PERFORMANCE EVOLUTION OF EYE AND HAND FUSION FOR DIAGONAL MOVEMENT GESTURE RECOGNITION. (A Review)

[1]Yogesh .C. Pathak, [2]S.A. Chhabria, [3]Rajiv V. Dharaskar

[1]M –Tech, [1]HOD OF IT DEPT., G.H.R.C.E. NAGPUR
[3]Director, MPGI Integrated Campus, Matoshri Pratishthan Group of Institute, Nanded
Email : yogeshpathak00712@rediffmail.com

Abstract:- Human–computer Interaction (HCI) involves the study, planning, and design of the interaction between people (users) and computers. It is often regarded as the intersection of computer science, behavioral, design and several other fields of study. HCI plays a very important role in interaction with the computers. This will be the new era in computation.

The project will compare different fusion techniques and use optimized fusion technique for fusing eye and hand gestures. It will be working on the movements towards diagonally right upwards, diagonally left upwards, diagonally right downwards, and diagonally left downwards.

The mechanism of the project is to make performance evolution of the fusion techniques; these fusion techniques will evaluate the ongoing performance using the testing phases and will show the results of fusion techniques for eye and hand gestures.

INTRODUCTION:

Human–computer Interaction (HCI) involves the study, planning, and design of the interaction between people (users) and computers. It is often regarded as the intersection of computer science, behavioral, design and several other fields of study. HCI plays a very important role in interaction with the computers. This will be the new era in computation.

The project will compare different fusion techniques and use optimized fusion technique for fusing eye and hand gestures. It will be working on the movements towards diagonally right upwards, diagonally left upwards, diagonally right downwards, and diagonally left downwards.

The mechanism of the project is to make performance evolution of the fusion techniques; these fusion techniques will evaluate the ongoing performance using the testing phases and will show the results of fusion techniques for eye and hand gestures.

PROBLEM DEFINITION:

AS THE PROJECT IS PERFORMANCE EVOLUTION OF FUSION TECHNOLOGY FOR EYE AND HAND DIAGONAL MOVEMENT GESTURE RECOGNITION.

The problem definition is to evaluate the performance of the fusion technique for eye and hand diagonal movements and to see which technique is the best suited related with the speed, accuracy, performance etc.

The mechanism of the project is to make performance evolution of the fusion techniques; these fusion techniques will evaluate the ongoing performance using the testing phases and will show the results of fusion techniques for eye and hand gestures.

LITERATURE SURVEY:

FOR FUSION MULTIMODAL INPUT FUSION IN HUMAN- COMPUTER INTERACTION

In this paper, we address the modality integration issue

on the example of a system that aims at enabling users to combine their speech and 2D gestures when interacting with life-like characters in an educative game context. In a preliminary limited fashion, we investigate and present the use of combined input speech, 2D gesture and environment entities for user system interaction.[6]

Multimodal human–computer interaction: A survey

In this paper, we review the major approaches to multimodal human–computer interaction, giving an overview of the field from a computer vision perspective. In particular, we focus on body, gesture, gaze, and affective interaction (facial expression recognition and emotion in audio). We discuss user and task modeling, and multimodal fusion, highlighting challenges, open issues, and emerging applications for multimodal human–computer interaction (MMHCI) research.[5]

FUSION

Fusion techniques are needed to integrate input from different modalities and many fusion approaches have been developed. Early multimodal interfaces were based on a specific control structure for multimodal fusion. For example, Bolt's ''Put-That-There'' system [3] combined pointing and speech inputs and searched for a synchronized gestural act that designates the spoken referent. To support more broadly functional multimodal systems, general processing architectures have been developed which handle a variety of multimodal integration patterns and support joint processing of modalities.

INTERMEDIATE FUSION TECHNIQUES

Since the early fusion techniques avoid explicit modeling of the different modalities, they fail to model both the fluctuations in the relative reliability and the asynchrony problems between the distinct (e.g., audio and visual) streams.

Moreover, a multimodal system should be able to deal with imperfect data and generate its conclusion so that the certainty associated with it varies in accordance to the input data. A way of achieving this is to consider the time instance versus time-scale dimension of human non-verbal communicative signals [11]. By considering previously observed data with respect to the current data carried by functioning observation channels, a statistical prediction and its probability might be derived about both the information that has been lost due to malfunctioning/inaccuracy of a particular sensor and the currently displayed action/reaction. Probabilistic graphical models, such as Hidden Markov Models (including their hierarchical variants),

Bayesian networks and Dynamic Bayesian networks are very well suited for fusing such different sources of information [1]. These models can handle noisy features, temporal information, and missing values of features by probabilistic inference. Hierarchical HMM-based systems [4] have been shown to work well for facial expression recognition. Dynamic Bayesian Networks and HMM variants [5] have been shown to fuse various sources of information in recognizing user intent, office activity, and event detection in video using both audio and visual information [5]. This suggests that probabilistic graphical models are a promising approach to fusing realistic (noisy) audio and video for context-dependent detection of behavioral events such as affective states.

"A performance evaluation of fusion techniques for spatio-temporal Saliency detection in dynamic scenes" In this paper a performance evaluation of fusion techniques for spatio-temporal saliency detection in dynamic scenes is presented. The nine fusion techniques are evaluated on a large dataset of twelve complex dynamic scenes. The results show the consistency of fusion approaches that base decision on the scene's characteristics as the final spatio-temporal saliency map takes the best of each individual saliency map (static and dynamic). This includes Mean, Scale Invariant, Max and Dynamic Weights fusion methods. On the other hand, fusion techniques which are based on a strong a priori such as Motion Priority fusion achieve good results only when the underlying assumption is satisfied. Thus, they performances vary depending on the sequence. It is clear that the accuracy of a spatio-temporal saliency map depends on the quality of both static and dynamic maps, which are based on the scene's contents. Therefore, it would be useful to derive the weights (fusion technique) based on the images contents.

FOR EYE DETECTION

A simple and efficient eye detection method in color

Images D. Sidibe, P. Montesinos, S. Janaqi

In this paper we propose a simple and efficient eye detection method for face detection tasks in color images. The algorithm first detects face regions in the image using a skin color model in the normalized RGB color space. Then, eye candidates are extracted within these face regions. Finally, using the anthropological characteristics of human eyes, the pairs of eye regions are selected. The proposed method is simple and fast since it needs no template matching step for face verification. It is robust because it can deal with rotation. Experimental results clearly show the validity of our approach. A correct eye detection rate of 98.4% is achieved using a subset of the AR face database.[9]

A Comparison of Face Detection Algorithms in Visible and Thermal Spectrums

Face Detection is the first step of facial recognition algorithms and has been widely researched in the visible spectrum. Current research has shown that thermal facial recognition is as accurate as the visible spectrum recognition algorithms. This paper presents three face detection algorithms in both long-wavelength infrared

(LWIR) images and visible spectrum images. The paper compares the Viola-Jones algorithm, Gabor feature extraction and classification using support vector machines, and a Projection Profile Analysis algorithm. The Gabor feature extraction method can detect faces in both spectrums with separate training, but the algorithm is extremely slow. The Project Profile Analysis method can find faces in LWIR images, but is not applicable to visible spectrum images. Our experimental results show that the Viola- Jones algorithm is the most reliable and efficient solution for the implementation of a real-time face detection system using either visible or thermal spectrum images.[4]

FOR HAND GESTURE DETECTION

A Review of Various Hand Gesture Recgnition Techniques Ginu Thomas

This paper presents simple as well as effective methods to realize hand gesture recognition. The methods used are pixel by pixel comparison, edges method and use of orientation histogram. First an image database is created which constitutes various static hand gesture images. These images are a subset of American Sign Language (ASL). Filtering of the image is done so as to reduce the amount of noise present in the image. Segmentation of image is done in order to make it easier to analyze. A pattern recognition system is used to transform an image into feature vector, which will then be compared with the trained set of gestures. The method used was successful to retrieve the correct match.[9]

A Stable Hand Tracking Method by Skin Color Blob Matching Jung-Ho Ahn and Jong-Hyoun Kim

Hand detection and tracking is one of the main research areas in computer vision for human computer interaction. But many research results are not wholly satisfactory for practical purposes. In this paper we propose a fast and stable hand detection and tracking method with human body model. We detected hand areas by combining the information of the difference image and skin colour area, and reconstructed hand shapes accurately. For hand tracking we suggest a skin colour blob matching method with some tracking rules. The experimental results show that the proposed algorithm performs well in real time.[8]

Vision-Based Finger Action Recognition by Angle Detection and Contour Analysis Daeho Lee and SeungGwan Lee In this paper, we present a novel vision-based method of recognizing finger actions for use in electronic appliance interfaces. Human skin is first detected by colour and consecutive motion information.

Then, fingertips are detected by a novel scale-invariant angle detection based on a variable k-cosine. Fingertip tracking is implemented by detected region-based tracking. By analyzing the contour of the tracked fingertip, fingertip parameters, such as position, thickness, and direction, are calculated. Finger actions, such as moving, clicking, and pointing, are recognized by analyzing these fingertip parameters. Experimental results show that the proposed angle detection can correctly detect fingertips, and that the recognized actions can be used for the interface with electronic appliances. [11]

We apply an appearance-based technique based on the work by Viola and Jones [1] for face detection. To improve the robustness of the face detector, we have developed a novel, linear-time mechanism to make the system invariant to variable lighting conditions: all features are scaled by a ratio of the average gray-level intensity of the training samples over the average gray-level intensity of the current search region. We use the integral image to efficiently compute feature block intensity levels. The best Haar-like features are selected with Adaboost [1]

Once the face is detected, we also perform eye detection using the so-called topographic features of eyes presented in our previous work [12]. The basic idea is to create a terrain map from the gray-scale image, effectively treating it like a continuous 3-D surface, and extract “pit” locations as pupil candidates. The actual eye pair is chosen using a Gaussian mixture model (GMM) and certain heuristics. The approximate centre of the two detected eye candidates is used as the focus of the camera. Note that we do not use the mutual information tracking approach presented in [12] and instead detect the eyes once every time the face is detected in the wide-angle view.

Hand region detection is the first step towards estimating the pointing direction. Motivated by the success of the face detection approach developed by Viola-Jones [13] using Haar-like features and an AdaBoost cascade detector, we extend the features

to hand region detection for the pointing gesture.

APPLICATION

- It could be used by paralyzed persons i.e. to move the wheel chair, or to operate the specific devices or even full computer operations.
- Could be further used by the patients or handicap persons to play certain games.

Future Scope: -

By developing this project it could be in future be implemented to help a disabled person or any handy cap person to move his or her wheel chair or even could be used to play certain games . This project will provide a different touch to the world of human computer interaction.

EXPECTED RESULT

Performance evolution of fusion techniques and fusion of eye and hand gestures like movements towards diagonally right upwards, diagonally left upwards,

diagonally right downwards, and diagonally left downwards.

REFERENCES:

[1] A Multi-Gesture Interaction System Using a 3-D Iris Disk Model for Gaze Estimation and an Active Appearance Model for 3-D Hand Pointing , by Michael J. Reale, Student Member, IEEE, Shaun Canavan, Student Member, IEEE, Lijun Yin, Senior Member, IEEE, Kaoning Hu, and Terry Hung IEEE TRANSACTIONS ON MULTIMEDIA, VOL. 13, NO. 3, JUNE 2011.

[2] Combining Eye Gaze and Hand Tracking for Pointer Control in HCI, Developing a more robust and accurate interaction system for pointer positioning and clicking, by Ngip-Khean Chuan, Ashok Sivaji Product Quality & Reliability Engineering (PQRE) MIMOS Berhad Kuala Lumpur, Malaysia ,nk.chuan@mimos.my, ashok.sivaji@mimos.my , 2012 IEEE Colloquium on Humanities, Science & Engineering Research (CHUSER 2012), December 3-4, 2012, Kota Kinabalu, Sabah, Malaysia.

[3] A Low-Computational Approach on Gaze Estimation With Eye Touch System , by Cihan Topal, Serkan Gunal, Onur Koc ¸deviren, Atakan Dog˘an, and O¨ mer Nezih Gerek , IEEE TRANSACTIONS ON CYBERNETICS 2013 IEEE.

[4] An Automated Hirschberg Test for Infants by Dmitri Model*, Student Member, IEEE, and Moshe Eizenman, IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 58, NO. 1, JANUARY 2011.

[5] Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques by Nasser H. Dardas and Nicolas D. Georganas, Fellow, IEEE. IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, VOL. 60, NO. 11, NOVEMBER 2011.

[6] Face and Eye Detection by CNN Algorithms DAVID BALYA AND TAMA´S ROSKA Analogical and Neural Computing Laboratory, Computer and Automation Research Institute, L´agym´anyosi u. 11, Budapest, H-1111, Hungary 2012.

[7] A Review of Various Hand Gesture Recognition Techniques 1Ginu Thomas VSRD-IJEECE, Vol. 1 (7), 2011, 374-383.

[8] A Stable Hand Tracking Method by Skin Color Blob Matching Jung-Ho Ahn and Jong-Hyoun Kim Pacific Science Review, vol. 12, no. 2, 2011,

[9] A Comparison of Face Detection Algorithms in Visible and Thermal Spectrums Kristopher Reese1, Yufeng Zheng2, Adel Elmaghraby1 BIOINFO Medical Imaging Volume 1, Issue 1, 2011,

[10] A simple and efficient eye detection method in color Images D. Sidibe, P. Montesinos, S. Janaqi LGI2P/EMA - Ales School of Mines Parc scientifique G. Besse, 30035 Nimes Cedex 1, France Desire. Sidibe, Philippe. Montesinos, stefan.Janaqi}@ema. 2006

[11] Vision-Based Finger Action Recognition by Angle Detection and Contour Analysis by Daeho Lee and SeungGwan Lee. 2010

[12] "Using geometric properties of topographic manifold to detect and track eyes for human-computer interaction," by J. Wang, L. Yin, and J. Moore, ACM Trans. Multimedia Computer. Commun. Appl., vol. 3, pp. 3:1–3:20, Dec. 2007.

[13] Robust real-time face detection by P. Viola and M. J. Jones Int. J. Comput. Vis., vol. 57, pp. 137–154, May 2004.