ISSN:0975-9646

Krishna Chauhan et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (4) , 2014, 5526-5528

Natural Feature Tracking (NFT) Enabled Low Cost Human Machine Interface (HMI) for Real Time Applications in Mobile

Prof. Krishna Chauhan
Sagar Institute of Reasearch and Technology
Rajiv Gandhi Technical University
Bhopal, Madhya Pradesh
India

Sandeep Singh Padan
Sagar Institute of Reasearch and Technology
Rajiv Gandhi Technical University
Bhopal, Madhya Pradesh
India

Abstract: **Use of Natural Feature Tracking (NFT) is imperative in order to avoid the use of any artificial parameter and to implement a low cost Human Machine Interface (HMI) for real time applications in mobile. However, accurate and precise tracking coupled with desired speed to obtain a fruitful output proves to be a technically expensive matter and sometimes this modern technology of man machine interaction is also away from the reach of academic study of the student's community owning to monetary expense over computational complexities. To cope with this basic hindrance, we propose a novel technique which not only reduces the amount of computational cost incurred but also stabilizes camera using light in-flow enabling the mobile for real time applications**

I INTRODUCTION

Natural features constitute parameters from unbiased scenes or platforms. Extraction of heavy features and complex procedures of matching has been a characteristic feature of most of the previous NFT methods [1]. Such practices require heavy computation and proves to be in efficient while its implementation on low performance devices for low cost analysis. Therefore to speed up this process, we propose a new tracking methodology based on flow of light from the feature itself. We implemented this proposed procedure on mobile devices in real time scenarios so that it can be implemented for AR applications. In addition to this, by replacing new feature points proportionally by the number of diminished/vanished ones we kept the number of feature points constant. The underlining principle is to restrict the extraction area of new features to small portioned sub regions and to recall the feature extraction matching loops less. The moment a potion of the input is matched to a predefined pattern; successive video frames are used to initiate the tracking. No further tracking is performed as long as desired results are being obtained at the output.

This paper introduces an approach to implement Augmented Reality tracking by using only natural features. It also highlights the unique combination of components and tools of recognition technology. Here, the idea of integration uncalibrated natural features with other artifacts at the scene in an Augmented Reality tracking system in mobile is unique.

II TRACKING NATURAL FEATURES

A well known method of tracking natural features is the feature tracker developed by Lucas and Kanade [2]. Let *A* and *B* be two consecutive grayscale images, and their scalar quantities $A(x, y)$ and $B(x, y)$ be the pixel intensities at image coordinates (x, y). The consecutive image frame *B* contains most parts of the first image frame *A*. The position (x, y) in *A* will be moved to a new position $(x + dx, y + dy)$ in *B*. The tracker must determine the disparity vector (dx, dy) at (x, y) from the intensity similarity of *A* and *B*. The similarity criterion is measured from the set of local neighborhood pixels, denoted by *W*, centered at the position (x, y). The disparity is commonly computed by minimizing the residual error due to the brightness differences [3]. An interest point detector and descriptor called Speeded Up Robust Features (SURF) to reduce the time for feature tracking have already been proposed by approximating second order Gaussian derivatives in the Hessian-Laplace detector with box filters, which can be evaluated very fast using integral images. SURF is currently regarded as a potentially more efficient descriptor then the previous descriptors such as SIFT and Fern [4]. Tracking features is an essential step before processing three dimension motions because AR applications severely restrict service domain in order to enhance speed. As a result of which previous studies have concentrated mainly on description and feature matching between two consecutive images. Their methods extract a new set of point features from each of the newly appeared image, instead of considering previous matched features. The extraction and matching of a new point set is time-consuming and should be avoided especially when the method is used for a real-time application.

We propose to utilize the matching features to speed up the process of tracking. In vision-based AR applications the purpose of NFT is to compute a homography between a planar scene and a projected image [5]. To ensure the existence of a pattern, there must be a large number of feature points for the planar scene pattern and also enough number of the feature points must be matched to points in the projected image. Conventional feature tracking techniques determine correspondences between two consecutive images, namely A_t and A_{t+1}. They extract an initial set of feature points from the initial frame and track their movements along the consecutive on coming image frames. This paper is the outcome of an attempt to utilize the previous matched features to considerably bring down the economical/computational expenses and development complexities of tracking process.

III PRE REQUISITES & SETUP

Visual Studio 2005 or 2008, or Visual Studio Express Edition, .NET 3.0 or greater and "WebcamLib.dll" are essential as software prerequisites. A standard webcam with 640x480 resolutions was used with an Android 4.2.2 Jelly Bean Mobile with 2 Mega Pixel Camera. The resolution proved more than needed for analyzing gestures of one hand, so frames were captured at half resolution (320x240) to improve processing speed and to make sure that our approach is able to handle low-resolution images.

Problems encountered: The windowed room had more noise issues. This was not surprising because of the widely varying lighting. It was generally acceptable but very difficult to get a clean image out when the lighting was exceptionally bad. In addition, incandescent light bulbs used in the room produced yellowish light, which made hue-based extraction more difficult.

IV BACKGROUND SUBTRACTION

Background subtraction is the first approach to extracting a finger that was attempted. It is a very simple approach that is highly susceptible to errors from noise, so the first fifty frames of the video feed were accumulated as an entire background sequence to improve robustness over using a single frame. Often a scene will have objects moving through it that are not supposed to be part of the background. By taking the mean of many frames, these objects can be averaged out and only the static background will remain. The mean and variance of each pixel through this initial sequence is calculated and used to find the Mahalanobis distance between the sequence and any new frame. If the distance exceeds a threshold then the new pixel is sufficiently different from the background and is turned on in the final binary finger image.
Problems Encountered: While good images were occasionally possible, generally the shadows overwhelmed the image. This essentially caused two overlapping hands to be extracted, which resulted in a blob that could not easily be analyzed.

V NEW POINT POSITION ESTIMATION, TIP CLASSIFICATION AND PHOTO RESIZING

Two different recursive filters have been designed and tested to estimate 3D positions of new points based on the camera pose and measurements of the feature image coordinates. The results of synthetic and real data showed that both filters converged and were stable. The EKF (Extended Kalman Filter) is known to have good characteristics under certain conditions, however the RAC (Recursive Average of Co-variances) filter gives comparable results, and it is simpler, operating completely in 3D-world space with 3D lines as measurements[6]. The RAC filter approach eliminates the linearization processes required in the EKF with Jacobian matrices. As the algorithm depends upon curvature calculation from the contour, thus, although the algorithm is immune to rotation, any single peak will always be classified as a thumb. Also, in case the left hand is used the system will treat it like a right hand. In other words the system has no way to determine whether hand is a left or right hand and to distinguish between the fingers. However, the algorithm works very well when there is only one tip in the image. In case the user shows two fingers, disturbance may occur. Also any single finger tip is detected as the thumb and the image is resized according to the motion of that finger. Again if the finger moves too close to any of the corners then the false tips may result. Tracking of an identified tip is an issue. Another constraint is that the hand should not move very swiftly.

VI RESEARCH OUTPUTS AND RESULTS

Fig 1: NFT enabled HMI using Mobile Phone
(Assembly No. 1)
Here the Golden Colored patterns on Purple Background Wall are getting recognized but the extra light coming from the window is acting as an unwanted external noise introduced deliberate into the scene to study its effects on the process of tracking
(Location: My Home/study room)

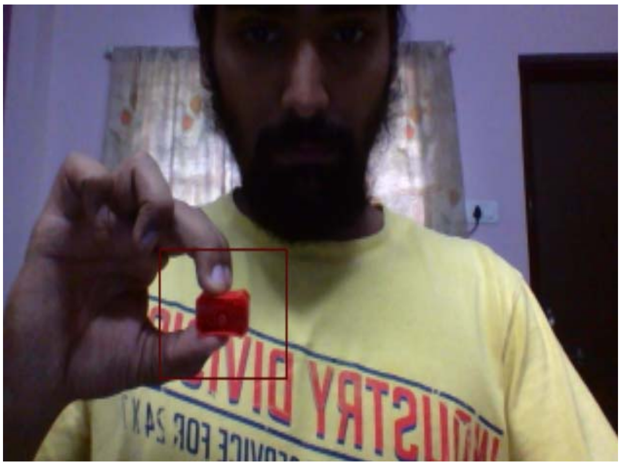

Fig 2: NFT enabled HMI using Mobile Phone. (Assembly No. 2)
Here I have used the yellow background of my T-shirt along with two different sources of light viz. light from the window and light from the tube light in the room. Here my fingures are in their natural form at the main portion of tracking DARK RED over LIGHT RED
(Location: My Home/study room)

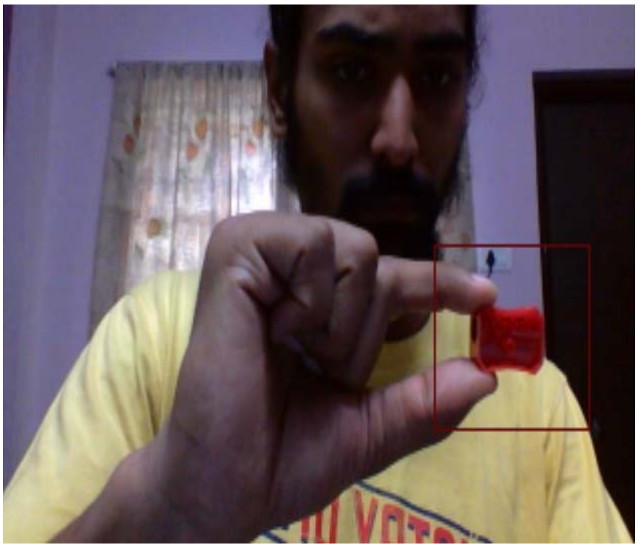

Fig 3: NFT enabled HMI using Mobile Phone.
(Assembly No. 3)
Here there are two light sources viz. light from the window and light from the tube light in the room. As compared to the earlier two assemblies the background is multi colored. However, the desired output is obtained (Location: My Home/study room)

VII RESEARCH OUTPUTS AND RESULTS

Table 1: Analysis of Processing and tracking time (in ms)

MTHD	AM	PROCESS	TRACK	TOTAL
AR toolKit	1	210.389	214.486	424.875
	2	235.486	231.598	467.084
	3	289.255	#	NA
SURF	1	203.524	2.598	206.122
	2	235.412	36.745	272.157
	3	298.252	107.336	405.588
PRO-POSED	1	3.919	24.486	28.405
	2	4.015	25.862	29.877
	3	5.326	25.685	31.011

Could not track; MTHD: Method; AM: Assembly;

Calculations:

Mean $_{(proposed)}$ = (28.405+29.877+31.011)/3 = 89.293/3

= 29.76

Deviation from mean $_{(proposed)}$

1. 28.405-29.76 = -1.355
2. 29.877-29.76 = +0.117
3. 31.011-29.76 = +1.251

VII CONCLUSIONS

Experimental results indicate that the method proposed by us significantly reduces the computational cost and also stabilizes the camera pose estimation process. While tracking, we analyzed the homography between the input frame and the matched pattern using the matched pairs. The accuracy is measured by the sum of differences of the evaluated corner positions from the homography and the actual corner positions which are manually specified. In the SURF method there are abrupt increases of error in some frames. In the proposed method the error is stable throughout the input stream. We compared the processing time on the three platforms between ARtoolKit, SURF and our proposed method. Table 1 shows the results of comparison. The feature extraction step takes more than half of the total processing time.
In the proposed method, the feature extraction and matching time is significantly reduced.

VIII FUTURE WORK

Noise remained an issue throughout testing, so finding a better way to extract the parameter and speed up the process of tracking would be very useful. Specially, a histogram-based HSV extraction could be tried. The classifier lacked precision, so improving peak and valley detection is also a goal. A larger Hu moment database could also be interesting because current testing has been done with very limited samples. Also, work was mainly done in a moderately constrained setting. Trying to deal with less ideal settings, such as one that could result in partially occluded gestures, could be very interesting.

ACKNOWLEDGEMENT

I gratefully acknowledge the contribution of Prof. Krishna Chauhan

REFERENCES

[1]. Lim, M.J., Jung, H.W., Lee, K.Y.: *Game-type recognition rehabilitation system based on augmented reality through object understanding.* The Journal of the Institute of Webcasting, Internet and Telecommunication 11(3) (2011) 93–98

[2]. Lucas, B., Kanade, T.: *An iterative image registration technique with an application to stereo vision.* In: Proceedings of the 7th International Joint Conference on Artificial Intelligence. (1981) 674–679

[3]. Krishna Chauhan, Sandeep Singh Padan: *"Implementation of low cost HMI using NFT"*, IJCA (0975-8887), Volume 86-No 17, Jaunary 2014

[4]. [CAUD92] T. P. Caudell, D. M. Mizell, *"Augmented Reality: An Application of Heads-Up Display Technology to Manual Manufacturing Processes,"* Proceedings of the Hawaii International Conference on Systems Sciences, pp. 659-669, 1992

[5]. FEIN93] S. Feiner, B. MacIntyre, D. Seligmann, *"Knowledge-Based Augmented Reality,"* Communications of the ACM, Vol. 36, No. 7, pp 52-62, July 1993

[6]. Lim, M.J., Jung, H.W., Lee, K.Y.: *Game-type recognition rehabilitation system based on augmented reality through object understanding.* The Journal of the Institute of Webcasting, Internet and Telecommunication 11(3) (2011) 93–98