



Interface to Interact with Physical World using Human Gesture

Prof. Suhas M. Patil¹, Nandkishor P. More², Sheetal S. Thokal³, Aniket S. More⁴, Bhagyavati V. Patil⁵

Lecturer, Dept. of Computer Engineering, KJ College of Engineering, Pune, India¹

Student, Dept. of Computer Engineering, KJ College of Engineering, Pune, India^{2,3,4,5}

ABSTRACT: The aim of this project is to develop an interface that can be used to connect the digital world with physical. By using the proposed system user can interact with system by making simple hand-gestures. There by the user is not bound to use any specific hardware making him free from the limitations and bounds set by these interfaces. The software takes video stream as input and seeks for the predefined gesture. When it recognizes that a gesture is made it process and stimulates the appropriate action associated with that particular gesture i.e. it collects information about the object & stores it into database. All that user needs is a device integrated with camera that can sense hand-gestures and a processor that process it into appropriate actions.

Keywords: OpenCV, MS-VS2010, object detection & recognition, gesture recognition, Information Retrieval.

I. INTRODUCTION

Any computational system having three major parts

1. Input
2. Processing Unit
3. Output

System takes inputs, process them & generate particular information i.e. Output. This output is further used by the user. For such a Input & Output purpose many devices are used like keyboard, mouse for input purpose & monitor, printer for output purpose. After these hardware devices touch pad, virtual keyboards are used as input device. These traditional devices are must be connected to the system via wired media or wireless media. These required complex electronic devices that increases cost of devices, also these devices are not so efficient to use for the users which does not know about the system.

The project bridges this gap. It uses natural hand gestures as an input device. Any mankind can make simple gestures to perform any action. The device is basically a wearable gestural interface that connects the physical world around with digital information and let's uses natural hand gestures to interact with this information. It implements a gestural camera that takes photos of the scene the user is looking at by detecting the 'framing' gesture. The user can stop by any surface or wall and flick through the photos he/she has taken. For example, if user put any product in mall then it will project all information about that product.

Computer vision:

Computer vision is used to make useful decisions about real physical objects and construction of meaningful description of object from picture or sequence of pictures.

Objectives of computer vision:

- **Segmentation:** Segmentation is breaking images and video into meaningful pieces.
- Reconstructing the 3D world from multiple views, shading and structural models
- **Recognition:** Recognizing objects in a scene
- **Control:** To control obstacle avoidance, robots, machines, etc.

The proposed system is based on Computer vision to fulfill Recognition objective of Computer vision and sees what user see but it lets out information that user want to know while viewing the object.



II.RELATED WORK

For Hand Gesture recognition & interaction systems there are mainly two phases hand detection and gesture recognition.

A) *Hand Detection & Gesture Recognition:*

There are many different approaches for hand detection. The straight forward way for the hand detection is to look for skin color region in the image. But skin color classification is difficult because it can also detect other body parts as a hand in the image. The use of markers attached to users hand can provide reliable detection therefore the proposed system uses red color strip on fingertip. The interface uses camera to track the movement of hand & recognizes gesture.

In computer interfaces, two types of gestures are distinguished.

- Offline gestures: Offline gestures are the gestures after performing these gestures menu gets activated. For example, if circle gesture is performing, image gets captured by camera.
- Online gestures: These gestures are used to rotate or resize the object.

In the proposed system offline gestures are used.

Steps for Gesture Recognition:

- A. User moves finger having red strip and draws a conceptual sign in front of camera.
- B. Finger movement of user is extracted from camera input stream.
- C. From that tracked hand movement conceptual meaning of gesture is determined.

B) *Object detection:*

Object detection is nothing but the localization of the object in input image. In the proposed system for the object detection, the technique SAD (Sum of Absolute Difference) is used. The sum of absolute difference is used as simple way for searching objects inside an image but it may be unreliable because of effect of factors like change in lighting, viewing direction, color, shape, size. SAD can be used in conjunction with other object recognition methods like edge detection to improve reliability of result. This algorithm measures similarity between image blocks. It first takes absolute difference between each pixel in original image and corresponding pixel in image which used for comparison and these differences are summed to create simple metric block similarity. Due to simplicity SAD is extremely fast metric. It is effectively the simplest possible metric that takes into account every pixel in a block. Therefore it is very effective for a wide motion search of many different blocks.

C) *Information Retrieval:*

There is some difference between information retrieval system and database system.

Information retrieval system works on unstructured data while database system works on structured data. Inverted index that is index of entries is used by information retrieval system. Main data structure used in database system is relational tables which have defined value for each row and column. Therefore database system work on data which is related to each other.

When user request for data, Information retrieval process is used to represents, search and stores collection of data for knowledge discovery. An object is an entity that is represented by information in a database. An information retrieval process begins when a user enters a query into the system. Queries are formal statements of information needs, for example search strings in web search engines. In information retrieval a query does not uniquely identify a single object in the collection. Instead, several objects may match the query, perhaps with different degrees of relevancy. User queries are matched against the database information. Depending on the application the data objects may be text documents, images, and audio.

III.WORKING

The proposed system contains mainly two modules.

1. Recognition of Gesture.
2. Detecting Object & Information Retrieval.

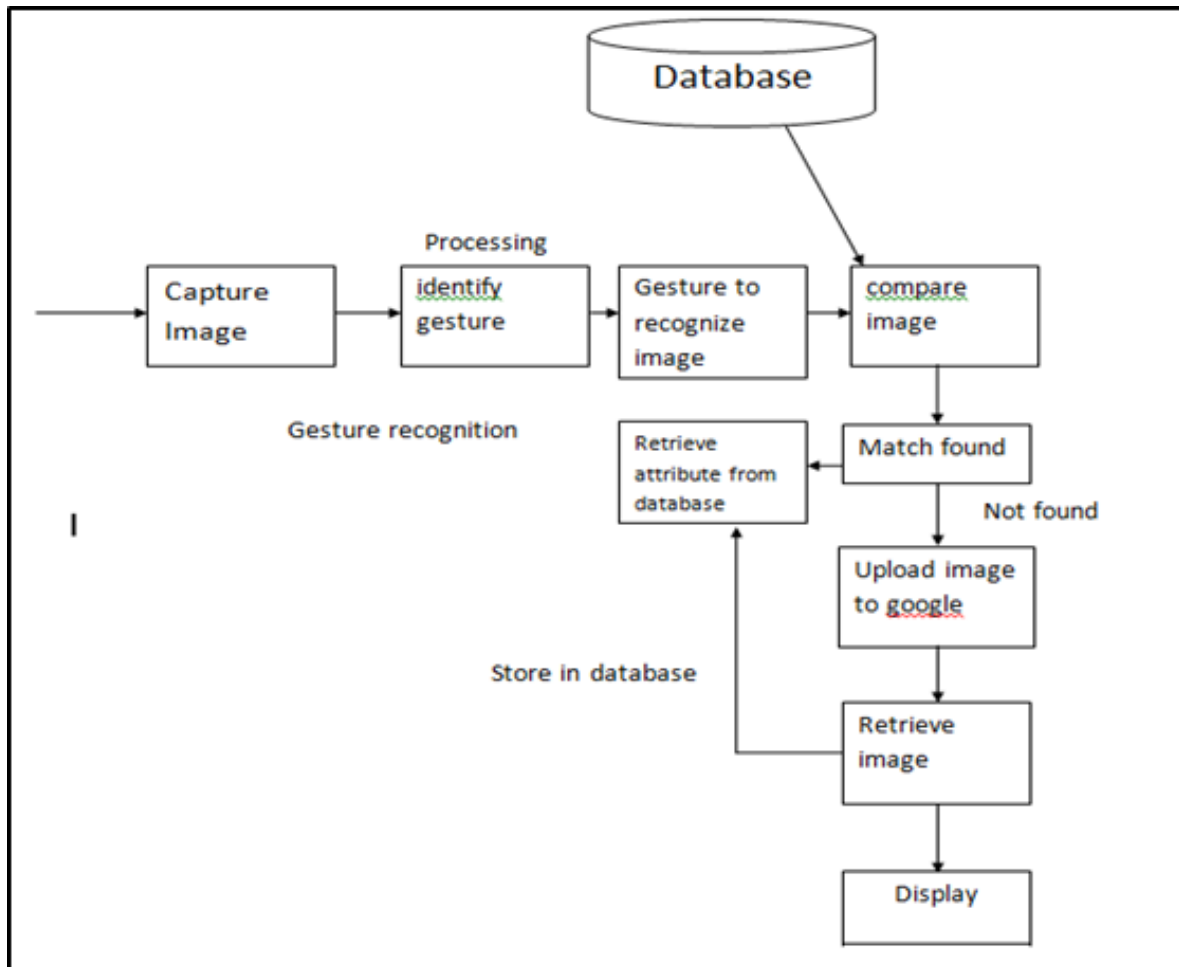


Figure 1: Working of system

Module 1: Recognition of Gesture:

User performs hand gesture in front of camera. Camera takes each frame and finds red color block of strip on finger in each frame. If colored block found then, it calculates its area and if the area is greater than 1000 then it calculate x & y coordinates of center point of that block. Area is calculated because image may contain another color block but it may be relatively smaller than our red strip on finger. After that system draws line from new point to previous one and checks if these lines forms a closed polygon or not. If yes then it finds min & max points of polygon and draws rectangle. This indicates that our gesture is recognized. Then it crops the current frame i.e. object rounded by gesture. Finally this cropped image is given to the second module as input.

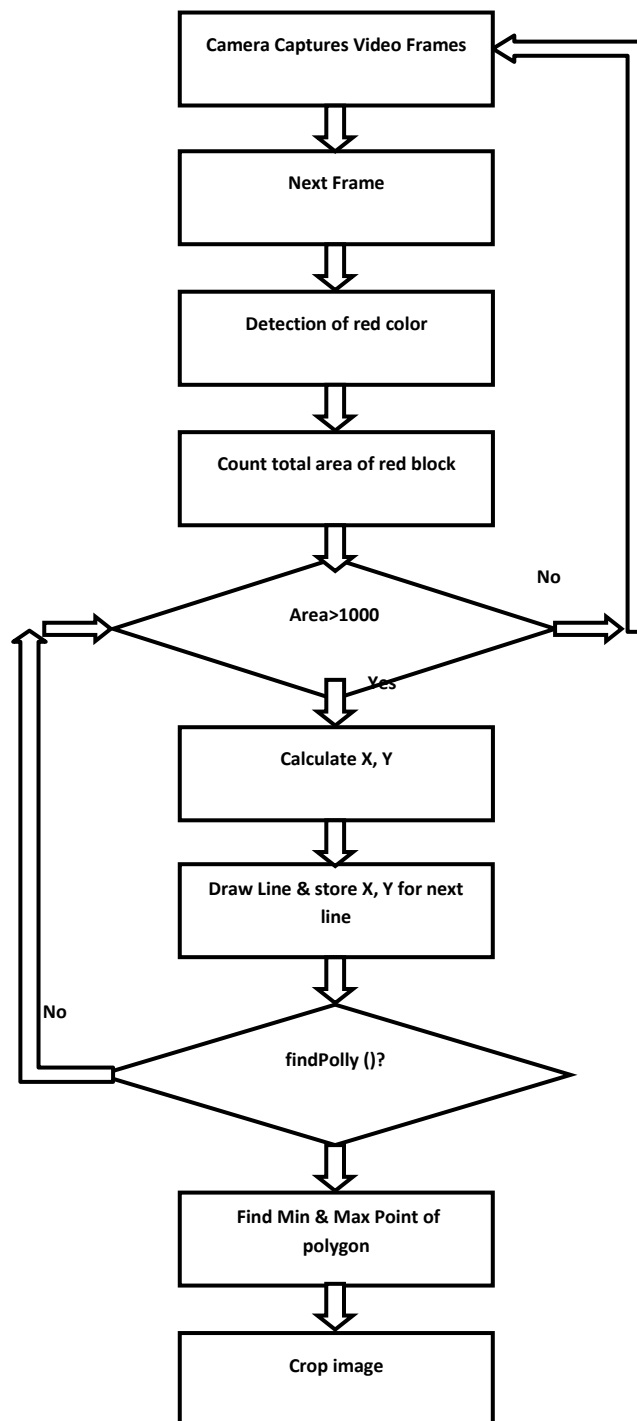


Figure 2: Gesture Recognition

Module 2: Detecting Object & Information retrieval

Following diagram describes flow of module two:

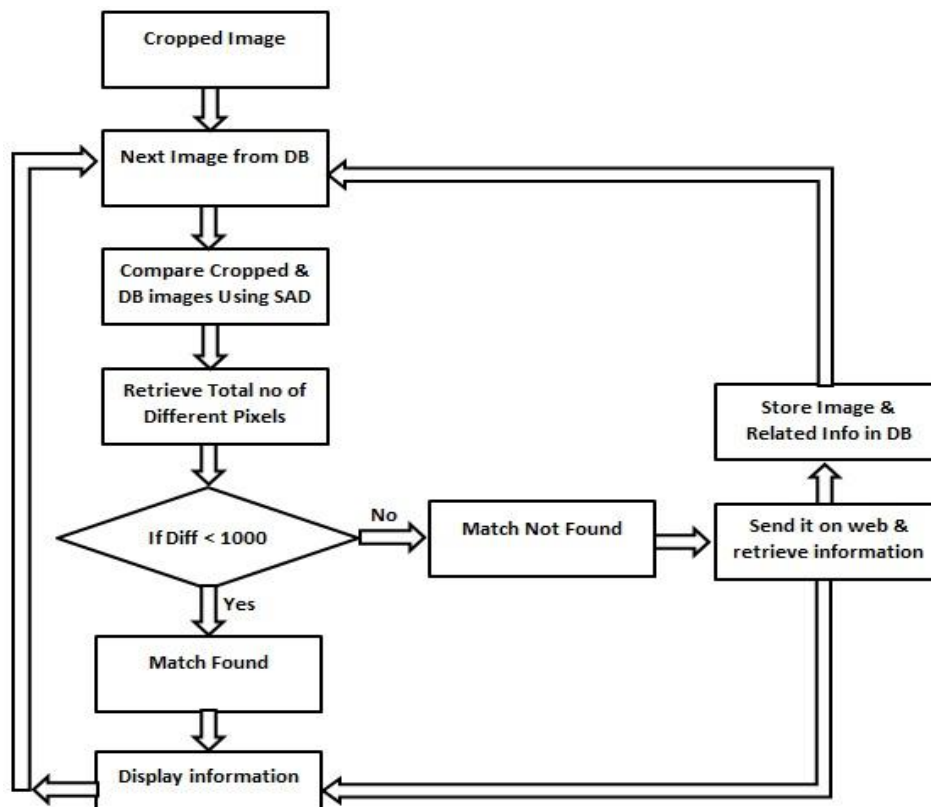


Figure 3: Object Detection

In this module system crops captured image. Then system compares this cropped image with images stored in the database one by one using SAD (Sum of Absolute Difference). If difference is greater than 1000 pixels then it displays message match not found and send this cropped image on web to retrieve information related to that object. After retrieval of information system stores that cropped image and information in the database. If difference is less than 1000 pixels, then it shows match found and displays related information from local database.

Implementation by using OpenCV:

Proposed system is implemented using OpenCV.

A) OpenCV Overview:

OpenCV is an open source computer vision library. The library runs under Linux, Windows and Mac OS X and it is written in C and C++. It contains active development on interfaces for Ruby, Matlab and other languages. OpenCV library has 500 functions which covers areas in vision like factory product inspection, user interface, camera calibration, robotics etc.

B) Goal of OpenCV:

- 1] To provide computation efficiency and strong focus on real time applications.
- 2] To provide simple to use computer vision infrastructure which helps people to build sophisticated vision applications.

C) Structure of OpenCV:

OpenCV contains mainly four components

- CV: CV contains basic image processing and computer vision algorithm.
- MLL: MLL is machine learning library which has many statistical classifiers and clustering tools. This library focused on statistical pattern recognition and clustering.
- High GUI: High GUI has functions for storing and loading video and images and I/O routines.
- CXCORE: It has basic data structures and contents.



IV.FUTURE WORK

As a future work we are planning to use robust algorithms for gesture recognition & object detection and going further we use voice commands to interact with system. We plan to implement OCR (Optical Character Recognizer) in the system to extract text from the image and provide related information.

Information system follows mainly three steps:

- 1] Data representation
- 2] Search and Match operation.
- 3] Returns appropriate result.

V.CONCLUSION

In this paper we have presented a prototype to interact with physical world using hand gestures. Our intention is to provide natural interface to retrieve information about any physical object from web. The proposed system is successfully implemented using hand gesture recognition and object detection techniques using OPENCV and information related to detected object is retrieved.

ACKNOWLEDGMENT

We will forever remain grateful for constant support and guidance extended by guide Prof. Suhas M. Patil, for Completion of project report. Through our many discussions he helped us to form and solidify idea we had invaluable discussions with him. Constant motivation of him led to the development of this project. We are also thankful to our family members for encouragement and support in this project work. We wish to express our sincere thanks to the project coordinator prof. Sumit A. Hirve as well as our Principal Dr. Sanjeev Wagh and departmental staff members for their support.

REFERENCES

- [1] G. Bradaski and A. Kaehler, "Learning OpenCV: computer Vision with the opencv library," O'Really Media, 2008
- [2] Wilhelm Burger and Mark J. Burge (2007). "Digital Image Processing: An Algorithmic Approach Using Java." Springer. ISBN 1846283795 and ISBN 3540309403. <http://www.imagingbook.com>.
- [3] Pedram Azad, TiloGockel, Rüdiger Dillmann (2008). "Computer Vision - Principles and Practice." Elektor International Media BV. ISBN 0905705718. <http://ivt.sourceforge.net/book.html>.
- [4] JörgAppenrodt, Ayoub Al-Hamadi, and Bernd Michaelis "Data Gathering for Gesture Recognition Systems Based on Single Color-, Stereo Color- and Thermal Cameras"International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 3, No. 1, March, 2010
- [5] PrateemChakraborty, PrashantSarawgi, Ankit Mehrotra, Gaurav Agarwal, Ratika Pradhan "Hand Gesture Recognition: A Comparative Study" Proceedings of the International MultiConference of Engineers and Computer Scientists 2008 Vol I IMECS 2008, 19-21 March, 2008, Hong Kong
- [6] R. Y. Wang and J. Popovic, "Real time hand-tracking with color glove," in ACM SIGGRAPH 2009 papers New Orleans, Louisiana: ACM2009
- [7] J. M. Rehg and T. Kanade, "DigitEyes: vision-based human hand tracking," school of Computer Science, Carnegie Mellon University.
- [8] Tsung-Han Tsai and Chung-Yuan Lin, Department of electrical engineeringNational central universityTaiwan, R.O.C"Visual Hand Gesture Segmentation UsingThree-Phase Model Tracking Technique forReal-Time Gesture Interpretation System"Journal of Information Hiding and Multimedia Signal Processing ©2012 ISSN 2073-4212Ubiquitous International Volume 3, Number 2, April 2012
- [9] Iván Gómez-Conde, David Olivieri, Xosé Antón Vila, Stella Orozco-Ochoa, "Simple Human Gesture Detection and Recognition Using a Feature Vector and a Real-Time Histogram Based Algorithm," Journal of Signal and Information Processing, 2011, 2, 279-286 doi:10.4236/jsip.2011.24040 Published Online November 2011 (<http://www.SciRP.org/journal/jsip>)
- [10] Craig Watman_, David Austin_y, Nick Barnesy, Gary Overett_ and Simon Thompsonz "Fast Sum of Absolute Differences Visual Landmark Detector" Robotic Systems Laboratory Department of Systems Engineering, RSISE, Australian National University Canberra, ACT 0200 Australia
- [11] Nadir NourainDawoud, BrahimBelhaouari Samir, Josefina Janier "Fast Template Matching Method Based Optimized Sum of Absolute Difference Algorithm for Face Localization" International Journal of Computer Applications (0975 – 8887) Volume 18– No.8, March 2011
- [12] William T Freeman and Michal Roth, "Orientation Histograms for Hand Gesture Recognition", IEEE Intl. Wkshp. on Automatic Face and Gesture Recognition, Zurich, June, 1995.
- [13] R K McConnell, "Method of and apparatus for pattern recognition", U. S. Patent No.4, 567,610, January, 1986.
- [14] Robert Laganière,"OpenCV 2 Computer Vision Application Programming Cookbook", Copyright © 2011 Packet Publishing