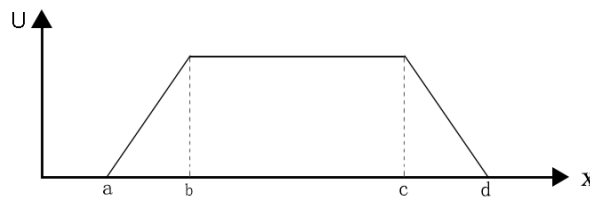

Fig.3. Membership function of fuzzy set

The choice gesture involves circle and ellipse. Use characteristic T1(P_c^2 / A_c)and T4(A_c / A_r) to

recognize. This paper defines a short axis and a long axis ratio greater than 0.6 is the right circle. When the gesture track is round, the maximum degree of membership function threshold T1 characteristics of value b=4π. When it is an ellipse, the threshold value of membership function is calculated c=17.0(As depicted in Fig.4.). In this range, that track is a circle. After a lot of gesture training, obtained confidence interval threshold, value a=12.6, value b=18.3. In the figure 4, the T1 characteristic ratio of the line is beyond the scope of the unit. The lower bound of the T1 characteristic of the triangle is higher than ellipse, and the upper bound is beyond the scope of the unit. The lower bound of the T1 characteristic of the cross is slightly higher than that of the ellipse, and the upper bound is much larger than the unit range. Therefor, if the T1 characteristic membership function value in the confidence interval of circle, or in the fuzzy interval of circle, and the track of the T4 characteristics value out of the case of the confidence interval of cross. Then $G_{tid}=1$.

The page add gesture involves triangle. Use characteristic T2（ A_t / A_c ）and T5 to recognize. In the ideal state of the input, $A_t \approx A_c$. To calculate the membership function of T2 value c=1. After a lot of gesture training, obtained confidence interval threshold value, a=0.82, b=0.85(As depicted in Fig.5). Because of there is no other geometry of the triangle fuzzy interval, so we can directly determine gesture for triangle. When the characteristic T5 value of the track is satisfied with the left upper screen coordinates. Then $G_{tid}=4$.
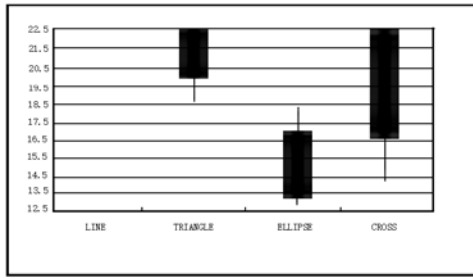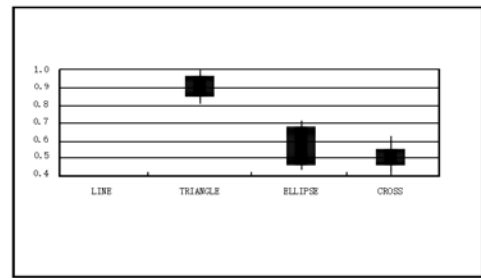

Fig.4. Percentiles for $P_c^{\,2} / A_c$ ratio   Fig.5. Percentiles for A_t / A_c ratio

The page turning gesture involves line. Use characteristic T3(H_r / W_r)and T5 to recognize. H_r and W_r are enclosing rectangle's height and width. If the stroke is line, the height and width ratio of the external rectangle is close to 0. This characteristic can be a good response to the characteristics of line(As depicted in Fig.6). For different turning direction, through the T5 characteristic calculated the cosine angle ψ of the forward direction of the X axis. Different range of ψ corresponding to different page turning gesture.

$$\cos\psi = \frac{T_5(x)}{\sqrt{T_5^{\,2}(x)+T_5^{\,2}(y)}} \tag{11}$$

$$G_{tid} = \begin{cases} 6 \quad, & -\dfrac{\pi}{4} \leq \psi < \dfrac{\pi}{4} & and \quad T_5(x) \geq 0 \\[2mm] 7 \quad, & \psi \geq \dfrac{3\pi}{4}\, or\, \psi < -\dfrac{3\pi}{4} & and \quad T_5(x) < 0 \\[2mm] 8 \quad, & \dfrac{\pi}{4} \leq \psi < \dfrac{3\pi}{4} & and \quad T_5(y) < 0 \\[2mm] 9 \quad, & -\dfrac{3\pi}{4} \leq \psi < -\dfrac{\pi}{4} & and \quad T_5(y) \geq 0 \end{cases} \tag{12}$$

The page delete gesture involves cross. Use characteristic T4(A_c / A_r)and T5 to recognize. T4 characteristic of cross is larger than the other three geometry. After a lot of gesture training, obtained confidence interval threshold value b=0.89, c=0.98, fuzzy interval threshold a=0.85, d=1.0 (As depicted in Fig.7). Through the characteristic T5 to calculation the dot product of begin vector v_b and end vector v_e divided the length of begin vector and end vector, get the angle θ of

direction. If $\theta \in \left[75^o, 105^o\right]$, then G_{tid} =5.

$$\cos\theta = \frac{v_{bx} * v_{ex} + v_{by} * v_{ey}}{\sqrt{v_{bx}^2 + v_{by}^2} + \sqrt{v_{ex}^2 + v_{ey}^2}} \tag{13}$$
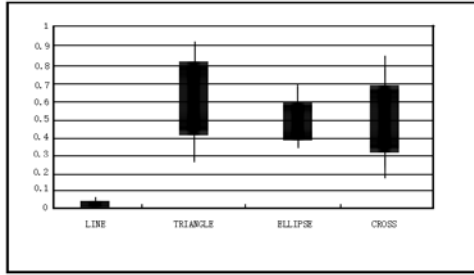

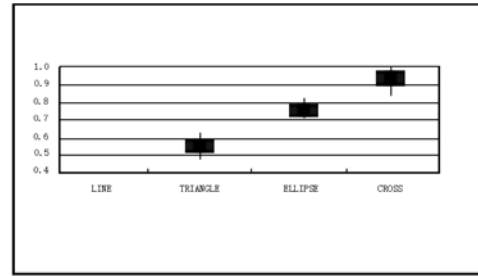Fig.6. Percentiles for H_r / W_r ratio

Fig.7. Percentiles for A_c / A_r ratio

For the situation that the number of primary point and effective points in point packet is three. If gesture movement is less than the threshold l_1, the duration of the gesture is greater than the time threshold t. Then determined the function menu gesture G_{tid} =3. For the situation that the number of primary point and effective points in point packet is more than 4, and the gesture track moving distance is greater than the threshold l_2, then determined the erase gesture, G_{tid} =2. The t for hold time threshold of t=2000, units in milliseconds. l for gesture moving distance threshold l_1=50, l_2 =20, unit in pixel.

If（$T_0 = 2$）

    If（Gesture is like line）   Then G_{tid} ={6,7,8,9}

    If（Gesture is like triangle）   Then G_{tid} =4

    If（Gesture is like x）   Then G_{tid} =5

    If（Gesture is like ellipse and is not like x）   Then G_{tid} =1

If（$T_0 = 3$ and $T_4 \leq l_2$ and $T_5 \geq t$）   Then G_{tid} =3

If（$T_0 \geq 4$ and $T_4 > l_1$）   Then G_{tid} =2

Through this method, we can effectively distinguish nine different kinds of gestures. The method has the advantages of small amount of calculation, strong robustness and wide application range.

Experimental Results and Analysis

This experiment used DELL P2314T touch screen as input device, starC cloud integration platform based on the whiteboard application. Three teachers and seven students who are familiar with the environment and gestures were invited. Each kind of gestures took the 100 data samples for recognition rate analysis. At the same time the operation time of each operation was done without using gestures. The efficiency analysis of the operation by dynamic gesture. Table.2 gives the recognition of nine kinds of gestures and the rate of false recognition, where TID is the number Aof gesture type, GES, the name of the gesture, REC indicates the rate of gesture recognition, ERR (TID) indicates the probability of error identification for TID. Table.3 compares the average operation time 10 times through gestures and non gestures, because the four page turn gesture is quite similar, so the four kinds of gestures were averaged, in order to reduce the error in the experimental process, using each operation time that experiment can complete 10 times as the statistical unit of time. In all operations, it must return to editing state after completing operations. The use of function menu gesture is to record the time to arouse toolbar.

Table.2. The recognition rate and recognition error rate

TID	GES	REC	ERR(1)	ERR(2)	ERR(3)	ERR(4)	ERR(5)	ERR(6)	ERR(7)	ERR(8)	ERR(9)	ERR(0)
1	Chose	94.9%	—	0.0%	0.0%	1.2%	0.6%	0.0%	0.0%	0.0%	0.0%	3.3%
2	Erase	95.4%	0.0%	—	1.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	3.5%
3	Function menu	93.8%	0.0%	3.4%	—	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.8%
4	Page add	94.2%	1.7%	0.0%	0.0%	—	0.0%	0.0%	0.0%	0.0%	0.0%	4.1%
5	Page delete	93.1%	2.1%	0.0%	0.0%	0.0%	—	0.3%	0.1%	0.2%	0.3%	3.9%
6	Next page	98.6%	0.0%	0.0%	0.0%	0.0%	0.0%	—	0.0%	0.1%	0.2%	1.1%
7	Previous page	97.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	—	0.2%	0.1%	1.9%
8	First page	97.4%	0.0%	0.0%	0.0%	0.0%	0.0%	0.3%	0.2%	—	0.0%	2.1%
9	Last page	97.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.3%	0.0%	—	2.3%
0	No sign	—	—	—	—	—	—	—	—	—	—	—

Table.3. Use gestures and don't use gestures of time spent

GES	Chose	Erase	Function menu	Page add	Page delete	Page turn
Use gestures	15.6s	12.3s	21.5s	14.3s	15.6s	11.1s
Without use gestures	23.5s	18.5s	19.5s	15.8s	17.3s	21.9s

The average accuracy rate of recognition of nine different gestures is 95.9%, page turn gesture average recognition accuracy is higher than 98%. Compared with the RBF neural network method rate of recognition 92.8% proposed by Wang Xiaoqing [15], the recognition effect is better. About dynamic gestures operating efficiency, compared to the operation use tool bar can save 70% of the time. The efficiency of classroom interaction better, only function menu gesture is higher than the operate by toolbar, because all the operations are in the 23 inch touch screen, in the actual classroom, teachers need to move to the edge of the electronic whiteboard to complete the operation, and the function menu gestures can be completed at any position. Gesture function have better experiences than toolbar.

This multi-touch dynamic gesture recognition method is applied in the whiteboard in starC cloud platform. Command mapping by gesture interaction, without affecting the multi handwriting input, using the gesture complete the operate in the select, erase, page add and delete, page turn operation. The gesture interactive function of electronic whiteboard is improved.

Conclusion

To solve the problems in classroom teaching environment with multi-touch human-computer interaction devices, we propose a hand gesture recognition method based on the number of track points and fuzzy pattern recognition. Compared with RBF neural network and Hidden Markov Model method in hand gesture recognition, we effectively distinguish multi handwriting input and gesture input, and achieve dynamic hand gestures recognition. Experiments also show that this method is implemented with high accuracy.

Acknowledgement

References

[1] Chao Sun, Zhiquan Feng, Yang Li, et al. A Survey of Gesture based Interaction[c]. CHCI, 20l0: 277-281

[2] Yaxiang Ren. Survey of human-computer interaction development based on hand posture and gesture recognition[J]. Computer Engineering and Design, 2006,27 (7): 1201-1204.

[3] Zongkai Yang. Ten years of Development of E-learning: future classroom, future school, future teachers, future education [J]. China Educational Technology, 2011(9).

[4] Yinghui Shi, Harrison Hao Yang, Zongkai Yang. Research on Application of Interactive Whiteboard in Foreign Education [J]. China Educational Technology, 2012(5):99-103.

[5] Michael J Smith．Human Computer Interaction：software and Hardware interface[J]．The fifth

international Conference On Human-Computer interaction(HCI International 93), 1993.

[6] Guangliu Zhang, Lin Li, Jianbo Tan, Research on single point gesture recognition method in touch device[J]. Computer application technology, 2011(1):7-11.

[7] Yunxiang Lin, Guohua Zhang, Ting Ye. Research on natural gesture recognition method based on multi-touch[J]. J Natl Univ Def Technol,2010(01):127-132.

[8] Dexin Wang, Chonglin Shi, Maojun Zhang. Research on Multi-touch gesture recognition based on Petri net and BPNN [J]. Pattern Recognition and Artificial Intelligence, 2010(09):408-413.

[9] Xiyin Wang, Guozhong Dai, Xiwen Zhang. Research on Complex dynamic gesture recognition based on HMM-FNN model[J]. Journal of Software, 2008(09):2303-2312.

[10] Lu Huang, Research on gesture recognition based on Interaction [D]. Central China Normal University. 2011

[11] Nielsen M, Störring M, Moeslund T B, et al. A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI [J]. Journal of Structural Geology, 2003, 17(17):1445-1453.

[12] ChenYan Fang, Weiqing Kong. Research on Multi-touch gesture recognition algorithm [J]. Software Engineer, 2014(11):27-30.

[13] Peide Zhou, Mengyin Fu. A linear time algorithm for the convex hull of a simple polygon [J]. Computer Engineering & Science, 2002,24(03):1-2.

[14] MJ Fonseca, JA Jorge. Using fuzzy logic to recognize geometric shapes interactively [J]. IEEE, 2000, 1:291-296.

[15] Xiaoqin Wang, Ge Chen. Research on multi-touch gesture analysis and recognition algorithm[J] Computer Science. 2012.6(39)522-525.