BLOB DETECTOR BASED GABOR DESCRIPTOR FOR FEATURE EXTRACTION

Szidónia Lefkovits[1], *László Lefkovits*[2]
[1]*Petru MaiorUniversity Tg-Mureș, Department of Computer Science*
[2]*Sapientia University Tg-Mureș, Department of Electrical Engineering*

Abstract: *The human visual system has the extraordinary capacity of recognizing a wide variety of objects or object categories from only two- or three-dimensional visual information. The aim of artificial vision is to model and to create robust automatic detection systems. Artificial vision is a branch of general object detection that processes two-dimensional images as a projection of three-dimensional space. The most up-to-date research has not led to a general system that could be useful for solving all practical applications. Each of the existing systems is created with a specific aim and work in certain given conditions. In this paper we describe a part-based local descriptor for facial feature classification. Further, we discuss the influence of interest point detection in part-based systems, by comparing supervised and unsupervised interest point selection methods. The blob like features implemented in our system are Laplacian of Gaussian operator, Kadir-Brady saliency detector and the circle Hough transform. These saliency point extraction methods determine the center point closest to the center of the eye. After this step we apply our Gabor descriptor presented in previous articles and create four different classifiers for the selected interest points with the GentleBoost algorithm. Finally, due to our experiments we draw conclusions on the applicability of the four interest point detectors for facial feature detection, by comparing the detection performances of the obtained descriptors and classifiers.*

Keywords: local descriptor, 2D Gabor wavelets, Gentle Boost, blob-like features: Kadir-Brady, LoG, Hough transform

1 Introduction

The task of object detection is one of the most widespread research domains in computer vision. Despite of the fact that humans have the extraordinary capacity to detect several types of objects or object categories; the general problem of object detection has not been solved yet. Nowadays, a high-performance system supposes a high rate of detection, and at the same time, a low rate of false detections.

The three main parts of such a system are: the interest points, the local descriptor and the object model. The interest points represent set of points that stand, in a way, out of their environment; in other words they capture the visual attention. The local descriptor represents a formal description of an image region in the neighborhood of the interest point. The local descriptors, applied in deformable object models, have the advantage of handling small deformation and partial occlusions. Based on the physical structure of the object, the model separates the target objects from other objects. Hence relying on the model the classification can be achived.

This paper presents a robust, part-based facial feature detection method. The selection of methods is done in several ways LoG (Laplacian of Gaussian), Hough transform. These unsupervised blob detectors are compared with the manually marked interest points, so with the supervised learning. The image parts detected around these interest points have been selected to be characterized by our 2D Gabor filter based patch descriptor created especially to detect the eye form the facial region. After obtaining the 2D Gabor filter responses in each case, we compare the detection performance of them. It is obvious that the supervised learning method is better than the unsupervised blob based interest point detection, but the goal of this article is exactly this, namely to compare the detection rates and to put an accent on the advantages and disadvantages of the above mentioned methods.

The paper is organized as follows: The first section presents an introduction, the second discusses about the interest point and local descriptors in specialized literature, the third consists of a theoretical review of three the blob like-features (LoG, Kadir-Brady and Hough transform) and Gabor wavelets for feature descriptor. Finally, the last section third section presents experimental results obtained.

2 Related work

Different interest point detectors have been proposed in the state of the art. Scale invariant detectors suppose that scale change is the same in all directions, although they are someway robust to little affine deformation. These types of interest point relay on the mathematical concept of local maxima, so they can be driven from the mathematical derivatives. Difference of Gaussian [17] feature point is computed from derivatives of the Gaussian. It is taken in consideration, if the obtained value is greater than a threshold. Lindenberg proposes the LoG operator [15], that is the Laplacian of Gaussians (second order derivatives). The Hessian detector [20] uses the second order matrix to determine interest points. Harris et al. proposes an operator [19] for corner detection, which considers the differential of corner score, with respect to the direction. Hessian affine region detector [19] is computed from the second order derivate and the Harris corner measure and the authors propose multiple scale iterative algorithm to spatially localize and select scale and affine invariant points. Maximally stable extremal regions (MSER) [18] and color MSER [2] are robust similarity measures for establishing tentative correspondences.

One of the histogram based feature descriptors that rely on the distribution of the gradient intensities is called Historgram of orinteted gradients [3]. Here, the features are obtained from little image parts, called cells and the more cells form a block descriptor. The scale invariant feature transform proposed by Lowe [17] is one of the most used scale invariant interest point detectors. It is based on the extrema point of the DoG and LoG in different scales of an octave. Speeded Up Robust Features are integer approximations of Hessian matrix detectors computed by the integral image [1]. The performances of these last two interest points are compared in a comprehensive overview article [21].

Kadir-Brady saliency detector [9] is a multi-scale algorithm for salient region detection. It is mostly used for object tracking or matching regions in similar images. For further details regarding the interest points and region detectors we suggest to consult the survey articles [20, 24].

After the study of state of the art interest points, we have chosen to determine the center of the eye with special circular-blob like features, i.e. LoG, Kadir Brady and Hough transform.

The second phase after the selection of interest points is the development of the most adequate local descriptor. In this case we propose a local image descriptor based on 2D Gabor wavelets [11-13]. Similar systems using Gabor filters were proposed [23] to obtain local features. The method proposed [25] defines 48 Gabor filters 6 frequencies and 8 directions for an image patch of 13×13 around the interest point. The majority of applications based on Gabor filters use a set of empirically chosen parameters regardless the object of interest. Our local descriptor is especially created for facial feature detection, concering a case study for the eye. The detailed description of the structure of local descriptor is presented in our previous articles [12].

3 Theoretical review

This section presents a short overview of circular region detectors used in our experiments. After determining the blob like features from the face, the interest points that most likely correspond to the eye centers are used in the training phase, in order to create the local feature descriptor that characterizes the region of the eye. These obtained region descriptors are finally classified and used as a detector.

3.1.1 LoG - Laplacian of Gaussian operator

This detector has been invented by Lindeberg [14,15] and is used for blob-like feature detection. The operator is obtained by the second order derivative filter, the Laplacian smoothed by a Gaussian. The analytical form of the 2D filter is

$$\nabla^2 G = \frac{x^2 + y^2 - 2\sigma^2}{2\pi\sigma^4} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

It can be observed that this is a negative kernel. It has a strong negative peak in the center and a positive ring surrounding it (figure1). Modifying the scale variable σ, it searches for the maximum similarity of the image region on which it is applied. It is a scale-invariant blob detector because it detects the scale of the region by computing the scale space extrema of a certain point.

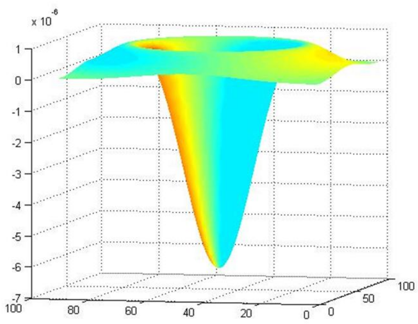


Figure 1 LoG operator

3.1.2 Kadir-Brady saliency detector

The Kadir-Brady saliency detector was presented in the article [9] and invented by the authors Timor Kadir and Michael Brady, the name of which comes from. The idea of this algorithm is based on the dissimilarity measure obtained from the information theory. The image regions having the intensity histogram with one peak are more complex than those with smooth histograms. This means that a complex local descriptor is able to determine interest points. The algorithm finds the position x and scale s obtaining circular image regions. In every circular region it calculates the probability density $p_{x,s}(d_i)$ of the descriptor chosen $d_i \in \mathcal{D}$. The used descriptor in this algorithm is the histogram of obtained intensities: $d_i \in \{0,1,\ldots,255\}$. The density function $p_{x,s}(d_i)$ is computed in every image region. For every possible interest point the algorithm computes the entropy of the image region considered, having the center point x and radius s.

$$\mathcal{H}_D(x,s) = -\sum\nolimits_{d_i \in \mathcal{D}} p_{x,s}(d_i) \log_2 p_{x,s}(d_i). \tag{2}$$

The scale of the image region will be the radius, which corresponds to the maximum entropy of image region centered on *x*. In order to differentiate the interest point the algorithm uses a saliency measure $\mathcal{Y}_D$. This is the weighted maximum entropy, which is computed by $\mathcal{Y}_D = \mathcal{H}_D(x,s) \cdot \mathcal{W}_D(x,s)$, where the $\mathcal{W}_D(x,s)$ weight is the differential value of the probability densities, respect to the scale.

$$\mathcal{W}_D(x,s) = -s \sum\nolimits_{d_i \in \mathcal{D}} \frac{\partial}{\partial s} p_{x,s}(d_i). \tag{3}$$

The last step is the elimination of useless saliency points. The points are ordered decreasingly based on the saliency measure $\mathcal{Y}_D$ and with the Greedy algorithm only those point are included in the final set of interest points, which are not in the interior of other circular saliency regions with greater saliency measure. This last selection was substituted with a k-nearest neighbor algorithm proposed in the next article of the authors [10].

The advantage of this interest point selection is the invariance to rotation, translation and scale. A robust version of this detector was designed and presented in article [22]. The modifications that are proposed in this paper are stabilizing the difference between consecutive scales when calculating the inter-scale saliency, overlap of pixels, partial volume estimation and windowing.

3.1.3 Hough transform

This algorithm was introduced by Paul Hough [8] initially only for straight lines. The detection of this types of objects is, in fact, a local extrema detection. The algorithm is the following: for every edge pixel in a binary image the corresponding parameters are computed and accumulated in an accumulation table. The fundamental idea of the algorithm is the parametric equation of the line, where *c* is the intercept and *m* is the slope of the line.

$$y = m \cdot x + c \ or \ c = -x \cdot m + y. \tag{4}$$

The *XY* space is transformed into the *CM* space, given the edge point in an image (x_i, y_i) and setting the number of angles m_k. For every point we create a concentric set of lines that all cross the considered point. To determine the line an accumulator array has to be declared. $A(C,M)$. This accumulator table is initialized to 0, first. Next, for every angle m_k and a current point (x_i, y_i) we compute the corresponding intercept *c*. Hence, the value of *c* obtained with the angle m_k make a vote in the accumulator table. The table at parameter $A(C,M)$ is increased with 1 unit. Because the slope can be arbitrary and infinite small Duda and Hart [4] proposed the use of the Hesse normal form

$$\rho = x sin\theta + y cos\theta. \quad (5)$$

The advantage of this type of representation is the parallel processing of the point, also a partially deformed and noisy shape can be also determined.

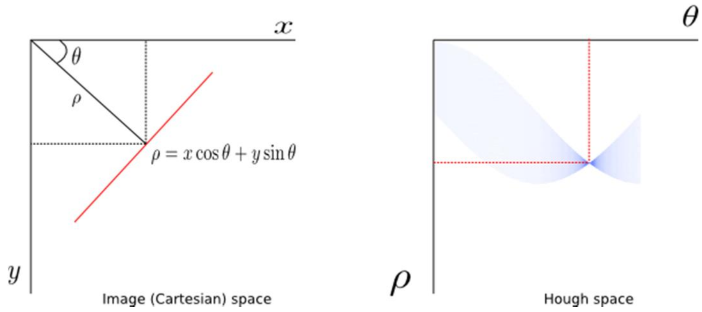


Figure 2 Line representation in Cartesian and Hough spaces [16]

The standard Hough transform (SHT) is the most know form, developed to detect lines, but over the years, several variations have been suggested to identify other analytical or even irregular shapes.

In our experiment we have used the Circle Hough Transform CHT used to determine analytical shapes as circle or ellipse. Here the local maxima have to be determine in 3D space $A(x, y, r)$. For every point in an image (x_i, y_i) and for every possible radius r_k we determine the parameters x, y, r of A obviously, from the equation of the circle:

$$(x - x_i)^2 + (y - y_i)^2 = r^2. \quad (6)$$

A comparative study of the CHT is presented in [7]. The comparison considers the SHT for circles, the fast Hough transform (FHT), and the space saving approaches devised by Gerig and Klein [6].

3.2 Gabor wavelets

In order to determine an adequate image descriptor the first step is to define the interest points of an object or object part. The Gabor wavelets have a wide area of use, especially in bioinformatics systems, because they work similar to the mammalian visual receptive field. The Gabor filter describes not only the interest point, corresponding to the center of the image patch, but creates a local image descriptor as well, covering the area of interest.

The 2D Gabor functions are defined as follows

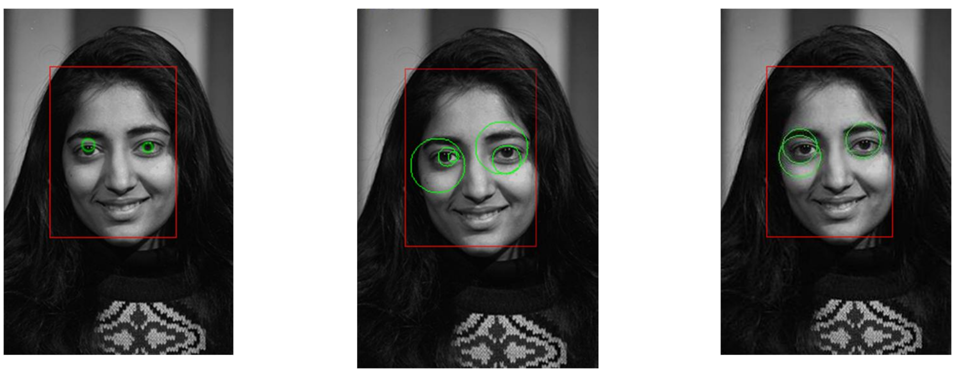
$$g(x, y) = \frac{1}{k} e^{-\pi\left[\frac{(x-x_0)_r^2}{\alpha^2} - \frac{(x-x_0)_r^2}{\beta^2}\right]} \cdot e^{j[\xi_0(x-x_0)+\nu_0(y-y_0)+P]}. \quad (7)$$

where r means the rotation of the envelope surface with θ_0 in trigonometric direction.

The Gabor wavelet is a plane wave modulated by a Gaussian envelope. This function is defined in a 9D parameter space. The number of parameters can be reduced considering the half magnitude profile.
Taking the obtained 4D space we define a considerable number of Gabor filters. Based on these, the system computes the filter response centered on the image patch. In order to choose only the most representative filters and the weight of each one in the final decision, a learning algorithm has to be applied. In our last paper we proposed the GentleBoost algorithm for this purpose [11,13]. In this paper we use the same algorithm for determining the most appropriate Gabor filters and to determine the final classifier. Hence, the goal here is to compare the three interest point selection methods with blob like features presented in section 3.1. and to classify each of them based on the descriptor proposed and presented in our previous research articles.

4 Experimental results

In our experiments we have compared supervised and unsupervised interest point detection techniques and combined those with our Gabor filter based feature descriptors. The Gabor descriptors obtained are used to create the final classifier by Boosting algorithm.
In our previous experiments we have manually marked facial points and created the Gabor descriptor [12] around these feature points. In this paper we discuss about blob like features that are able to determine the center of the eye in an unsupervised or semi-supervised manner. These methods are used in order to reduce computation time and to get rid of manual annotation of facial interest points. In our experiments we have used four methods first the supervised eye center points, than the blob like features that extract circular regions: Laplacian of Gaussian, Kadir-Brady saliency detector and Circle Hough Transform. Figure 3 shows the detection of interest points used in the next step. These interest points are, in fact, the circle centers.



a. LoG **b. Kadir-Brady saliency detector** **c. Hough transform**

Figure 3 Interest point detection example

For all these we have created Gabor descriptors separately. The parameters of the four dimensional feature-space (λ,S,bw,θ) have been fine-tuned in the training process in order to find the most adequate filters based on the positive and negative training examples. Similar to our previous paper [13], we have defined 3024 Gabor filters: 14 frequencies, 6 aspect ratios, 3 bandwidths and 12 orientations taking in account the real part, imaginary part, magnitude and the distribution of the complex responses. The final classifier has been crated with the GentleBoost algorithm that selects 33 weak classifiers based on the Gabor filters defined. Surely, in each case different 33 filters or filters with different weights have been selected, and here comes the difference between the final classifiers obtained. The experiments have been carried out for the eye extracted from the FERET [5] database. The training set consists of 730 positive and 2000 negative examples and the test set of 160 and 500 patches. The image patch used in the training phase is 33×33 pixels centered on the eye and the negative images have been extracted randomly from the face, but not the eye. In order to compare the interest point detectors in our system training and validation conditions have been ensured.

Interest points are detected based on certain geometrical properties that present some kind of consistency, ex. scale or rotation invariance. The LoG (Laplacian of Gaussian), Kadir-Brady saliency detector and Hough transform searches the image and selects parts that have a circular aspect. Due to the circular aspect of the eye, we have decided to use these interest point detectors The experiments done on test sets show a detection rate of 95% after searching for the first 200 LoG points per image. The average distance of the closest LoG pixel to the marked (correct location) is approximately 2.27 pixels. This property suggests the applicability of the LoG detector in the training process as well as the final detection process.

Compared to the LoG detector Kadir-Brady detector detects several saliency points, from these we have to select the most appropriate interest point, closest to the eye. The computational complexity of this detector is greater than that of the LoG's. The position obtained from this interest point detector has a deviation from the correct location of 6.73 pixels in average. The descriptor based on these interest points has a detection rate of 83%.

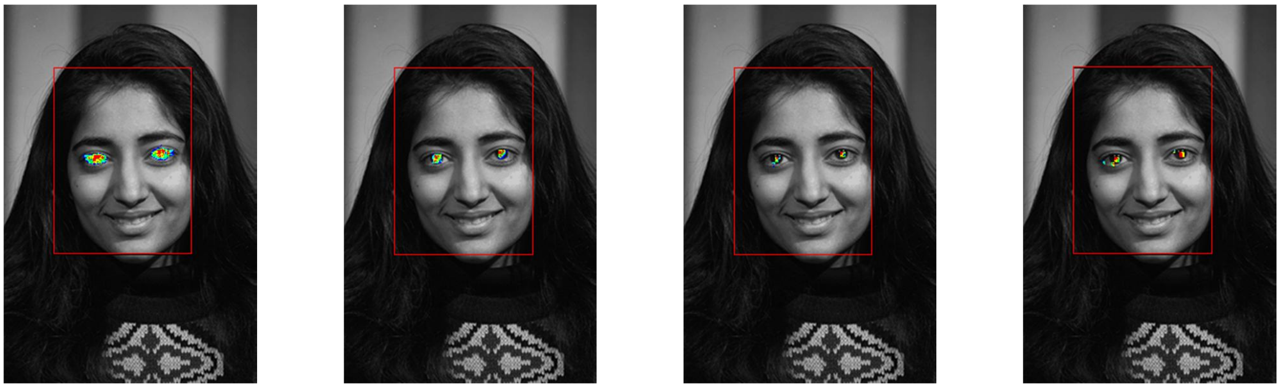
The last interest point that has been compared in this study was the well-known Hough transform for circle points. This type of points can be detected and implemented in an easy way, but they detect only perfect circles, or circles with little deviation thanks to the clustering algorithm applied after detecting the circles. The detection rate of this classifier is 88%.

Based on the extraction of these interest points different image databases have been obtained. On these databases we have created the corresponding Gabor descriptor formed of 3024 Gabor wavelets. The detection rates have been measured by the classification obtained from the GentleBoost algorithm which has selected the first 33 best weak classifiers. The table 1 shows the detection performances measured: the detection rate, the false positive error and the false negative error.

	Supervised	**LoG**	**Kadir-Brady**	**Hough transf.**
DetRate[%]	98	95	83	88
ErrFP[%]	0.31	0.72	6.93	6.94
ErrFN[%]	3.05	7.30	20.45	16

Table 1 Comparison of interest point detectors

Figure 4 shows the obtained Gabor descriptors obtained from the three interest point selection methods presented. It is obvious that the supervised method with manually marked points overtakes all the saliency point based detection methods. But the most important advantage of unsupervised interest point detection is the automatic selection of the facial points included in the training set.



a. Marked interest point **b. LoG** **c. Kadir-Brady points** **d. Hough transform**

Figure 4 Gabor descriptor detection example

5 Conclusion and future work

This paper presents a facial feature detector based on Gabor filter responses. Compared to our previous papers [11-13] in this article we analyze the supervised learning method with unsupervised blob like detection. In our experiments we discuss about the LoG, Kadir-Brady and Hough transform for feature detection. These features detect circular parts form the image, especially the eye. From the extracted image parts we careate four different databases (one from supervised and three from unsupervised) and based on these images we manage to construct four different Gabor descriptors. It is obvious that the manually marked image database has the best detection performance, but the advantage of the blob like detectors is the fully-automated way of cratering the underling training image samples. For the future we propose to build a descriptor which is able to detect more significant facial features.

Acknowledgement

Financial support of the work of Szidónia Lefkovits was provided from programs co-financed by The Sectorial Operational Program of Human Resources Development, Investing in people! Key Area of Intervention 1.5 “Doctoral and Postdoctoral scholarship in support of research”, Project title: “Integrated system for quality improvement of the doctoral and post-doctoral research in Romania and promoting the role of science in society”, Contract: POSDRU/159/1.5/S/133652.

The work of László Lefkovits in this paper is supported by The Sectorial Operational Programme Human Resources Development POSDRU/159/1.5/S/137516 financed by the European Social Found and by the Romanian Government.

References

[1] Bay H., Tuytelaars T. and Gool L. van: "SURF: Speeded Up Robust Features". *Proceedings of the 9th European Conf. on Computer Vision, Springer LNCS* vol. 3951, part 1. pp. 404–417, 2006

[2] Chavez, A.; Gustafson, D.: "Color-Based Extensions to MSERs". *International Symposium on Visual Computing ISVC*, pp. 358–366, 2011.

[3] Dalal, N. and Triggs, B., "Histograms of Oriented Gradients for Human Detection". *Computer Vision and Pattern Recognition, CVPR 2005. IEEE Computer Society Conf*, pp. 886-893, 2005.

[4] Duda, R. O. and P. E. Hart: "Use of the Hough Transformation to Detect Lines and Curves in Pictures," *Comm. ACM*, vol. 15, pp. 11–15, 1972.

[5] FERET Database: http://www.itl.nist.gov/iad/humanid/feret/feret_master.html

[6] Gerig, G. and Klein. F.: "Fast Contour Identification through Efficient Hough Transform and Simplified Interpretation Strategy" *8th International Conference on Pattern Recognition - ICPR* , Paris, pp. 498-500, 1986

[7] Hassanein, A. S.; Mohammad, S.; Sameer, M. and Ragab, M. E.: "A Survey on Hough Transform, Theory, Techniques and Applications". *International Journal of Computer Science IJCSI* vol. 12, issue 1, January 2015

[8] Hough, P. V. C.: "Method and Means for Recognizing Complex Patterns", US Patent 3,069,654, Ser. No. 17, 7156 Claims, 1962.

[9] Kadir, T. and Brady, M.: "Scale, Saliency and Image Description". *International Journal of Computer Vision*. vol.45, no. 2, pp:83–105, 2001.

[10] Kadir, T., Zisserman, A. and Brady, M.: "An affine invariant salient region detector". *Proceedings of the 8th European Conference on Computer Vision*, 2004.

[11] Lefkovits, Sz. and Lefkovits, L.: "Enhanced Gabor Filter based Facial Feature Detector". *Contemporary Perspectives on European Integration Between Tradition and Modernity*, vol. 1., no. 1, pp. 1196-1206, 2013.

[12] Lefkovits, Sz.: "Improvements on Gabor Descriptor Retrival for Patch Detection". *Computing and Infromatics* (accepted, in editing).

[13] Lefkovits, Sz.: "Novel Gabor filter-based patch descriptor". *IEEE 10th Jubilee International Symposium on Intelligent Systems and Informatics*, pp. 373-378, 2012.

[14] Lindeberg, T. :"Scale invariant feature transform". *Scholarpedia* vol. 7 no. 5:10491, 2012.

[15] Lindeberg, T.: "Feature detectionwith automatic scale selection", *International Journal of Computer Vision*, vol. 30 no. 2 pp.79-116, 1998.

[16] Linear Hough Transform Using Python http://www.nabinsharma.wordpress.com

[17] Lowe, D. G., "Distinctive Image Features from Scale-Invariant Keypoints", *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[18] Matas, J.; Chum, O.; Urban, M; and Pajdla, T: "Robust wide baseline stereo from maximally stable extremal regions." *Proc. of British Machine Vision Conference*, pp. 384-396, 2002.

[19] Mikolajczyk, K.; Tuytelaars, T.; Schmid, C.; Zisserman, A.; Kadir, T and Gool, L. Van: "A Comparison of Affine Region Detectors". *International Journal of Computer Vision*, vol. 65, no.1-2, pp. 43-72, 2005.

[20] Mikolajczyk, K. and Schmid, C.: Scale & affine invariant interest point detectors. *International Journal on Computer Vision* vol. 60, no. 1. pp:63-86,2004.

[21] Panchal, P.M.; Panchal, S.R. and Shah, S.K.: "A Comparison of SIFT and SURF ", *International Journal of Innovative Research in Computer and Communication Engineering* vol. 1, issue 2, 2013.

[22] Shao. L.; Kadir, T. and Brady, M.: "Geometric and Photometric Invariant Distinctive Regions Detection". *Information Sciences*. vol. 177 no. 4, pp:1088-1122, 2007.

[23] Tamminen, T. and Lampinen J.: "Sequential Monte Carlo for Bayesian Matching of Objects with Occlusions", *IEEE Trans Pattern Anal Mach Intell*. vol. 28, no. 6, pp. 930-941, 2006.

[24] Tuytelaars, T. and Mikolajczyk, K.: "A survey on local invariant features". *In: Foundations and Trends in Computer Graphics and Vision*, vol. 3 no. 3. Now Publishers, Boston, 2008.

[25] Vukadinovic, D. and Pantic, M., "Fully automatic facial feature point detection using Gabor feature based boosted classifiers". *Systems, Man and Cybernetics, 2005 IEEE International Conference on*, vol. 2, pp. 1692-1698, 2005