

HGRIHCI: Hand Gesture Recognition for Intuitive Human Computer Interaction

Arindam Sarkar
Department of Computer Sc. &
Engineering, NSEC,
Kolkata-700152

Shilpi Bose
Department of Computer Sc. &
Engineering, NSEC,
Kolkata-700152

Chandra Das
Department of IT, NSEC,
Kolkata-700152

ABSTRACT

Gesture Recognition using Computer Vision opens up a whole new frontier in Human-Computer Interaction. Hand gestures are natural, intuitive, and require almost no learning, or remembering whatsoever. The proposed work involves developing a system to translate gestures (predetermined) performed by the user to control the active application on a computer. Viola-Jones Cascade Object Detector with Histogram of Orientation Gradients (HOG) features is used to detect Hand candidates (Open Palm or Fist). In this paper several heuristics are used to filter out false-positives. Drag patterns made by user, are then interpreted as one of the gestures in the database, using a simple pixel-wise distance based algorithm to match patterns. This approach results to robustness, and higher invariance to illumination changes, as compared to earlier works, which depend purely on known distribution of human skin-color.

Keywords

Gesture Recognition; Computer Vision; histogram; segmentation; open-palm; Fist.

1. INTRODUCTION

According to a famous and famously misused rule, communication is only 7 percent verbal and 93 percent non-verbal. The non-verbal component is made up of body language (55 percent) and tone of voice (38 percent). But that is in the world of humans. Along the million years of our existence, we have developed a quite sophisticated system of communication. However, since the advent of computer systems, we have been striving to make them more powerful, easy to use, and human-like. A part of that effort has gone into creating more intuitive ways of interacting with computers, past the mice and keyboards. Touch-Screens have taken us a step closer. There are already quite well developed Speech-Recognition systems. They have become better with time, and are part of many systems now. In the past decade, a lot of work has gone into developing computer vision based systems for interacting with computers. A system which can understand well known human-gestures, and perform suitable actions, could be used in 3D Visualization, TV/desktop-control, Robotics, Gaming, Sign-Language Recognition, Video Surveillance etc. [1].

Gestures could be Static or Dynamic. For example, if one wants to control a Media Player, he could perform a STOP like the Traffic Police-man does to stop the currently playing track, or swipe open hand to left, to go to previous track. Thus what a Gesture Recognition system should be able to do, is to capture the user's gesture with web-cam or other suitable video-capture device, and convert it to an intended action, which would be some system command to control the Computer or TV or some other Remote Devices.

The Objective of this paper is to develop a system to control mouse via hand gestures, and to recognize patterns of mouse-drags performed by the user. So, the aim is to detect two classes of objects in image – open-palm facing webcam, and Fist. The target environment is well-illuminated, and stable. Also, the user must not be wearing gloves. The user can be at a distance of 1-4 ft. from the webcam. Web-cam needs to support at least 320x240 resolution of image capturing capability in RGB. Also, better quality webcam will result into better detection. User can perform one of the gestures by dragging Fist in front of webcam: Scroll (up/down) by moving fist up/down, Press Left/Right Arrow Keys for navigation by moving fist left/right, Reload/Refresh by making a circle-like pattern, Close current application by making a Cross (X) like pattern.

Here, Viola-Jones Object Detector [2] with Histogram of Orientation Gradients (HOG) [3] features is used for detecting Hand (Open Palm/Fist) in present Image Frame. Object Detector for every frame, in addition to correct hand candidate, returns some non-hand objects as well. To remove these, heuristics based on skin-color range (Explicit Skin Color-Space Thresholding) in YCbCr color space [4], and position of hand-candidates are used. Open palm is used as cursor controller, and fist as single mouse click. When user moves fist around, a drag pattern is created, this pattern is matched against the predetermined patterns, using aspect ratio and nearest distance based matching.

2. RELATED WORK

Hand Gesture Recognition has a vast literature of its own, which shows how challenging accurate, real-time hand-gesture recognition is. There are many approaches that were followed by different researchers like vision based, data glove based using Artificial Neural Network, motion based gesture recognition utilizing GPU (Graphics Processing Unit), Hidden Markov Model, Support Vector Machines etc.

Many researchers used vision based approaches for identifying hand gestures. Kapuscinski [5] found out the skin colored region from the input image captured and then this image with desired hand region was intensity normalized and histogram was found out for the same. Feature extraction step was performed using Hit-Miss Transform and the gesture was recognized using Hidden Markov Model (HMM). Recognition rate obtained was 98%. Yu [6] used YCbCr color model to distinguish skin colored pixels from the background. The required portion of the hand was extracted using this color model and filtered using median filter and smoothing filter. The edges were detected and features extracted were hand perimeter, aspect ratio, hand area after which Artificial Neural Network (ANN) was used as classifier to recognize a gesture. Accuracy rate obtained was 97.4%.

Most of the complete hand interactive systems can be considered to be comprised of three layers: detection, tracking and recognition. The detection layer is responsible for defining and extracting visual features that can be attributed to the presence of hands in the field of view of the camera(s). The tracking layer is responsible for performing temporal data association between successive image frames, so that, at each moment in time, the system may be aware of "what is where". Moreover, in model-based methods, tracking also provides a way of maintaining estimates of model parameters, variables and features that are not directly observable at a certain moment in time. Lastly, the recognition layer is responsible for grouping the spatiotemporal data extracted in the previous layers and assigning the resulting groups with labels associated to particular classes of gestures. In this section, research on these three identified sub problems of vision-based gesture recognition is reviewed.

The primary step in gesture recognition systems is the detection of hands and the segmentation of the corresponding image regions. This segmentation is crucial because it isolates the task-relevant data from the image background, before passing them to the subsequent tracking and recognition stages. In [7], a comparative study on the performance of some hand segmentation techniques can be found. In skin color segmentation to increase invariance against illumination variability some methods [8], [9], [10], [11], [12], [13] operate in the HSV [14], YCrCb [15], or YUV [16], [17] color-spaces, in order to approximate the "chromaticity" of skin (or, in essence, its absorption spectrum) rather than its apparent color value. For detection of motion in hand, several approaches have been utilized. Early works (e.g. [18], [19], [20]) 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. For example in [21], [22] such information is utilized to distinguish hands from other skin-colored objects and cope with lighting conditions imposed by colored lights. 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.

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. In [23], [24] correlation-based template matching is utilized to track hand features across frames. Once the hand(s) have been detected in a frame, the image regions in which they appear is utilized as the prototype to detect the hand in the next frame. Again, the assumption is that hands will appear in the same spatial neighborhood. This technique is employed for a static camera in [25], to obtain characteristic patterns (or "signatures") of gestures, as seen from a particular view. The work in [26] deals also with variable illumination. A target is viewed under various lighting conditions. Then, a set of basis images that can be used to approximate the appearance of the object viewed under various illumination conditions is constructed. Tracking simultaneously solves for the affine motion of the object and the illumination.

The overall goal of hand gesture recognition is the interpretation of the semantics that the hand(s) location, posture, or gesture conveys. Template matching, a fundamental pattern recognition technique, has been utilized in the context of both posture and gesture recognition. In the

context of images, template matching is performed by the pixel-by-pixel comparison of a prototype and a candidate image. The similarity of the candidate to the prototype is proportional to the total score on a preselected similarity measure. For the recognition of hand postures, the image of a detected hand forms the candidate image which is directly compared with prototype images of hand postures. The best matching prototype (if any) is considered as the matching posture. Clearly, because of the pixel-by-pixel image comparison, template matching is not invariant to scaling and rotation.

Template matching was one of the first methods employed to detect hands in images [27]. To cope with the variability due to scale and rotation, some authors have proposed scale and rotational normalization methods (e.g. [28]), while others equip the set of prototypes with images from multiple views (e.g. [29]). In [28], the image of the hand is normalized for rotation based on the detection of the hands main axis and, then, scaled with respect to hand dimensions in the image. Therefore, in this method the hand is constrained to move on a planar surface that is front in parallel to the camera. To cope with the increased computational cost when comparing with multiple views of the same prototype, these views were annotated with the orientation parameters [30]. Searching for the matching prototype was accelerated, by searching only in relevant postures with respect to the one detected in the previous frame. A template comprised of edge directions was utilized in [31]. Edge detection is performed on the image of the isolated hand and edge orientations are computed. The histogram of these orientations is used as the feature vector. The evaluation of this approach showed that edge orientation histograms are not very discriminative, because several semantically different gestures exhibit similar histograms.

3. PROPOSED WORK

3.1 Basic Idea

The Viola Jones algorithm, with HOG features is used here as HOG features capture overall shape of the object better than the haar-like features proposed originally for face detection by Viola et al. Our own dataset is being formed for training a detector for detecting open-palm, and another for detecting a fist. MATLAB's cascade-detector for object detection, in Computer Vision module, which implements viola-jones algorithm, with HOG features [32] is used for implementation. The detection stage returns a few hand candidates, and some of them may not actually be hands. Especially in case of open-palm, it was found that there were multiple hand candidates surrounding the actual hand, containing parts of hand. In our application, robustness is the most desirable property. Although, the underlying detector is not highly accurate, and gives various false detections per stage, the proposed method aims to improve the performance by employing certain empirical measures. To filter out the false positives, certain heuristics are used.

3.2 Viola-Jones Object Detection Framework

Viola-Jones Object Detection Framework [2], was a seminal work in the literature of Computer Vision, and provided competitive object detection rates for real usage. They used 'Integral Images' as Image representation, using Haar-like features. Using AdaBoost as learning algorithm, their algorithm yields extremely efficient classifiers [33]. They propose training a cascade of classifiers, each one being a weak classifier in itself, becoming increasingly complex, as number of stages increase, where each preceding stage, filters

negatives, so that only promising object candidates are passed on to next stage. This arrangement yields a very fast, and quite accurate object detector in general. Viola-Jones Algorithm is mainly known for its performance in face detection.

3.3 Histogram of Orientation Gradients (HOG)

Histogram of Orientation Gradients are feature descriptors, first described in [3], where the authors N.Dalal and Bill Triggs applied it successfully for the problem of human detection (upright) in images. The essential thought behind the Histogram of Oriented Gradient descriptors is that local object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions. The implementation of these descriptors can be achieved by dividing the image into small connected regions, called cells, and for each cell compiling a histogram of gradient directions or edge orientations for the pixels within the cell. The combination of these histograms then represents the descriptor. HOG descriptors resemble edge orientation histograms [34] [35], SIFT [36] descriptors, and shape contexts [37], but they are computed on a dense grid of uniformly spaced cells and use overlapping local contrast normalizations for improved performance. The HOG descriptor operates on localized cells, thus is robust to geometric and photometric transformations, except object orientation.

3.4 Heuristics

1. User is quite unlikely to move hands over large distances over very small time frames (typically time between consecutive image frames).
2. A sudden Click (Fist detection) should be ignored, when a palm movement is taking place.
3. If a 'Drag' gesture is in progress, give priority to retaining Fist candidates over open-palm candidates. We have used a counter for this purpose, so that intermittent false positives don't result into higher priority to open-palm all the time.
4. While user is making drag-patterns, Fist positions too far away from last position, are either due to wrong detections, or unintentional. Same goes for open-palm positions.
5. A valid open-palm/fist candidate should have at least some amount of skin-like pixels.
6. A valid open-palm/fist candidate should have a minimum/maximum area that can be determined in advance, as range of application, and resolution of camera are known.



Fig.1. A detected hand is shown



Fig.2. A detected Fist is shown

3.5 Skin-Like Pixels

Several researches have shown that skin-color of human-beings across a number of races, lie in a very small area of total color space. Various researches have established empirical thresholds for detecting skin in different color spaces. We have used color-based thresholding in YCbCr color space, for detecting skin. Y component has been left, as it accounts for luminance, to provide some illumination invariance. The pixel values in the range $C_b = [77, 127]$, and $C_r = [133, 173]$ are defined as skin pixels.

Skin-color ratio is defined as:

$$skin_ratio = \frac{count(skin - like - pixels)}{count(pixels - in - image)}$$

A skin-threshold is used to filter out false positives. We have taken $skin_ratio_threshold=0.15$

3.6 Tracking the Hand

Hand (open-palm/fist) can be tracked by comparing centroids from last accepted hand candidate, and currently accepted hand candidate.

$$dx = |last_centroid.x - current_centroid.x|$$

$$dy = |last_centroid.y - current_centroid.y|$$

If, $dx < motion_threshold$ and $dy < motion_threshold$, then respective motion along X-Axis, and Y-Axis is assumed.

3.7 Drag Patterns

User creates a drag pattern when he/she has held the Fist up for more than 2 consecutive frames. All the centroids, are enqueued in a list. When Fist is no more detected, algorithm tries to interpret the gesture. The centroids are connected by straight lines, and the produced pattern is passed on to the gesture recognizer stage.

Since set of gestures is very limited, and fast performance is desired, two basic pieces of information about patterns are used here. The aspect ratio of patterns is used, i.e. width/height ratio, as well as minimum distance from patterns stored in database.

The patterns supported are Scroll Up/Down, Left/Right Flick, Reload/refresh, Close current window.

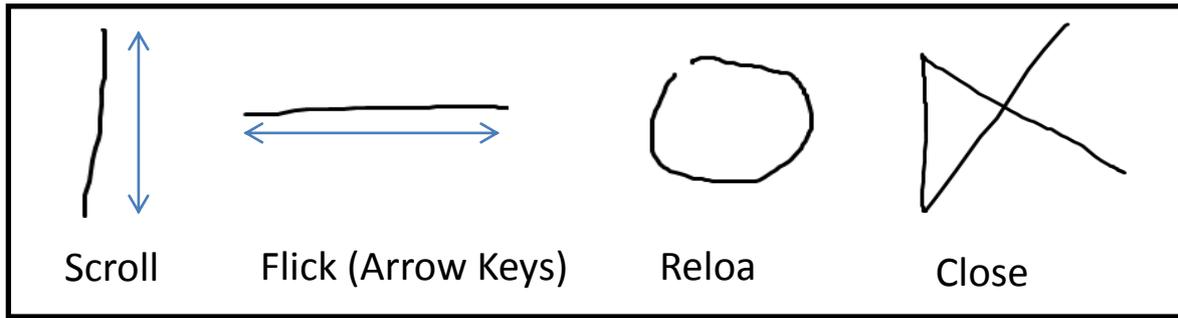


Fig.3. Supported Gestures by the Application

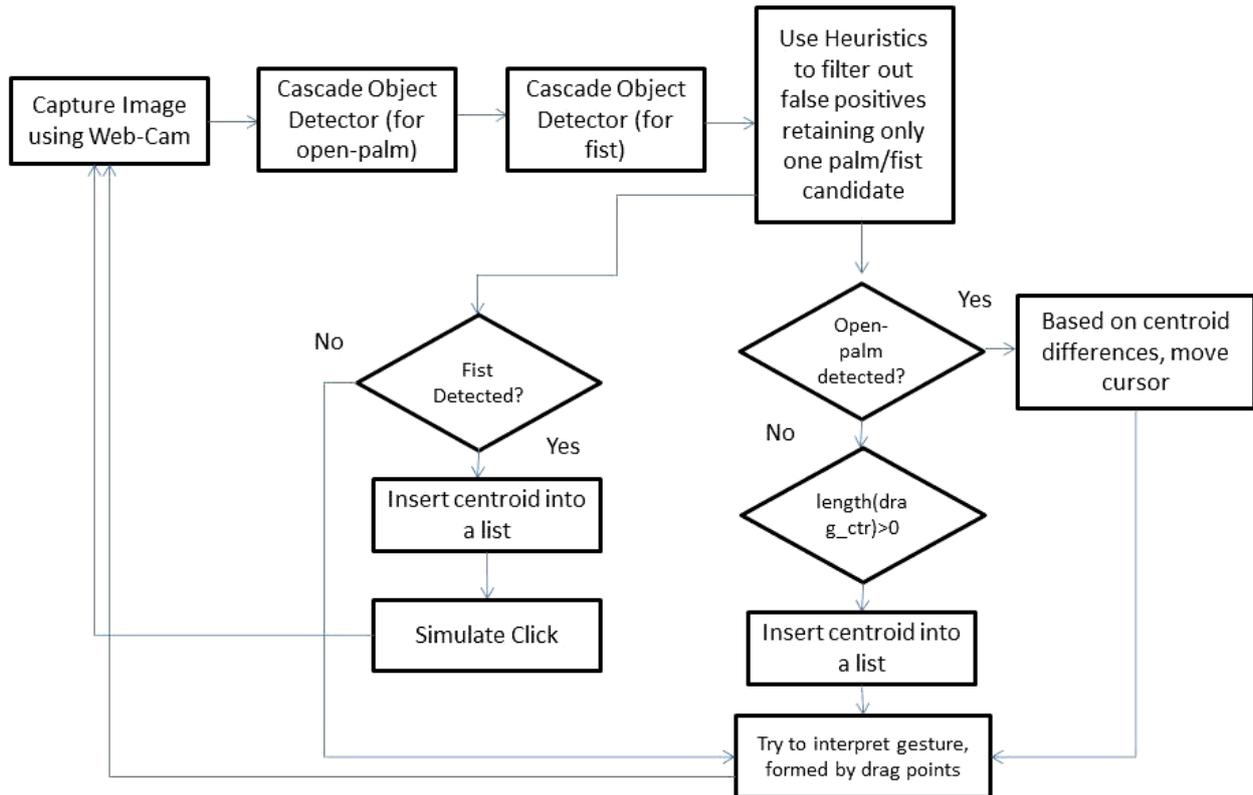


Fig.4. Flow Diagram of the Application

3.8 Sending Commands to the Operating System

Java's Robot Class has been used to automate cursor movements, and clicks. Whenever open-palm is detected, and is moved, cursor is moved. Whenever a Fist is made by user, left mouse click is simulated. When a Reload/Refresh Pattern is made, F5 key press is simulated. Similarly, when a Cross (Close) pattern is made by user, ALT+F4 key press is simulated.

The proposed method is more robust to illumination changes than Gesture recognition approaches which rely wholly on skin-segmentation results. The proposed method uses a low threshold 15-20%, for ratio of skin-like pixels to all pixels, for filtering out hand candidates which are unlikely to be so.

4. RESULTS

The Application on images collected from web and captured by webcam is considered as benchmark to measure accuracy of detection of open-palm and Fist, as well as detection of drag-patterns.

Two sets of experiments have been performed. In one setting, for every frame, the proposed method evaluates number of correct and incorrect detections are present. In another, it evaluates how many patterns are interpreted correctly by the system.

Open-Palm detection rate was nearly 95%, and Fist Detection rate was 72%, owing to smaller dataset used. Pattern detection accuracy was found to be 80% for the simple gestures – scroll up/down, flick left/right. And for tougher ones (with respect to the simple algorithm used), i.e. close and reload, it was nearly 60%. This low rate of pattern detection is acceptable, as a weak pattern-matching algorithm was used, and thus is fast to evaluate.

Compared to other recent works in gesture recognition, like [38], where the authors attained Dynamic Gesture Recognition rates of 95.47% (only UP, DOWN, LEFT and RIGHT), and 95.07% for Static Gestures, our accuracy rates are lower. But they utilize motion history images, which is susceptible to breakdown, in environments with frequent background movements. A similar work is [39], which is a bit

older, where authors have used face detection results for skin-segmentation, thus segmenting hand from image, where they achieved average detection rates of 70%. For static gestures, the proposed method can be compared with [50], where authors report 99-100% accuracy in simple backgrounds; they use template matching with features like Hu's moments, number of fingers, convex defect characteristics. In [51], authors report accuracy of 93%, but they used a very constrained experimental setup.

Table 1. Table captions should be placed above the table

Method	Accuracy	Pros	Remarks
Real time hand gesture recognition system using motion history image [38]	Dynamic: 95.47% Static: 95.07%	Fast, Real-time	Susceptible to breakdown, in environments with frequent background movements.
Real-time Hand Gesture Detection and Recognition using Boosted Classifiers and Active Learning [39]	(Palm, Fist, Pointing, Five) - 70%	Fast	Depends on Face Detection results
Static Gesture recognition Using combination of object features [40]	Up to 100% in simple backgrounds	-	Template matching is usually not robust to variations in environment concerned.
An invariant-based approach to static Hand-Gesture Recognition [41]	Up to 93%	-	Very constrained Experimental Setup
HGRIHCI	Open Palm – 95% Fist – 72% Dynamic – 70%	Fast, Real-time, Robust to illumination changes, mostly unaffected by background clutter.	Use of skin-threshold in one of the filtering heuristics, reduces illumination invariance.

5. CONCLUSION

Using viola-jones detector with HOG features the proposed HGRIHCI method can able to construct an accurate, robust

and fast hand detector with respect to existing techniques. Better results can be obtained with extensive training data. To further decrease false positive rates, a HOG+SVM approach can be adapted. A sophisticated pattern matching can be used, like SVM, for matching patterns. Also, the entire algorithm can be run on GPU, which will speed up the processing greatly.

6. ACKNOWLEDGMENTS

Our thanks to the experts who have contributed towards development of the template

7. REFERENCES

- [1] William T. Freeman, Craig D. Weissman. Television Control by Hand gestures. 1995. IEEE Intl. Workshop on Automatic Face and Gesture Recognition, Zurich, June.
- [2] Viola, Paul and Michael J. Jones, Rapid Object Detection using a Boosted Cascade of Simple Features. 2001. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. Volume: 1, pp.511–518.
- [3] Dalal, N., and B. Triggs, Histograms of Oriented Gradients for Human Detection. 2005. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Volume 1, (2005), pp. 886–893.
- [4] D. Chai, K.N. Ngan. Face segmentation using skin-color map in videophone applications. 1999. IEEE Trans. Circuits Syst. Video Technol. 9 (4).
- [5] T. Kapuscinski and M. Wysocki. Hand Gesture Recognition for Man-Machine interaction. 2001. Second Workshop on Robot Motion and Control, October 18-20, 2001, pp. 91-96.
- [6] C. Yu, X. Wang, H. Huang, J. Shen and K. Wu. Vision-Based Hand Gesture Recognition Using Combinational Features. 2010. IEEE Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 543-546.
- [7] M. Cote, P. Payeur, and G. Comeau. Comparative study of adaptive segmentation techniques for gesture analysis in unconstrained environments. 2006. In IEEE Int. Workshop on Imaging Systems and Techniques pages 28-33.
- [8] R. Francois and G. Medioni. Adaptive colour background modelling for real-time segmentation of video streams. 1999. In Int. Conference on Imaging Science, Systems, and Technology, pages 227-232.
- [9] J. Martin and J. Crowley. An appearance-based approach to gesture-recognition. 1997. In Int. Conf. on Image Analysis and Processing, pages 340-347, Florence, Italy.
- [10] G. Bradski. Real time face and object tracking as a component of a perceptual user interface. 1998. In IEEE Workshop on Applications of Computer Vision, pages 214-219.
- [11] M. Kampmann. Segmentation of a head into face, ears, neck and hair forknowledge-based analysis-synthesis coding of video-phone sequences. 1998. In Proc. International Conference on Image Processing (ICIP) volume 2, pages 876{880, Chicago, IL .
- [12] R. Herpers, G. Verghese, K. Darcourt, K. Derpanis, R. Enenkel,J. Kaufman, M. Jenkin, E. Milios, A.Jepson, and J.1999. Tsotsos.An active stereo vision system for

- recognition of faces and related hand gestures. In Int. Conf. on Audio- and Video-based Biometric Person Authentication, pages 217-223, Washington, D. C.
- [13] T. Kurata, T. Okuma, M. Kourogi, and K. Sakaue. The hand mouse: Gmm hand-color classification and mean shift tracking. In Int. 2001. Workshop on Recognition, Analysis and Tracking of Faces and Gestures in Real-time Systems, pages 119-124, Vancouver, BC, Canada.
- [14] D. Saxe and R. Foulds. Toward robust skin identification in video images. 1996. In IEEE Int. Conf. on Automatic Face and Gesture Recognition, pages 379-384.
- [15] D. Chai and K. Ngan. Locating the facial region of a head and shoulders color image. 1998. In IEEE Int. Conference on Automatic Face and Gesture Recognition, pages 124-129, Piscataway, NJ.
- [16] J. Yang, W. Lu, and A. Waibel. Skin-color modeling and adaptation. 1998. In ACCV, pages 687-694.
- [17] A. A. Argyros and M. I. A. Lourakis. Real-time tracking of multiple skin-colored objects with a possibly moving camera. 2004. In Proc. European Conference on Computer Vision, pages 368-379, Prague, Czech Republic, May.
- [18] W. Freeman and C. Weissman. Television control by hand gestures. 1995. In Int. Workshop on Automatic Face and Gesture Recognition, pages 179-183, Zurich, Switzerland.
- [19] F. Quek. Eyes in the interface. 1995. *Image and Vision Computing*, 13(6):511-525.
- [20] Y. Cui and J. Weng. Hand sign recognition from intensity image sequences with complex background. 1996. In Proc. IEEE Computer Vision and Pattern Recognition (CVPR), pages 88-93.
- [21] R. Cutler and M. Turk. View-based interpretation of real-time optical flow for gesture recognition. 1998. In Proc. International Conference on Face and Gesture Recognition, pages 416-421, Washington, DC, USA, IEEE Computer Society.
- [22] J. Martin, V. Devin, and J. Crowley. Active hand tracking. 1998. In IEEE Conference on Automatic Face and Gesture Recognition, pages 573-578, Nara, Japan.
- [23] J. Crowley, F. Berard, and J. Coutaz. Finger tracking as an input device for augmented reality. 1995. In International Workshop on Gesture and Face Recognition, Zurich, June.
- [24] R. O'Hagan and A. Zelinsky. Finger Track - a robust and real-time gesture interface. 1997. In Australian Joint Conference on Artificial Intelligence, pages 475-484, Perth, Australia, November.
- [25] T. Darrell, I. Essa, and A. Pentland. Task-specific gesture analysis in real-time using interpolated views. 1996. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18(12):1236-1242.
- [26] G. Hager and P. Belhumeur. Real-time tracking of image regions with changes in geometry and illumination. 1996. In Proc. IEEE Computer Vision and Pattern Recognition (CVPR), pages 403-410, Washington, DC.
- [27] W. Freeman and C. Weissman. Television control by hand gestures. In Int. 1995. Workshop on Automatic Face and Gesture Recognition, pages 179-183, Zurich, Switzerland.
- [28] H. Birk, T. B. Moeslund, and C. B. Madsen. Real-time recognition of hand alphabet gestures using principal component analysis. 1997. In Proc. Scandinavian Conference on Image Analysis, Lappeenranta, Finland, June.
- [29] T. Darrell and A. Pentland. Space-time gestures. 1993. In Proc. IEEE Computer Vision and Pattern Recognition (CVPR), pages 335-340, New York, NY.
- [30] H. Fillbrandt, S. Akyol, and K. F. Kraiss. Extraction of 3D hand shape and posture from image sequences from sign language recognition. 2003. In Proc. International Workshop on Analysis and Modeling of Faces and Gestures, pages 181-186, Nice, France, October.
- [31] W. Freeman and M. Roth. Orientation histograms for hand gesture recognition. 1995. In Proc. International Conference on Automatic Face and Gesture Recognition (FG), pages 296-301, Zurich, Switzerland.
- [32] <http://in.mathworks.com/help/vision/ref/vision.CascadeObjectDetector-class.html>
- [33] Yoav Freund and Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. 1995. *International Conference on Computational Learning Theory: Eurocolt '95*, pages 23-37. Springer-Verlag.
- [34] W. T. Freeman, K. Tanaka, J. Ohta, and K. Kyuma. Computer vision for computer games. 1996. 2nd International Conference on Automatic Face and Gesture Recognition, Killington, VT, USA, pages 100-105, October.
- [35] D. G. Lowe. Distinctive image features from scale-invariant keypoints. 2004. *IJCV*, 60(2): pp. 91-110.
- [36] S. Belongie, J. Malik, and J. Puzicha. Matching shapes. 2001. The 8th ICCV, Vancouver, Canada, pages 454-461.
- [37] F. Fleuret, J. Berclaz, R. Lengagne and P. Fua. Multi-Camera People Tracking with a Probabilistic Occupancy Map. 2008. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, Nr. 2, pp. 267 - 282, February.
- [38] Chen-Chiung, Dung-Hua Lion, David Lee. A Real Time Hand Gesture Recognition System Using Motion History Image. 2010. 2nd International Conference on Signal Processing Systems (ICSPS).
- [39] Hardy Francke, Javier Ruiz-del-Solar and Rodrigo Vershae. Real-time Hand Gesture Detection and Recognition using Boosted Classifiers and Active Learning. 2007. *Advances in image and video Technology*, vol. 4872, pp. 533-547.
- [40] Hong Duan, Yang Luo. A Algorithm for Static Gesture recognition Using combination of object features. 2013. *Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013)*.
- [41] Pujan Ziaie, Alois Knoll, An invariant-based approach to static Hand-Gesture Recognition. 2008. 18th International Conference on Artificial Reality and Telexistence, ICAT2008.