International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Movement of Hand Motion Trajectory through Mid-Air using Stroke Technique

Tanaya Chowdhury[1], Dr. Sanjeev Kumar Sharma[2]

[1]Technocrats Institute of Technology (Adv)
tanayachowdhury18[at]gmail.com

[2]Technocrats Institute of Technology (Adv)
sanjeevsharma10091977[at]gmail.com

Abstract: ***In this paper, we propose a regular HCI mid-air based method based on trajectory motion. A key issue in distinguishing uninterrupted gestures is that performance of predictable algorithms is slower through this method, without knowing initial and destination of a gesture or differences in gesture interval. These issues become particularly difficult for those methods that rely on time-based information. For the improvement the issues of regular gesture recognition, we propose a outline that simultaneously performs both segmentation and 3D form of the motion with Depth Sensing Camera. The movement of the hand motion store in the frames through video and improve the performance and correctness of the gesture with large time-based deviation. Gesture trajectories perform a single stroke in the frame of hand movement in angular form. Dynamic trajectory segment varied in various frames with DTW to improve the Performance. For recognition, these trajectory segments are examined to determine whether the segment belongs to a class among intended gestures or a non-gesture class based on fusion of shape information and temporal features. In order to assess performance, the proposed algorithm was evaluated with a database of digit hand gestures. The experimental results indicate that the proposed algorithm has a high recognition rate while maintaining its performance in the presence of continuous gestures.***

Keywords: Hand Gesture, Trajectory, Feature Extraction, DTW, Machine Learning

1. Introduction

Human gestures [1] are expressive human body motions, which generally contain spatial and temporal variation. To handle these variations, an appropriate representation must be chosen. A vast amount of work in gesture recognition has been performed in the area of computer vision, and is reviewed in [2]. These works can be divided into two categories: trajectory-based and dynamics model-based analysis. The trajectory-based approach matches curves in configuration space to recognize gestures [3]. The dynamics model-based approach learns a parametric model of gestures. Gesture recognition systems in general are composed of three main [4] components: image preprocessing, tracking, and gesture recognition. In individual systems some of these components may be merged or missing, but their basic functionality [5] will normally be present. Image preprocessing is the task of preparing the video frames for further analysis by suppressing noise, extracting important clues about the position of the object of interest (for example hands) and bringing these on symbolic form. This step is often referred to as feature extraction. Tracking – on the basis of the preprocessing, the position and possibly other attributes of the object (hands) must be tracked from frame to frame. This is done to distinguish a moving object of interest from the background and other moving objects, and to extract motion information for recognition of dynamic gestures. Gesture recognition decides if the user is performing a meaningful gesture based on the collected position, motion and pose clues.

The classical algorithms from the field of pattern recognition are Hidden Markov Models (HMM), correlation, and Neural Networks. Especially the first two have been used successfully in gesture recognition while the Neural Networks often have the problem of modeling non- gestural patterns [6]. HMM is a typical dynamics model and was proven to be robust in its recognition of gestures [7]. The HMM model has been extended to a more general model named Dynamic Bayesian Networks [8]. Black and Jepson extended the Condensation algorithm [9], to recognize gestures and facial expressions in which human emotions were modeled [10] as temporal trajectories of some estimated parameters (which describe the states of a gesture or an expression) over time. Many gesture recognition methods used colored gloves or markers to track hand movements. Felsand Hinton used data gloves and Polhemus sensors to extract 3D hand location, velocity, and orientation [13]. Bobick and Wilson [14] extract 3-D location of hands using stereo cameras and skin color to recognize a set of 32 size gesture using parameterized hidden Markov Model and data glove.

2. Various Techniques

(a) Motion

Motion is a cue utilized by a few approaches to hand detection. The reason is that motion-based hand detection demands for a very controlled setup, since it assumes that the only motion in the image is due to hand movement. Indeed, early works assumed that hand motion is the only motion occurring in the imaged environment. In more recent approaches, motion information is combined with additional visual cues. In the case of static cameras, the problem of motion estimation reduces to that of background maintenance and subsequent subtraction. The difference in luminance of pixels from two successive images is close to zero for pixels of the background. By choosing and maintaining an appropriate threshold, moving objects are detected within a static scene.

(b) Tracking

Tracking, or the frame-to-frame correspondence of the segmented hand regions or features, is the second step in the process towards understanding the observed hand movements. The importance of robust tracking is twofold. First, it provides the inter-frame linking of hand/finger appearances, giving rise to trajectories of features in time. These trajectories convey essential information regarding the gesture and might be used either in a raw form (e.g., in certain control applications like virtual drawing the tracked hand trajectory directly guides the drawing operation) or after further analysis (e.g., recognition of a certain type of hand gesture). Second, in model-based methods, tracking also provides a way to maintain estimates of model parameters variables and features that are not directly observable at a certain moment in time.

3. Motivation

Challenges of Hand Gesture Recognition

There are many challenges associated with the accuracy and usefulness of gesture recognition software. For image-based gesture recognition there are limitations on the equipment used and image noise. Images or video may not be under consistent lighting, or in the same location. Items in the background or distinct features of the users may make recognition more difficult.

The variety of implementations for image-based gesture recognition may also cause issue for viability of the technology to general usage. For example, an algorithm calibrated for one camera may not work for a different camera. The amount of background noise also causes tracking and recognition difficulties, especially when occlusions (partial and full) occur. Furthermore, the distance from the camera, and the camera's resolution and quality, also cause variations in recognition accuracy.

In order to capture human gestures by visual sensors, robust computer vision methods are also required, for example for hand tracking and hand posture recognition or for capturing movements of the head, facial expressions or gaze direction.

4. Problem Formulation

Our purpose is to track the hand Gesture Trajectory for writing the character and numeric values and discriminate and composed of multiple quasi-linear segments (strokes). The algorithm must be able to distinguish the gestures based on the number of segments, the angles between segments and the horizontal axis, the angles between consecutive segments and segments proportionality. Using these parameters, it is possible to define a variety of gestures which can be easily discriminated: square, wide/tall rectangle, triangle, Z, N vertical/horizontal multiple strokes, cross etc. Each gesture may have a different meaning depending on the movement direction. For example, one action can, be associated with drawing a square clockwise, and another (possibly opposite) action can be associated with drawing a square counterclockwise. The Hue, Saturation, Value (HSV) color space is usually preferred [17] instead of the Red, Green, Blue (RGB) color space in human skin tracking applications because Hue is less sensitive to different skin colors and because it is more robust to illumination changes. As the overall computational cost must be low, complex tracking solutions like those based on particle filters are not appropriate, and a tracking algorithm based on target representation and localization must be used. Such a solution is offered by the CamShift (Continuously Adaptive Mean Shift) algorithm introduced by Bradski [17]. An implementation of this algorithm is available in the OpenCV library [19]. The CamShift Algorithm mainly consists of the following steps:

1) Choose the initial location of the search window;
2) Mean Shift (one or many iterations); store the zeroth moment;
3) Set the search window size equal to a function of the zeroth moment found in Step2;
4) Repeat Steps 2 and 3 until convergence.

The algorithm outputs the center, size and orientation of the tracked object. The object trajectory can be obtained from the positions of the center of the object in successive frames.

5. Problem Solution

The method we propose, presented in figure 1, uses a CamShift tracker to track the hand of the user and saves a trajectory obtained from the centers of the tracked region. The saved trajectory is then segmented into strokes. Considering all the gestures used are composed of a reduced number of strokes, information like number of strokes, average angles with horizontal axis, angles with neighboring segments and segments proportionality can be easily derived from the segmented trajectory. Finally, based on the extracted features the gesture is uniquely identified.

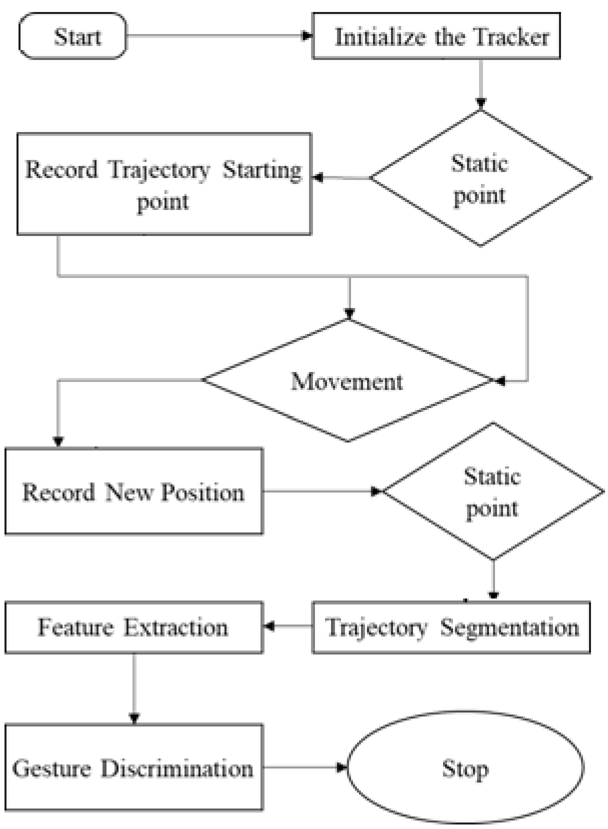

Figure: Flow of Hand Gesture

6. Conclusion & Future Work

Hand gesture segmentation and recognition involves several inherent difficulties; as it mentions no clear point that where a phase begins. So, different researchers may present different segmentations for the same input video. Also, there is a difficulty in identifying the rest and hold positions. For better learning of classifier and its performance, the gesticulation behavior must be recorded in different sessions. We developed a framework that addresses three associated problems. The experimentation and evaluation are done by spotting, segmentation and recognizing hand gestures in video.

The recognition of hand gestures is done using Deep Learning Networks after re-sampling the frames using a KNN based method and achieved better accuracy in comparison to other benchmark learning algorithms. It is shown that the gestures of interest that are embedded in video streams can easily be learned and recognized using re-sampling the frames. The performance of framework is evaluated on the basis of various metrics including F- score and categorical accuracy.

References

[1] S. Wang, H.Y. Shum, Y. Q. Xu and N. N. Zheng, Unsupervised Analysis of Human Gestures, IEEE Pacific Rim Conference on Multimedia, 2001, pp. 174-181

[2] Y. Wu, T. Huang, Vision-Based Gesture Recognition: A Review, Proceedings of the International Gesture Recognition Workshop, 1999, pp. 103-115.

[3] L. Campbell and A. Bobick.: Recognition of Human Body Motion Using Phase Space Constraints, Proceedings of the Fifth International Conference on Computer Vision, 1995, pp. 624- 630.

[4] T. Moeslund and L. Nrgaard, A Brief Overview of Hand Gestures used in Wearable Human Computer Interfaces, Technical Report CVMT 03-02, Computer Vision and Media Technology Laboratory, Aalborg University, DK, ISSN: 1601-3646, 2003.

[5] V.I.Pavlovic, R. Sharma, and T.S. Huang, Visual Interpretation of Hand Gestures for Human Computer Interaction: A Review, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, 1997, pp. 677–695.

[6] H.K.Lee and J.H. Kim, An HMM-based Threshold Model Approach for Gesture Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 10, 1999, pp. 961–972.

[7] T.Starner and A.Pentland.: Visual Recognition of American Language Using Hidden Markov Models, Proceedings of the International Workshop on Automatic Face and Gesture Recognition, Zurich, 1995, pp. 189-194.

[8] V. Pavlovic, J. M. Rehg and J. MacCormick, Impact of Dynamic Model learning on Classification of Human Motion, Proceedings of the International Conference on Computer Vision and Pattern Recognition, 2000, pp. 788-795.

[9] M. J.Black and A.D. Jepson, A Probabilistic Framework for Matching Temporal Trajectories: Condensation-Based Recognition of Gestures and Expressions, Proceedings of the Fifth European Conference on Computer Vision, 1998, pp. 909-924.

[10] M. H. Yang, N. Ahuja and M. Tabb, Extraction of 2D Motion Trajectories and Its Application to Hand Gesture Recognition, IEEE Transactions on pattern analysis and machine intelligence, Vol. 24, No. 8, 2002, pp. 1061- 1074.

[11] L. Nørgaard, Probabilistic Hand Tracking for Wearable Gesture Interfaces, Master's thesis, Laboratory of Computer Vision and Media Technology, Aalborg University, Denmark, 2003.

[12] J. Lin, Y. Wu, and T.S. Huang. Capturing human hand motion in image sequences. Proceedings of IEEE Workshop on Motion and Video Computing, 2002, pp. 99-104.

[13] S.S. Fels and G.E. Hinton, Glove-Talk II: A Neural Network Interface which Maps Gestures to Parallel Format Speech Synthesizer Controls, IEEE Transactions on Neural Networks, Vol. 9, No. 1, 1997, pp. 205-212.

[14] A.D. Wilson and A.F. Bobick, Parametric Hidden Markov Models for Gesture Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 9, 1999, pp. 884-900.

[15] S. Lenman, L. Bretzner and B. Thuresson, Computer Vision Based Recognition of Hand Gestures for Human-Computer Interaction, Technical report TRITANA -D0209, CID- 172, ISSN 1403-0721, Department of Numerical Analysis and Computer Science, 2002.

[16] M. J. Jones and J. M. Rehg, Statistical Color Models with Application to Skin Detection, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1999, 274-280.

[17] G. R. Bradski, Computer vision face tracking as a component of a perceptual user interface, Intel Technology Journal Q2 1998, http://developer.intel.com/technology/itj/archive/1998.htm.

[18] D. Comaniciu, V. Ramesh, P. Meer, Kernel-Based Object Tracking, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 5, 2003. http://opencvlibrary.sourceforge.net.

[19] L. Bretzner, I. Laptev and T. Lindeberg, Hand gesture recognition using multi-scale colourfeatures, hierarchical models and particle filtering, Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002, pp 405-410.

[20] J.H. Shin, J.S. Lee, S.K. Kil, D.F. Shen, J.G Ryu, E.H. Lee, H.K. Min and S.H. Hong, Hand Region Extraction and Gesture Recognition using entropy analysis, IJCSNS International Journal of Computer Science and Network Security, Vol. 6 No. 2, 2006, pp. 216-222.