

Gesture browsing method for geological data

Huang Chen¹, Li Jianhua², Yu Hanchao³, Shen Jianghai⁴

Abstract: In big screen environment, careful observation of geological data is significant for geological workers. However, traditional mouse is not suitable for free operation in big screen environment. Influenced by illumination, background and some other things, gesture recognition based on ordinary camera is neither real-time nor robust. Therefore, considering the geological workers' common gestures in the actual manipulation for geological data, an adaptive Kalman filter method based on the Kinect is proposed, which can not only obtain real-time and continuous hand depth data, but also further identify the geological data manipulation gestures. Finally we realize a geological data control browsing system with high robustness. The results show that adaptive Kalman filter method is conducive to manipulate for geological data.

Key words: geological data; gesture recognition; Kinect; adaptive Kalman filter

1 Introduction

Detailed observation of geological data plays an important role in the study of Petroleum Geology. Digital geological data^[1], on the one hand, avoid the solid core disorder deletion and destruction which caused by frequent sampling, observation to describe, photography, natural weathering and other factors. On the other hand, due to the rapid development of big screen display system and a wide range of applications, the digitalized geological data can be displayed and observed on a big screen systems which greatly improves the efficiency in the use of core data^[2]. However, in big screen environment, the traditional mouse mode is not conducive to observe digital geological data. Based on the existing geological data digitization technique^[3], we implement geological data manipulation human-computer interaction interface in big screen system from the perspective of geological research in this paper.

This paper aims to deal with two problems as follows. On the one hand, in big screen projection environment, the traditional mouse and keyboard operation is not conducive to geological data translation, zooming and other operations. Natural gesture can be used to compensate for this

¹ Huang Chen(✉)

School of Geosciences, Yangtze University, Wuhan 430100, China
e-mail: huangchen@ict.ac.cn

¹ Huang Chen , ² Li Jianhua

School of Geosciences, Yangtze University, Wuhan 430100, China

³Yu Hanchao, ⁴Liu Junfa

Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China
Beijing Key Laboratory of Mobile Computing and Pervasive Device, Beijing 100190, China

⁵Shen Jianghai

College of Computer Science, Yangtze University, Wuhan 430100, China

deficiency. On the other hand, fingers and hand gestures have to be accurately identified and tracked, so as to realize the freedom gestures manipulation method. However, influenced by illumination, skin color, background and some other things, the gesture recognition based on ordinary camera is not real-time or robust.

In this paper, Kinect is used to acquire the 3D spatial data of hands and fingers. Then, the parameters of hand and fingers are extracted accurately in real-time. Finally, user's operating intentions are recognized. The existing hand tracking methods based on Kinect are affected by computer hardware condition and transient cover of environment. This weakness will bring the lack of frame and frame losing, which can cause failure in operation identification and so on. Therefore, an adaptive Kalman filter method is proposed, which is able to compensate for the inadequacy of Kinect lost frames or temporary shelter. The new method greatly improves gesture recognition accuracy, tracking robustness and the accuracy of manipulation the geological data.

2 Related work

In practical application, a lot of gesture interaction is needed, but building a gesture recognition system with robustness is still a challenging research issue^[4]. Since the gesture form changes a lot, variety of tagged gesture recognition methods are widely utilized, meanwhile the method of gesture recognition based on vision has become a hot research topic.

Using Hidden Markov Model(HMM) to test the dynamic gesture trajectory is the key point of gesture recognition. Wu^[5], Jiang^[6], Zhang^[7], by hand gesture segmentation, did continuous dynamic gesture recognition using HMM, and obtained a certain effect. Silanon^[8] employed HMM for dynamic gesture motion analysis to identify the Thai alphabet. However, the traditional gesture recognition method based on RGB camera is susceptible to light, background, camera characteristics. Zhang^[9] implemented the gesture track recognition based on the depth image information get by Kinect. He put forward a gesture trajectory recognition method based on HMM, which is adopted to control the movement of intelligent wheelchair.

Kinect can solve the illumination, background, camera characteristics on the impact of gesture trajectory recognition. However, the use of HMM can only realize the recognition of a signal gesture sequence, but cannot identify the gestures in real time, which leads to hysteresis of gestures. In the case of a short block, there will be a great influence on the precision of gestures trajectory identification. The gesture recognition work using depth information by Pedro Trindade^[10], Jesus^[11], Suarez and Daniela Ram´irez-Giraldo^[12] obtain satisfactory results. But it is easily affected by the lost frame and shade to use the depth information directly provided by Kinect.

Therefore, combining the digital geological data display system, the trend of big screen display and the requirements of the geological workers in the practical application, we propose an adaptive Kalman filter method and implement gesture points of steady tracking, which can make up for the deficiency of mouse and keyboard operation. Finally, according to geological data of commonly used functions in big screen and the practical gesture used by geological workers, we design three kinds of gestures which are enlarging, shrink and moving. Experiments showed that the proposed adaptive Kalman filter method can improve the accuracy in gesture trajectory recognition and geological data browsing.

3 Adaptive Kalman filter method

3.1 The definition of the trajectory

In the control process of browsing geological data, we use the information provided by Kinect to obtain feature points of hand gesture. The feature point trajectory taking Kinect as the original coordinates is defined as:

$$G(T) = \langle (x_1, y_1, z_1, t_1), \dots (x_i, y_i, z_i, t_i), \dots (x_n, y_n, z_n, t_n) \rangle \quad (1)$$

where $t_1 < \dots, t_i < \dots, t_n, t_i$ represents the time point in the process of gesture movement. x_i, y_i, z_i is the three-dimensional coordinate which taking Kinect as the original coordinates.

We used K-means clustering algorithm and the depth information obtained by Kinect to realize hand segmentation, then utilized the palm contour and the convex hull of the public as the tip of the finger points, and located the palm ^[13-15] (Figure 1.1). According to the tip of the finger point, palm of hand, we realized the gestures of posture and gesture trajectory tracking, finally obtained geological data manipulation gesture as needed.

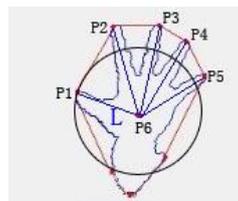


Figure 1.1 the palm contour figure

Fingertips and palm centre will be lost at the case of lost frame or temporary shelter, so we set up detection coefficient:

$$F(P_i) = \begin{cases} 1, & \text{Detect the } i \text{ th finger or palm center} \\ 0, & \text{the } i \text{ th fingertip or palm center is missing} \end{cases} \quad (2)$$

Where $1 \leq i \leq 6$.

3.2 Adaptive Kalman filter method

There are lost frames and temporary shelter problems in the Kinect, which caused fingers and palm centre lost or can't be detected. But in the geological data browsing it need real-time access to accurate and robust location information of the fingertips or the palm centre, so a new algorithm tracking the fingertips and the palm centre is urgently needed. The standard Kalman filter can be used to predict fingertips and palm centre of current frame by corresponding information of previous frame. Therefore, a new method adaptive Kalman filter is proposed, which can predict fingertips and palm centre under the condition of lost frames and temporary shelter. Compared with standard Kalman filter, the adaptive Kalman filter has lower complexity.

First, let us introduce a discrete control system and use a linear stochastic equation to describe:

$$X_{ki} = FX_{(k-1)i} + Bu_k + w_k \quad (3)$$

The observation equation of the dynamic system is:

$$Z_{ki} = HX_{ki} + v_k \quad (4)$$

Where $1 \leq i \leq 6$, X_{ki} is the state of the system for i -th fingertip or palm centre at k -time. u_k is the control variable of the system at k -time. F and B are the system parameters. Z_{ki} is the measured value of k time. H is the parameter of measurement system. w_k and v_k represent the process and measurement noise, they are usually assumed to be Gauss white noise, their covariance respectively is Q and R .

3.2.1 System prediction stage

1) The current frame of fingers and palm centre position:

$$\hat{X}_{ki} = (F_k \hat{X}_{(k-1)i} + B_k u_k)(1 - F(P_i)) + X_{ki} F(P_i) \quad (5)$$

Where \hat{X}_{ki} is the position of i -th finger or palm centre which predicted by the $k - 1$ -th frame. $\hat{X}_{(k-1)i}$ is the position of i -th finger or palm centre at the $k-1$ th frame. X_{ki} is the observed value at k -th frame. F_k and B_k are the parameter matrix of the system. u_k is the control variable of the system at k -th frame.

2) The covariance matrix of the k -th frame:

$$P_{ki} = F_k [P_{(k-1)i}(1 - F(P_i)) + P_{(k-1)i} F(P_i)] F_k^T + Q_k \quad (6)$$

In the above equation, P_{ki} is the corresponding covariance of \hat{X}_{ki} . $P_{(k-1)i}$ is the corresponding covariance of $\hat{X}_{(k-1)i}$. $P_{(k-1)i}$ is the corresponding covariance of X_{ki} . F_k^T is the transposed matrix of F_k . Q_k is the covariance of system process.

3.2.2 System update stage

The updated state estimation is calculated by

$$\hat{X}_{ki} = [\hat{X}_{(k-1)i} + K_k (Z_{ki} - H_k \hat{X}_{(k-1)i})](1 - F(P_i)) + X_{ki} F(P_i) \quad (7)$$

Where K_k is the Kalman Gain:

$$K_{ki} = P_{ki} H_{ki}^T (H_{ki} P_{ki} H_{ki}^T + R_{ki})^{-1} \quad (8)$$

Updated covariance estimation:

$$P_{ki} = (I - K_{ki} H_{ki}) P_{ki} \quad (9)$$

Where H_{ki} is the parameter of measurement system, I is a unit matrix.

After the adaptive Kalman filter processing, the gesture of fingertips and palm centre point trajectory in browsing the geological data as follows:

$$G(T_i) = \langle (x_{ic1}, y_{ic1}, z_{ic1}), \dots (x_{ick}, y_{ick}, z_{ick}), \dots (x_{icn}, y_{icn}, z_{icn}) \rangle \quad (10)$$

Where $1 \leq i \leq 6$, when $i = 6$, $G(T_6)$ is the trajectory of palm centre. After dealing with equation (2), adaptive Kalman filter which has the function of the values of prediction decreased the

number of operation.

4 Gesture recognition of geological data manipulation

The hand gestures that geological workers used in the practical application include enlarge, shrink, move up, move down, move to the left and move to the right. Therefore the method mainly made up of two sets of gestures. The first is single hand gesture which includes geological data moving in four directions. The other one is two-hand gestures, including geological data enlarge and shrink.

4.1 Gesture model

During the manipulation process, Kinect is regarded as the original coordinates, and the distance of human body to Kinect is Z_D . When facing with Kinect, we take the plane which distance to the human body is Z_0 as the touch starting surface. In order to achieve the adaptive effect, the value of Z_0 varies with the different size of human body. F_I (figure 2.1) is composed of a series of points in space which is taking Kinect as the original coordinates, defined as:

$$F_I = \{P | Z_D - Z_0 \leq Z(P)\} \quad (11)$$

Where P is pixels, $Z(P)$ is the depth value of P . When $I = 1$, the region corresponding to the left hand, when $I = 2$ it means the right one.

$$F_I = \begin{cases} 0, & Z_D - Z_0 > Z(P) \\ 1, & Z_D - Z_0 \leq Z(P) \end{cases} \quad (12)$$

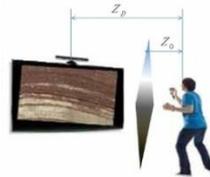


Fig. 4.1 the gesture area map

4.2 Geological data manipulation gesture

4.2.1 Moving browse for geological data

According to the geological worker's habits, they often use the right hand operation to realize the geological data in the four directions of movement. When $F_2 = 1$ and $F_1 = 1$, the right hand is detected.

$$R = \frac{1}{5} \sum_{i=1}^5 |\overrightarrow{P_6 P_i}|$$

The initial state is $L = |\overrightarrow{P_6 P_1}|$ (Figure 1), palm open when $R > L$, when $R \leq L$ for grasping

movements. Thus when $\leq L$, geological data with the movement of hand.

4.2.2 Scaling browse for geological data

According to the geological worker's habits, hands close means geological data shrink and hands away expresses the geological to enlarge. When $F_1 = 1$ and $F_2 = 1$, both hands are detected. According to the movement trend of hands to determine the geological data is enlarging or shrink.

When $i = 6$, by Equation (10) and Equation (11) to the left palm and the right palm at the same time. Equation (13) is the trajectory consist of a series of palm points in the area F_l .

$$G(T_{j6}^i) = \langle (x_{j6c1}, y_{j6c1}, z_{j6c1}), \dots (x_{j6ck}, y_{j6ck}, z_{j6ck}), \dots (x_{j6cn}, y_{j6cn}, z_{j6cn}) \rangle \quad (13)$$

Where $z_{j6cn} < Z_D - Z_0$, $j = 1$ is the left palm centre, $j = 2$ is the right palm centre.

Hence, through the left palm centre and right palm centre move trend, we can control the geological data zoom in and out.

5 Experiment and results

5.1 Experiment platform

1) Processor: Intel i5-3210M @ 2.50GHz dual-core; 2) Memory: 4GB; 3) Graphics card: NVIDIA GeForce GT 630M; 4) Depth sensors: Kinect.

5.2 Geological data gesture performance analysis

The algorithm complexity of standard Kalman filter algorithm for tracking palm centre is $O(5N^2 + 4N)$, and the hand gesture trajectory which controls the geological data in four directions of movement need to follow up with 6 points, the algorithm complexity is $O(6(5N^2 + 4N))$, the algorithm complexity is $O(2(5N^2 + 4N))$ when controls the geological enlarge or shrink (N is the number of filter coefficients). The average algorithm complexity of proposed Adaptive Kalman filter algorithm is $O(0.5N^2 + 0.4N)$ when $F_1 = 1$ and $F_2 = 1$, $O(1.4(5N^2 + 4N))$ when $F_2 = 1$ and $F_1 = 0$.

Compared to the standard Kalman filter algorithm, the adaptive Kalman filter algorithm retains the function of prediction value and reduces the number of operations in the trajectory tracking process. At the same time, the adaptive Kalman filter algorithm overcomes the unstable of gesture recognition which caused by lost frames and fingers temporary shelter in the process of hand movement.

The two cases $F_1 = 1, F_2 = 1$ and $F_1 = 0, F_2 = 1$ above have detailed introduced. When $F_1 = 1$ and $F_2 = 0$, the left hand is detected, but the right hand is not detected, at this case there is no processing. When $F_1 = 0$ and $F_2 = 0$, neither of the hands are detected. So we realize the complex gestures to switch naturally and without interference.

5.3 Experiment results

Combining the actual gesture recognition based on adaptive Kalman filter operation, we do the experiment for geological to enlarge, shrink, move up, move down, move to the left and to the right operation (Figure 4.1).

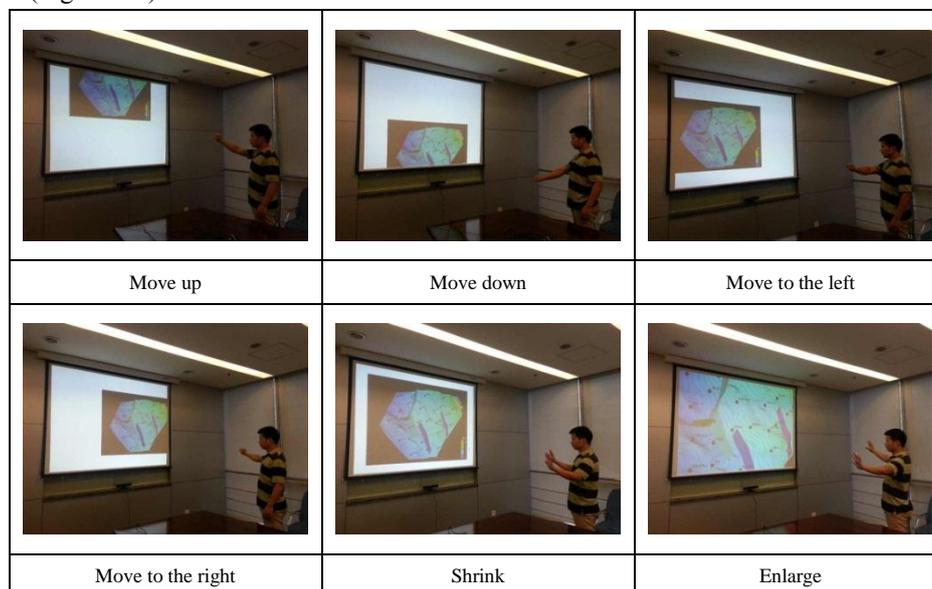


Figure 4.1 Gestures for browsing geological data

In the experiment, we record the trial result of 10 people. Each person used all the gestures to operate the geological data for 10 times, therefore, we obtained 100 times' data. The final experiment results are showed in table 4.1. The correct recognition rate of move up, move down, move to the left and move to the right is above 93%, but the operation of enlarge and shrink is less than 89%. The main reason is the palm posture based on multi-point depends on the angle between different fingers of the user's. When the palm is open and the angles are large, Kinect can be very good for the identification and distinction; when the palm of the hand in a closed state, the fingers angle is too small, less than the minimum resolution, which leads to inaccurate identification. The palms are more easily out of shape in the operation of enlarging and shrink, which result in the decline of identification rate.

Table 4.1 Experiment data statistics

Functions	Number of correct operation	Correct rate(%)
Enlarge	133	88.67
Shrink	132	88.00
Move to the left	140	93.33
Move to the right	142	96.67
Move up	145	96.67
Move down	147	98.0

6 Conclusion

Considering the geological workers' common gestures in practical manipulation of geological data, we put the standard Kalman filter algorithm into fingers identification and palm centre tracking.

The proposed adaptive Kalman filter method not only compensates the inadequacy of lost frames and temporary shelter, but also greatly improves the gesture recognition accuracy and tracking robustness., which is conducive to manipulate geological data.

For future work, hand gesture recognition method based on precision detection of angles between fingers will be well studied.

Reference:

- [1] Hu Shulong, Xu Jing. Improved design of core surface and profile image acquisition apparatus[J]. Journal of Jiangnan Petroleum Institute, 2003, 25(3): 127-128.
- [2] Wu Chonglong, He Wenzhen, Weng Zhengping, et al. Property, classification and key technologies of three-dimensional geological data visualization[J]. Geological bulletin of China, 2012, 30(5): 642-649.
- [3] Hu Shulong, Shen Jianghai. Development of a surface and profile core image acquisition unit[J]. Journal of Jiangnan Petroleum Institute, 2002, 24(4): 79-80.
- [4] Z. Ren, J. Meng, J. Yuan. Depth Camera Based Hand Gesture Recognition and its Applications in Human-Computer-Interaction[C]. Proceedings of the 8th International Conference on Information, Communications and Signal Processing (ICICS), 2011:1-5.dio:[10.1109/ICICS.2011.6173545](https://doi.org/10.1109/ICICS.2011.6173545)
- [5] Wu Xiaojuan, Zhang Boyang, Zhang Zhen, et al. A new threshold model for gesture recognition based on HMM[J]. Pattern Recognition and Artificial Intelligence, 2003, 16(1): 122-125.
- [6] Jiang Chao, Ai Jiaoyan. Dynamic gesture trajectory recognition and its application design for camera based on OpenCV[J]. Journal of Computer Applications, 2012, 32(S1): 128-133.
- [7] Zhang Boxiang, Wu Xiaojuan, Ge Qingguo, et al. Performance Research of Training Dynamic Gesture Track Based on HMM [J]. Singal Processing, 2004, 20(6): 662-666.
- [8] Kittasil Silanon, Nikom Suvonvorn. Hand motion analysis for Thai alphabet recognition using HMM[J]. International Journal of Information and Electronics Engineering, 2011, 1(1): 65-71.
- [9] Zhang Yi, Zhang Shuo, Luo Yuan, et al. Gesture Track Recognition Based on the Kinect Depth Image Information and It's Applications [J]. Application Research of Computers, 2012, 29(9): 3547-3550.
- [10] P. Trindade, J. Lobo, JP. Barreto. Hand gesture recognition using color and depth images enhanced with hand angular pose data[C]. 2012 IEEE International Conference on Multi sensor Fusion and Integration for Intelligent Systems, 2012:71-76. doi:[10.1109/MFI.2012.6343032](https://doi.org/10.1109/MFI.2012.6343032)
- [11] Jesus Suarez, Robin R. Murphy. Hand gesture recognition with depth images: A Review[C]. The 21st IEEE International Symposium on Robot and Human Interactive Communication, 2012:411-417. doi: [10.1109/ROMAN.2012.6343787](https://doi.org/10.1109/ROMAN.2012.6343787)
- [12] D. Ram´irez-Giraldo, S. Molina-Giraldo, Andr´esM.Alvarez-Meza. Kernel based hand gesture recognition using kinect sensor[C]. Image Signal Processing, and Artificial Vision, 2012: 158-161. doi: [10.1109/STSIVA.2012.6340575](https://doi.org/10.1109/STSIVA.2012.6340575)
- [13] V. Frati and D. Prattichizzo. Using kinect for hand tracking and rendering in wearable haptics[C]. 2011 IEEE World Haptics Conference (WHC), 2011: 317-321. dio:[10.1109/WHC.2011.5945505](https://doi.org/10.1109/WHC.2011.5945505)
- [14] Yu Hanchao, Liu Mingjie, Liu Junfa, et al. Hand gesture recognition method based on double-channel heterogeneous sensors fusion[J]. Computer Application and Software, 2012, 21(11): 1-3.
- [15] Yi Li. Hand gesture recognition using kinect[C]. 2012 IEEE 3rd International Conference on Software Engineering and Service Science (ICSESS), 2012: 196-199. doi: [10.1109/ICSESS.2012.6269439](https://doi.org/10.1109/ICSESS.2012.6269439)