An Image Matching and Object Recognition System using Webcam Robot

Sanjana Yadav
Assistant professor
Dept.of Computer Science & Engineering
Hindustan college of Science & Technology, Farah, Mathura
sanjana.engg21@gmail.com

Archana Singh
Assistant professor
Dept.of Information Technology
Hindustan college of Science & Technology, Farah, Mathura
archisingh.15@gmail.com

***Abstract*— Finding relation among multiple images is a core and vital step in the field of computer vision. This can be achieved if we can track correspondence over consecutive frames in an image and then matching among them can be identified. This paper proposes a precise approach for image matching in real time scenario and also in the field of ROBOTICS. We are constructing an application based Image matching model that can detect images that are exactly the same, as well as images that have been edited in some ways. Implementation of this Image Matching and object recognition system is based on tracking an object, calculating its feature Points, and classification with the help of trained Data Sets. The system is capable to perform matching of images, both automatically and manually. On the other hand, to operate the proposed system manually, user itself takes the images of an object or something, and will store it in database. After that system will through some steps, and image matching will be performed. Black & White point calculation of an image, Chamfer matching Algorithm, 3-4 Distance Transformation with canny edge detector is applied in the application which is well suited for calculating the pixel values. The whole system is reliable and capable to match the two images in Digital Image Processing**.

Keywords-Chamfer matching , Gradient, Edge detection, Image matching, Thresholding.

I. Introduction

Computer Vision deals with automatic information extraction, from an image or series of images, and then analyzes it for understanding information and high dimensional data from the real world. Computer Vision is a science, that makes a machine capable to perceive the world around it in a similar way as human eyes and brains visually sense it. Matching is a key problem in computer vision, image analysis, and pattern recognition: objects, units, or other features in the image must be recognized and named [10].The concept of finding matches among images of a database has been automated and quite efficient. Image alignment is one of the most fundamental tools in many areas of computer vision that may involve target tracking, motion analysis, object recognition, and automated visual inspection. For target tracking, image matching is used to accurately locate target objects in temporal image sequences [11][12]. There are several image matching methods to find region correspondence among image sequences that is to track image motion analysis. As industries are evolving, the need of handling of small parts is also increasing the demand of intelligent vision robots to replace human inspectors [13].

These intelligent robots have better tolerance towards long working hours and more accuracy and repeatability in comparison to human inspectors. Consequently, there has been extensive research in computer vision, especially in the area of template image matching.

Suppose we have some part of a real or an artificial world that we call a scene and we suppose that objects of the scene are static. This means that objects don't move, and also don't change their shapes in time. Next we have a camera which moves in the scene. In certain short time intervals, the camera takes images of the scene [9].

Suppose there are two or more images of one scene, each taken from different angles of the camera, then image matching problem will refer to the process of establishing mapping between each pair of visible homologous image points on a given pair of images. That means the process is applied on two-dimensional (2-D) discrete image functions.

Image Matching exemplifies two fundamental problems of image matching.

a) If the matching problem is handled using only local information then ambiguous solutions may occur.

b) The problem needs better approximation.

Studies prove that usually different image matching algorithms, consider the transformational factors like translation, rotation, scaling and perspective invariance.

However, the inputs taken at the same time from different sensors and inputs to same sensor at different times will always result in unclear output, and matching occluded images are always a challenge [1][13].

For expeditious and promising object recognition, the relevant features and criterion for best matching must be known. For this purpose, through some comparative studies, an optimal method is discovered as a result that method is known as chamfer matching [5]. In this method we have some local features that may be points, edges and black and white points. When we try to automate the system another problem may arise, that is interfacing the robot with the software. Sometimes it may be challenging for a robot to capture the images in unsuitable topological setup and process those images, and there may be times where signalization or factors like these may cause problems in automated image matching.

In the proposed method, our main concern is primarily to find the black and white points and secondarily, the edge points, and whenever a match is detected, the signal automatically sent to the robot to pick the object. The requirement for these types of applications is generally seen in labs and industries where they have authentication systems, and huge machines to pick the objects without or with minimal human involvement.

Therefore we have to build here an automated system of Image Matching where a webcam robot will be allowed to acquire the images and send it to system for processing.

Some preprocessing methods like Image gray scaling, binarization and edge detection are essentially applied. For edge detection and to calculate edge points, out of many edge detection methods available, that are generally and most commonly applied, we have chosen canny edge detector due to its accuracy and robustness.

II. PROPOSED METHOD

The goal of proposed system is to match the object images, taken from different distances, using different image matching algorithms. This paper can either be implemented with any hardware equipment to capture the images automatically or can be manually done by the user. In this work, we are manually inserting different images of a single object as well as, we are using a webcam robot as hardware and for image matching and recognition MATLAB code is used.

The proposed system is shown in flowchart (Fig. 1.) which is also described by the following Steps:

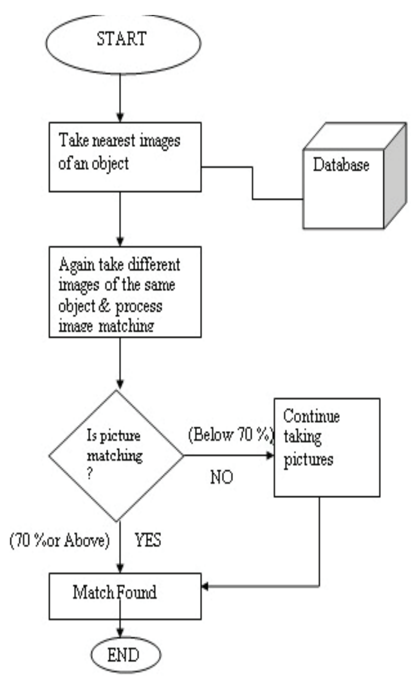

Fig. 1. *Flow chart of proposed system*

a) **Step1: All nearest pictures of the objects are stored in database:**

This is the very first step of proposed system, where the nearest images of the object are stored manually in the database. These images are feature extracted and used for classification. Processing Algorithm will be implemented after the images are stored in the database and are matched with the current images which will be taken by robot at different time interval and from different distances.

b) **Step2: image acquisition at regular interval and motion control of robot:**

System is also interfaced with a mobile robot fitted with a CCD camera and is used to take the images from different distances [4]. The images stored in the database act as source image and the current images taken afterwards by robot act as target images. Matching will be done between these two images. The motion of the robot is controlled with the help of signalization. CCD camera work as eye of robot and the system as a whole is controlled by user so that a proper operation can be executed.

c) **Step3: Matching process within the images using Matching Algorithm**:

In this system we are dealing with Chamfer Distance Transformation because of its simplicity and reliability. But before implementing 3-4 DT, the image is converted to grey scale and binarization is performed to count the black white points in the image. There is one more thing to keep in mind, that we are going to use Canny Edge Detector, to detect the edge points of the template.

Steps Involved in image matching system are following:

- Image Acquisition
- Create a database folder in c: drive where images are to be loaded.
- Images loaded in database can be taken itself by user or by webcam robot.
- Image Preprocessing: Grey Scaling, Filtration and Binarization
- Feature Extraction: Canny Edge Detector to count edge points and Calculating Black –White Points in an Image.
- Classification: Classification on the basis of Image Matching using database, and minimum matching points around 70 percent.
- Automated system interfaced with webcam robot.

A. *Canny Edge Detection*

For edge detection process (Fig. 1.), canny edge detector is an operator that works on multi stage algorithm to find wide range of edges in an image.

Canny's aim was to discover the optimal edge detection algorithm, where being optimal means the following

- strong detection – the algorithm should mark as many real edges in the image as possible.
- good localization – edges marked should be as close as possible to the edge in the real image.
- minimal response – a given edge in the image should only be marked once, and where possible, image noise should not create false edges.

Fig. 2. *Edge detection*

B. Chamfer Distance Transformation

The process of converting a binary image to an approximate distance is called distance transformation (DT). Chamfer distances are the distances between horizontal/vertical neighbors and two local distances in a 3x3 neighborhood.

Distance transform (DT) is such an operation which measures the distance of non-edge pixels to the nearest edge pixel while the edge pixels get the value zero. The purpose of the distance transformation is to produce numeric image whose pixels are labeled with distance between each of them and their closest border pixel [8]. It is important that the DT used in the matching algorithm is a reasonably good approximation of the Euclidean distance, otherwise the discriminating ability of the matching measure, computed from distance map (DT image), and becomes poor [14]. Borgefors [6] indicated that the root mean square average produces better edge distance for measuring the fit between the model and target images and yield fewer false alarm minima than any of the other averages. The root mean square average can be formularized as:

$$D_c = \frac{1}{3}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\mathrm{d_i})^2}$$

where d_i is the Chamfer distance value and n is the number of edge points in model.

In binary image first we set each edge pixel to zero and non edge pixel to infinity. To compute DT using parallel propagation of local distances, the following expression is used and at each iteration each pixel obtains a new value [3][7].

$$v_{i,j}^{k} = minimum(v_{i-1,j-1}^{k-1} + 4, v_{i-1,j}^{k-1} + 3, v_{i-1,j+1}^{k-1} + 4, v_{i,j}^{k-1}, v_{i,j+1}^{k-1} + 3, v_{i+1,j-1}^{k-1} + 4, v_{i+1,j}^{k-1} + 3, v_{i+1,j+1}^{k-1} + 4)$$

Where, $v_{i,j}$ is the value of the pixel in position (i,j) at iteration k. The iteration continues until no value changes and is corresponding to it's longest distance in image [14].

C. Matching Measures

A list is made of the edge pixels that are extracted and converted to pair of coordinates as shown in Fig. 3., where each pair represents the row and column of that edge pixel. From this list the edge points that are actually used are later chosen according to some criterion, which is application dependent [14].The coordinate pairs are superimposed on the distance image. An average of the pixel values that the polygon hits is the measure of correspondence between the edges, called the edge distance. A perfect fit between the two edges will result in edge distance zero, as each polygon point will then hit an edge pixel. The actual matching consists of minimizing the edge distance. There are many variants of matching measure averages, e.g. arithmetic, root mean square and median [2][6][14].

Arithmetic average distance= (v1 + v2 + . . . + vN)/N

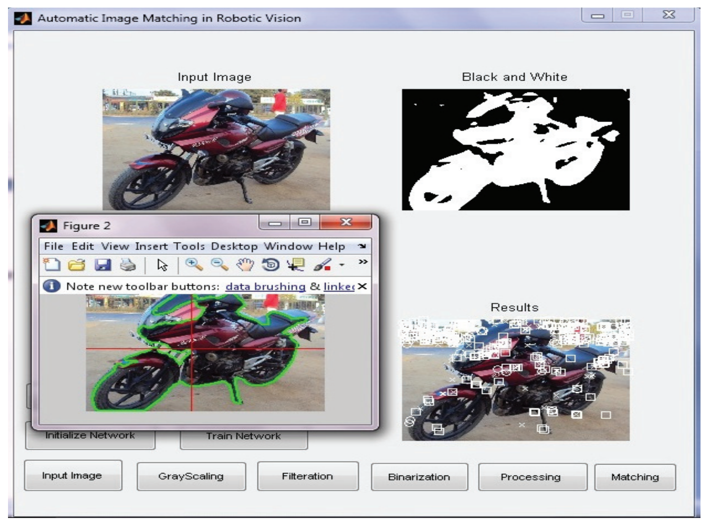

Fig.3. *Image matching*

Root mean square average distance = $\{(v_1^2 + v_2^2 + \ldots\ldots.. + v_n^2)/N\}^{1/2}/3$

where N is the total number of points in coordinate pairs and $v_1 + v_2 \pm \ldots\ldots\ldots. v_N$ are pixel values of corresponding coordinate pairs. Each position of the coordinate pairs corresponds to an edge distance. The position with minimum edge distance is a matching position.

D. Classification via Training dataset:

For Image classification, supervised machine learning procedure is used where a set of features are extracted from the images and passed to previously trained image dataset that maps the features to the defined classes. The purpose of doing this is to detect the matching between an image and all those images present in the trained dataset. After that the class with highest image similarity is find out and chosen. The classification procedure is shown in Fig. 4.

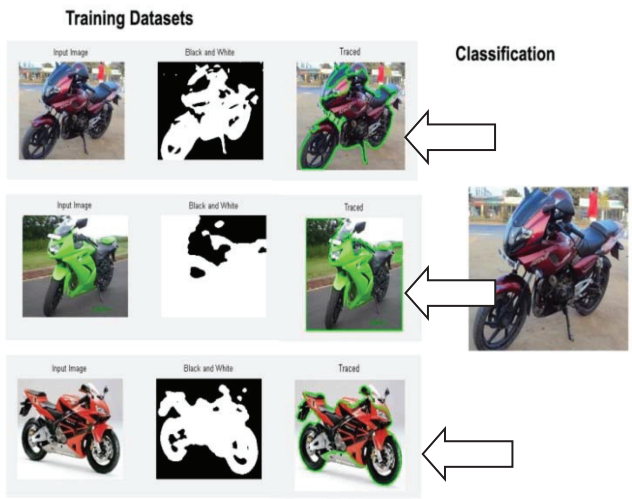

Fig.4. *Classification by image matching*

E. *Orientation Distance transformation (ODT)*

Traditionally the distance transformation method was invented to generate a distance map of two binary images that can be used to measure the likelihood between these images. But to achieve object recognition in video, the idea is to track an object with the given template. Thus, we introduce an Orientation Distance Transform (ODT) which assigns a complex number to each pixel [15].

1) Likelihood measurement: For measurement of likelihood between two images of same size, the idea is to treat one image as mask while the other one is treated as template, and then processed by ODT [15].

2) Shift Estimation: When the mask is superimposed over the orientation distance map, the values of hit pixels in the map can be used to not only measure a likelihood value, but also estimate a shift vector that implies where the object may lie. Consider a binary image with positive pixels standing for edge points. We superimpose it onto its orientation distance transform result after moving it along an orientation for several pixels. Generally, there are several positive pixels in a mask. Each of them suggests a pixel shift vector. A histogram of these vectors is calculated [15]. The main shift vector for the whole mask is determined by the histogram's peaks, which can be located by mean-shift method. Ideally when the mask is moved according to this main shift vector, most positive pixels will hit nearer to or exactly on the edge points. Thereafter a higher likelihood value may be obtained. And the mask can be led to a closer position to the template iteratively.

d) **Step4: Results;**

If it is not matched, then the processed image will be automatically discarded or removed and the same process will be repeated with other images till we get the match.

III. RESULTS OF EXPERIMENT AND ANALYSIS

When the user inserts images manually into the database, the manual system goes through the following steps:

Result 1: We first click on source, and input an image then we select our target image. Here input image and target image are different in terms of contrast. To match both images we click on Match button. The calculated matching percentage is 43% almost that is below 70% which is the threshold value we have set for matching two images. Hence, it will display an Output message "Images not Matching" as shown in Fig. 5.

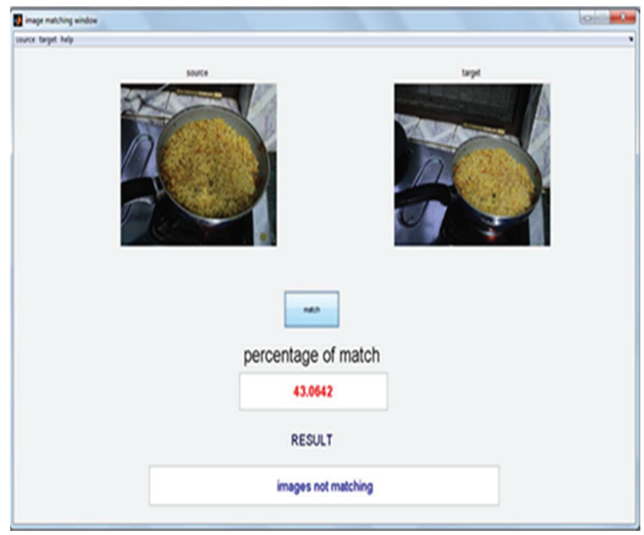

Fig. 5. *Results on first test image*

Result 2: We first click on source, and Input an image then we click and select the target image. Here the input image and target image are different in terms of orientation. To match both images we click on Match button. The matching percentage here is approx 71% that is above 70% which is the threshold value we have set for matching two images. Hence, it will display an Output message "Almost Same" as shown in Fig. 6.

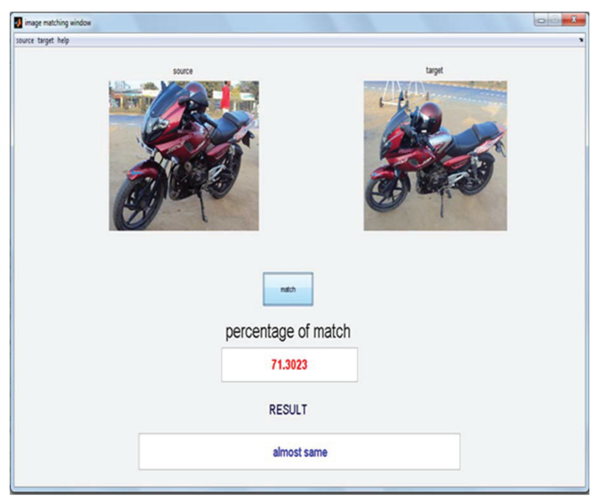

Fig. 6. *Result of second test image*

Result 3: We first click on source, and input an image then we input the target image. Here both the images are exactly same. To match the both images we click on Match button. The matching percentage is 100% that is above than set threshold value. Hence, it will display an Output message "Exactly same" as shown in Fig. 7.

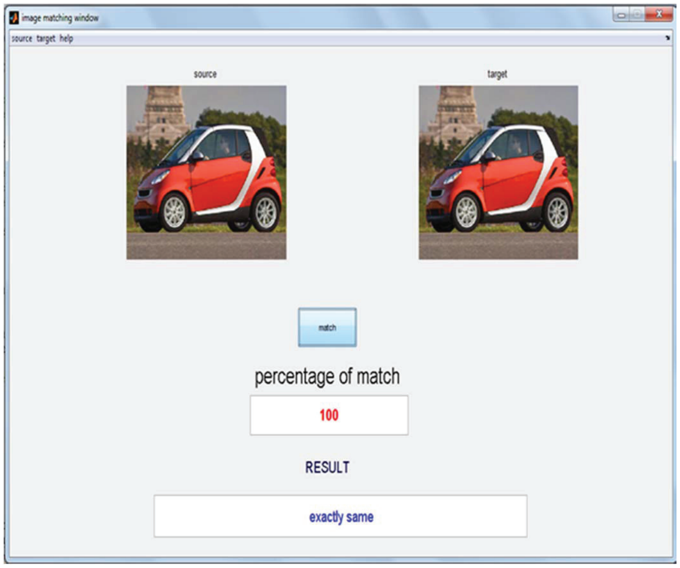

Fig. 7. *Result of third test image*

IV. CONCLUSION

This paper proposes a technique to find out matching percentage among multiple images of an object. The proposed system is capable to calculate percentage of matching among the images that are exactly the same, as well as between the images those are edited in some ways and slightly different from each other. Though in this paper we have presented the manual approach for inserting the images in database, the system can be automated also, where a webcam robot acquires images of an object and send it to database automatically for further processing. Here we have used Chamfer Distance Transformation, that results in efficient and high performance method for object detection, due to its pixel based co-relation approach. For detecting low level features like edges, canny edge detection is used here because of the fact that this algorithm is not exposed to noise, and that makes it able to identify the weak edges. But the algorithm used here is slightly time consuming, because of the number of grey levels involved. This disadvantage can be overlooked if the achieved results are better.

REFERENCES

[1] A. Ghafoor, R. Naveed Iqbal and S. Khan, "Image Matching Using Distance Transform", Springer-Verlag Berlin Heidelberg, pp. 654-659, 2003.

[2] A. Ghafoor, R. N. Iqbal,, S. A. Khan, "Modified Chamfer Matching Algorithm", IDEAL 2003, LNCS 2690, pp 1102-1106, 2003, springer-Verlag Berling , Heidelbarg 2003.

[3] Chuang, Chun- Hsing, C-C J. Kuo and H. M. Hang, " Wavelet deformable model for shape description and multimedia elastic matching", visual Communication and Image Processing 93, 1993.

[4] D. G. Lowe, "Object recognition from local scale-invariant features" proceeding of the seventh IEEE International Conference on Computer Vision pp. 1156-1157 1999.

[5] E. A. Barrow, J. M. Tenenbaum and R. C Bolles, and H. C. Wolf, "Parametric correspondence and chamfer matching: Two new techniques for image matching", SRI International pp. 659-667, 1977.

[6] G Bogerfors," Hierarchichal Chamfer Matching: A Parametric Edge-Matching Algorithm", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.10, pp. 849-866, November 19888

[7] Huang, Aiming, Z. Gao, B. Dai, and L. Luo."Real time object matching",Application of digital Image Processing XXI, 1998.

[8] J. Chen,P. Tan,T. Goh , "Refinement of the chamfer matching for a "center-on", Singapore Technologies Dynamics,I mage and vision computing, N Z, Nov 2003.

[9] J. Walder, "Using 2D wavelet analysis for matching two images, CESCG 2000,

[10] Sanjana, M. Diwakar and A. Sharma, "Image Matching in Robotic Vision in Application of Tracking an Object" in proceedings of International Conference on Emerging Trends in Computer Science and Information Technology, 2012.

[11] S. H. Lai, "Robust Image Matching under Partial Occlusion and Spatially Varying Illumination Change", Computer Vision and Image Understanding 78, pp. 84-98, 2000.

[12] S. H. Lai, M. Fang, "Robust and efficient image alignment with spatially varying illumination models", IEEE, 1063-69,1999.

[13] S. Khan, " Robust image matching algorithm", Proceedings EC-VIP-MC 2003 4th EURASIP Conference focused on Video/Image Processing and Multimedia Communication(IEEE Cat No. 03EX667)VIPMC, 2003.

[14] S. M. Smith, and J. M. Brady, "SUSAN A new approach to low level image processing", International Journal of Computer Vision, pp. 45-78, 1997.

[15] Wu, Tao, X. Ding, S. Wang, K. Wang and D.Fofi, " Video object tracking using improved chamfer matching and condensation particle filter", Image Processing Machine Vision Applications, 2008.