Defence Science Journal, Vol. 63, No. 3, May 2013, pp. 292-297, DOI: 10.14429/dsj.63.2763

A Distributed Weighted Voting Approach for Accurate Eye Center Estimation

Gagandeep Singh* and Maheshkumar H. Kolekar

Indian Institute of Technology, Patna-800 013, India

**E-mail: g.singh@iitp.ac.in*

ABSTRACT

This paper proposes a novel approach for accurate estimation of eye center in face images. A distributed voting based approach in which every pixel votes is adopted for potential eye center candidates. The votes are distributed over a subset of pixels which lie in a direction which is opposite to gradient direction and the weightage of votes is distributed according to a novel mechanism. First, image is normalized to eliminate illumination variations and its edge map is generated using Canny edge detector. Distributed voting is applied on the edge image to generate different eye center candidates. Morphological closing and local maxima search are used to reduce the number of candidates. A classifier based on spatial and intensity information is used to choose the correct candidates for the locations of eye center. The proposed approach was tested on BioID face database and resulted in better Iris detection rate than the state-of-the-art. The proposed approach is robust against illumination variation, small pose variations, presence of eye glasses and partial occlusion of eyes.

Keywords: Eye center estimation, iris recognition, biomarkers, biometrics

1. INTRODUCTION

Eye center location information is used in detecting facial features. Eyes are one of the salient and stable features of human face and play an important role in face recognition systems and in understanding facial expressions. The position of the eyes is used to obtain a normalized face which is robust to translation, rotation and scaling and can improve the recognition rate of a face recognition system. Eye center location is also an important step in gaze estimation. Gaze estimation can provide useful information regarding area of interest of a person looking at a computer system. This information can be exploited in the design of systems involving interactions between humans and computers. For example, Gaze can be used as a pointing device for disabled persons, which have limited means of communication. Knowledge of the direction of a user's gaze may help a computer to ascertain certain cognitive states of a user such as confusion, excitement, fatigue or boredom. Besides these, eye detection has important applications in eye tracking and video surveillance.

Any eye center location system has to deal with challenges such as changes in illumination, changes in pose of the subject, occlusion of the eye by the eyelids, presence of eye glasses etc. The various techniques applied for eye center detection can be classified into shape based[1-3], appearance based[4] and hybrid methods[5]. Shape based methods use the geometric properties of the eyes to define a geometric model of the eye which is used for eye detection. In appearance based methods, photometric characteristics of the eye are used for detection. In hybrid based methods, advantages of different eye detection methods are combined within a single system to overcome their respective shortcomings.

A detailed survey of various eye detection approaches is provided in Hansel[6], *et. al*. Under the shape based methods, voting based systems are quite popular because they are simple, easy to implement and have low computational complexity. Kothari and Mitchell[2] proposed a voting scheme that uses spatial and temporal information to detect the location of the eyes. In their approach, pixels vote to a subset of pixels in a direction opposite to the gradient with equal weightage. The generated candidates are then classified using a classifier. A similar voting scheme is suggested by Valenti and Gevers[1]. Their method is based on isophote curvatures in the intensity image and uses edge orientation directly in the voting process. The approach relies on a prior face model and anthropomorphic averages to limit false positives. Among the various voting schemes proposed for eye center location, the approach of Valenti and Gevers[1] is considered state-of-the-art. They compute displacement vector based on radius of curvature at each pixel according to the following formula:

$$D(x,y) = \frac{-\{I_x, I_y\}(I_x^2 I_y^2)}{\{I_y^2 I_{xx} - 2I_x I_{xy} I_y + I_x^2 I_{yy}\}} \tag{1}$$

where I_x and I_y are the first image derivatives, I_{xx} , I_{xy} , and I_{yy} are the second image derivatives and $\{I_x , I_y\}$ represents gradient vector. In this approach, first left and right eye regions are separately extracted using anthropometric measurements. Pixels having non-positive curvature are allowed to vote for the eye center. The votes are weighted according to the following importance measure:

$$\text{Curvedness} = \sqrt{I_{xx}^2 + 2I_{xy}^2 + I_{yy}^2} \tag{2}$$

The major drawback of the above approach is that it

Received 20 October 2012, revised 22 March 2013, online published 16 May 2013

requires accurate determination of radius of curvature which is very difficult in presence of noise. Moreover, for lower solution images, computation of second derivatives is often error prone. The anthropometric measures used for extracting eye regions make the method less suitable for cases in which face model is not fixed. In this paper we propose a Distributed weighted voting based approach for eye center location. The advantage of the proposed method over the approach of Valenti and Gevers[1] are as follows:

1 The approach works very well for low resolution image.
2 There is no need to calculate second derivatives in our approach,
3 Instead of allowing a pixel to vote for only one candidate, we distribute its vote according to a Gaussian kernel.
4 No anthropometric measurements are used to separate eye regions.
5 The algorithm operates on whole face image.

2. PROPOSED APPROACH

The salient stages of the proposed approach are shown in Fig.1. We first normalize the image, this is followed by edge detection and our distributed weighted voting mechanism. The candidates are generated next which are then classified using a classifier.In this section, we outline proposed approach for eye center estimation.

(a) (b) (c)

Figure 1. Photometric normalization steps (a) gamma corrected image (b) difference of gaussian filtering (c) contrast equalized image.

2.1 Photometric Normalization

This is a preprocessing phase incorporated to remove the effects of illumination variations, local shadowing and highlights while preserving the essential elements of visual appearance[4,7,8]. This phase is divided into following steps:

(i) *Gamma correction:* In this step a non-linear gray level transformation is applied on the input image in which gray level I is replaced by I^{γ} where $\gamma \; \varepsilon \; [0,1]$ is a user defined parameter. This transformation enhances the local dynamic range of image in dark or shadowed regions while compressing it in bright regions and at highlights. The value of γ is fixed to 0.2 in our experiments.

(ii) *Difference of Gaussian filtering:* Difference of Gaussian[9] is a band pass filter that is used to increase visibility of edges. In this step, image obtained in the previous step is convolved with two Gaussian kernels having different standard deviations to obtain two blurred versions of the image. One of the images is less blurred than the other. The more blurred version is subtracted from the less blurred version to obtain the final image. The values of standard deviations for this step are fixed to 1 for narrower Gaussian and 2 for wider Gaussian for all images. The equation for two dimensional Difference of Gaussian kernel is given in Eqn (3).

$$I(x,y) \leftarrow I * \frac{1}{2\pi\sigma_1^2} e^{\frac{-(x^2+y^2)}{2\sigma_1^2}} - I * \frac{1}{2\pi\sigma_2^2} e^{\frac{-(x^2+y^2)}{2\sigma_2^2}} \tag{3}$$

(iii) *Contrast equalisation:* A drawback of using difference of Gaussian is a reduction in contrast of the resulting image.

The resulting image also contains some extreme values produced by highlights, garbage at the image borders and small dark regions such as nostril. To counter this effect, contrast equalization is applied. In this step the following operations are applied one after another (i.e., Eqn (3) followed by Eqn (4)) on every pixel. The effect of applying these equations is the reduction in range of values present in the resulting image. The outputs of individual steps of the photometric normalization process id given in Fig.1.

$$I(x,y) \leftarrow \frac{I(x,y)}{\left[mean\{\lambda, |I(x',y')|^{\alpha}\}\right]^{1/\alpha}} \tag{4}$$

$$I(x,y) \leftarrow \frac{I(x,y)}{\left[mean\{\min(\lambda, |I(x',y')|)^{\alpha}\}\right]^{1/\alpha}} \tag{5}$$

Here, *mean* is taken over entire image, *a* is a strongly compressive exponent that reduces the influence of large values, τ is a threshold used to truncate large values and the mean is over the whole image. The value of $a = 0.1$ and $\tau = 5$ in our setting. Note we use absolute values while calculating mean in Eqns (3) and (4). The reason for this is that the image obtained after applying difference of Gaussian contains negative values for some pixels.

2.2 Edge Detection

Canny Edge detector is applied on normalized image to obtain an edge map of the image. The edge map is convolved with a Gaussian filter. This makes the edge map smooth for further processing. Figure 2 shows the difference between edge detection performed on normalized and un-normalized images. It can be observed that the edges corresponding to the eye region are clearer in Fig. 2(b).

(a) (b)

Figure 2. Canny edge detector output on (a) un-normalized image (b) normalized image.

2.3 Distributed Voting

This stage is divided into following steps:

(i) *Estimation of gradient direction*: Accurate estimation of image derivatives is very essential for the success of our approach. The first derivatives of the resulting image are calculated using the 7-tap coefficients given by Farid and Simoncelli[11]. The direction of gradient is estimated by the following formula:

$$\varphi(x, y) = \tan^{-1} \frac{I_y(x, y)}{I_x(x, y)} \tag{6}$$

The gradient direction is quantized into eight directions as shown in Fig. 4 with different colors representing bins.

(ii) *Calculation of voting kernel: A* one dimensional Gaussian kernel is used as voting kernel. The width *n* of the kernel is calculated as follows. A random subset of database is selected. The radius of left and right eye of each subject in the subset is manually calculated. The average of radius of left and right eye for each image in the subset is calculated. The resulting radius is averaged over all the subjects in the subset. The value of *n* is chosen close to the averaged radius.

(iii) *Voting:* An accumulator array A is initialized to zero. Each edge pixel whose gradient magnitude is greater than a certain threshold β takes part in voting. The value of β in our work is 0.2* (mean of gradient magnitude over whole image). Since eye region is darker than surrounding regions and gradient points in direction of increasing intensity, each qualified pixel votes to *2*n* closest pixels which lie in a direction opposite to its quantized gradient direction. The contribution by pixel (x, y) to t^{th} closest pixel located in opposite gradient direction *is* given by $I(x, y) *g(t)$. Here $I(x, y)$ is the intensity at edge pixel (x, y),$g$is the one dimensional Gaussian Kernel and $1 \leq t \leq 2*n$.

(iv) *Post-processing:* The resulting accumulator array *A* is then replaced by A^{α} where $\alpha > 1$ is used to eliminate non-radially symmetric features such as lines. The resulting image is again convolved with a Gaussian Kernel.

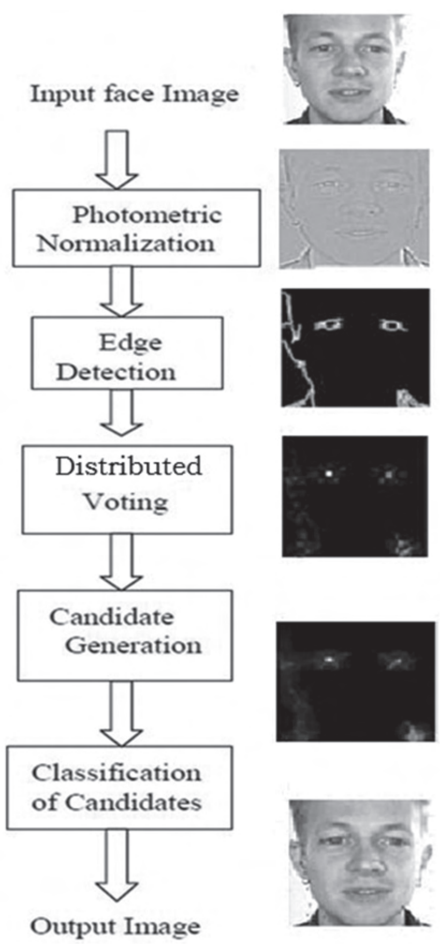

Figure 3. Block diagram of proposed system.

Figure 4. Quantization of gradient direction into eight bins.

2.4 Candidate Generation

The voting phase results in many candidates which are located close to each other. In order to reduce the number of candidates, we merge regions containing pixels which are close to each other. In this phase, morphological closing operation is applied to fill out the gaps in the resulting image which results connected regions. Local maxima are then found in each region. These maxima are the potential candidates for eye centers.

2.5 Classification of Candidates

All resulting pairs of candidates are classified on the basis of following measure:

$$|d - d`| + |y_{1-} y_2| + 0.5*|max - B(x_1, y_1)| + 0.5*|max - B(x_2, y_2)| \tag{7}$$

Here *d* is Euclidean distance between eye centers measured manually, *d`* is Euclidean distance between the candidates, $B(x_1, y_1)$ and $B(x_2, y_2)$ are the intensities at the two candidates at locations (x_1, y_1) and (x_2, y_2) respectively in the image *B* obtained after fourth stage and *max* is the global maximum intensity value in *B*. The pair of candidates having minimum value on this measure is selected as the estimated eye centers.

3. RESULTS AND DISCUSSIONS

3.1 Experimental Results

We have used Bio ID database[12] for testing our approach. The dataset consists of 1521 gray level images with a resolution of 384 x 286 pixels and has been taken in uncontrolled illumination condition. Besides changes in illumination, the position of the subjects is variable both in scale and pose. Bio-ID dataset consists of the challenging cases such as subject wearing glasses, closed eyes, eyes turned away from the camera, eyes are completely hidden by strong highlights on the glasses. A ground truth of the left and right eye centers is provided with the dataset. Figure 3 shows the images during each step of the proposed algorithm. Figure 5 shows how proposed system has accurately estimated eye centers even in the presence of glasses and partial occlusion cases. Red

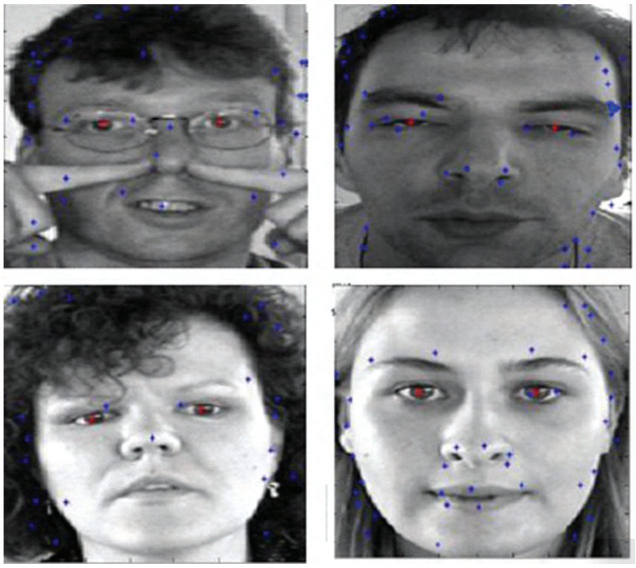

Figure 5. Sample results on Bio ID dataset. Red dots in the images represent the estimated eye centers.

dots and blue dots indicate the estimated eye center and the candidates for eye center respectively. The normalized error[13], indicating the error obtained by the worse eye estimation, is used as the accuracy measure for the found eye locations. The measure is defined as:

$$e = \frac{\max(d_{left}\ d_{right})}{w} \tag{8}$$

where d_{left} and d_{right} is the Euclidean distance between the located eyes and the ones in the ground truth, and w is the Euclidean distance between the eyes in the ground truth. In this measure $e \leq 0.25$ roughly corresponds to the distance between the eye center and the eye corners, $e \leq 0.10$ corresponds to the range of the iris, and $e \leq 0.05$ corresponds to the range of cornea. Table 1 shows that the accuracy of eye detection for proposed approach is better than those reported in the literature. For comparison with state of the art, worst case, best case and average case error estimates are considered in Fig. 6. In worst case estimation only worst eye estimates are considered, in best case, only the best eye estimates are considered, in average case, the average of errors in estimating both eyes are considered. It is clear from Fig. 6(a) that the proposed approach performs well achieving an accuracy of approximately 80 per cent for $e \leq 0.05$, for $e \leq 0.10$ the approach yields accuracy of approximately 98 per cent.

Table 1. Accuracy *vs* Normalized error for different methods in percentage

Method	Accuracy ($e \leq 0.05$)	Accuracy ($e \leq 0.10$)	Accuracy ($e \leq 0.25$)
Proposed Approach	80.00	97.98	100
MIC[1]	77.15	82.11	96.35
MIC + MS[1]	79.56	85.27	97.45
MICs + SIFT[1]	84.10	90.85	98.49
Asteriadis[15]	74.00	81.70	97.40
Jesorsky[13]	40.00	79.00	91.80
Cristinacce[14]	56.00	96.00	98.00
Turka[15]	19.00	73.68	99.46
Bai[16]	37.00	64.00	96.00
Campadelli[17]	62.00	85.20	96.10
Hamouz[18]	59.00	77.00	93.00

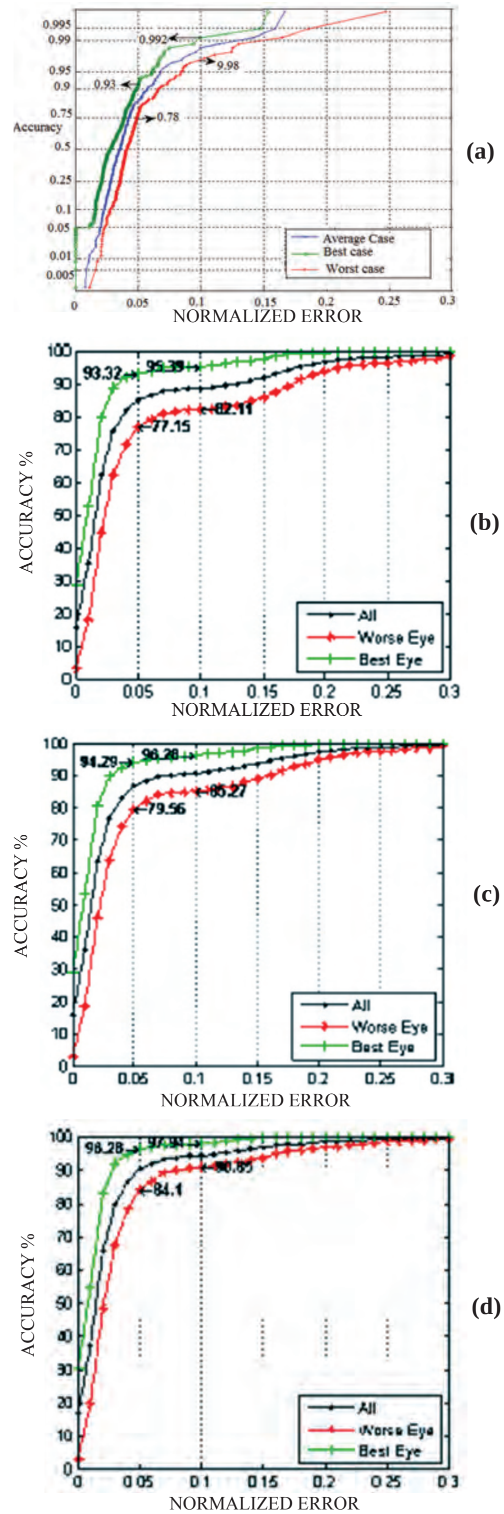

Figure 6. Accuracy vs. minimum (best eye) and maximum (worse eye) normalized error (a) obtained with the proposed method, (b) MIC, (c) MIC+MS, (d) MICs+ SIFT KNN obtained from Valenti and Gevers[1].

3.2 Parameters

There are two variable parameters to our system

3.2.1 Standard Deviation σ of the Gaussian

The value of this parameter affects the amount of blurring produced in the image and varies from image to image. A low value results in less blurring whereas a higher value causes more blurring. We generally use value of $\sigma = 1.0$ for most of the images in our experiments.

3.2.2 Threshold t for the Canny Edge Detector

The value of threshold varies from image to image. A low threshold results in the generation of too many candidates while high threshold can result in the removal of information necessary for generating correct candidates. The Gaussian Kernel is used as voting kernel because of its symmetric and discriminating properties. The proposed classifier performs very well. It fails only when there are too many candidates. Sometimes, problems may occur when the face is tilted too much. The proposed classifier assumes that prior knowledge of distance between the eye centers is available.

3.3 Complexity Analysis

Let N be the total number of pixels in the face image, σ be the width of Gaussian, *p* be the number of pixels in the structuring element used for morphological closing which is a disk of radius *n*, and *q* be the number of candidates generated. Table 2 shows complexity of various stages of the proposed approach. The computational complexity of the proposed approach is quite low as compared to the other existing methods.

Table 2. Computational complexity of the proposed algorithm.

Stage	Complexity
Photometric normalization	$O(\sigma N)$
Edge detection	$O(\sigma N)$
Distributed weighted voting	max((O(σN)), (O(2*n*N))
Candidate generation	$O(pN)$
Classification of candidates	O(q^2)

4. CONCLUSION

The proposed distributed weighted voting approach for eye center estimation is accurate and robust against illumination variation. It also works well for low resolution images and able to cope with partial occlusion of the eye and slight variations of pose. The proposed technique has low computational complexity and achieves significantly better results than the state-of-the-art methods. The accuracy of our method can be further increased if we are able to localize the eye regions. For this, we are interested in using structured learning techniques[19].

ACKNOWLEDGEMENT

Authors would like to mention special thanks to Prof I. Kakadiaris, University of Houston, USA for providing valuable suggestions for improvement in this work.

REFERENCES

1. Valenti, R. & Gevers, T. Accurate eye center location and trackingusing isophote curvature. *In* the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2008, pp. 1-8.
2. Kothari, R. & Mitchell, J.L. Detection of eye locations in unconstrained visual images. *In* the International Conference *Image Process.*, 1996, **3**, pp. 519-522.
3. Valenti, R.; Sebe, N. & Gevers, T. Combining head pose and eye location informationfor gaze estimation. *IEEE Trans. Image Process.*, 2012, **21**(2), 802-815.
4. Bourlai, T.; Whitelam C. & Kakadiaris, I. Pupil detection under lighting and pose variations in the visible and active infrared bands. *In* the IEEE International workshop on Information Forensics and Security, 2011, pp. 1-6.
5. Cootes, T.F.; Edwards, G.J. & Taylor, C.J. Active appearance models. *In* the *European* Conference Computer Vision, 1998, **2**(6), pp 484-498.
6. Hansen, D.W. & QiangJi. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2010, **32**(3), 478-500.
7. Tan, X. & Triggs, B. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans. Image Process*., 2010, **19**(6), 1635-1650.
8. Lloyds, Raja G. & Kolekar, Maheshkumar H. Illumination normalization for image restoration using modified retinex algorithm. *In* the Annual IEEE India conference (INDICON), 2012, pp. 941-946.
9. Marr, D. & Hildreth, E. Theory of edge detection. *Proc. Royal Society of London, series B, Biological Sci.*, 1980, **207**(1167), 215-217.
10. Geusebroek, J.M.; Smeulders, A.W.M. & Weijer, J. van de. Fast anisotropic gauss filtering. *IEEE Trans. Image Process.*, 2003, **12**(8), pp 938-943.
11. Farid, Hany & Simoncelli, Eero. differentiation of discrete multi-dimensional signals. *IEEE Trans. Image Process.*, 2004, **13**(4), 496-508.
12. BioID technology research. The BioID face database. http://www.bioid.com, 2001.
13. Jesorsky, O.; Kirchbergand, K.J. & Frischholz, R. Robust face detection using the hausdorff distance. *In* the International Conference on Audio- Videobased Biometric Person Authentication, Springer Lecture Notes in Computer Science, 2001 pp. 90-95.
14. Cristinacce, D.; Cootes, T. & Scott, I. A multi-stage approach to facial feature detection. *In* the British Machine Vision Conference, 2004, pp. 277-286.
15. Turkan, M.; Pard´asand, M. & Etin, A. C, Human eye localization using edge projection. *In* the International Conference on Computer Vision Theory and Applications, 2007, **1**, pp. 410-415.
16. Bai, L.; Shen, L. & Wang, Y. A novel eye location algorithm based on radial symmetry transforms. *In* the International Conference on Pattern Recognition, 2006, **3,** pp.511-514.
17. Campadelli, P.; Lanzarotti, R. & Lipori, G. Precise eye localization through a general-to-specific model definition. *In* the British Machine Vision Conference, 2006, pp.1-20.

18. Hamouz, M.; Kittlerand, J.; Kamarainen, J.K.; Paalanen, P.; Kalviainen, H. & Matas, J. Feature-based affine-invariant localization of faces. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2005, **27**(9), 1490-1495.
19. Blaschko, M.B. & Lampert, C.H. Learning to localize objects withstructured output regression. *In* the European Conference Computer Vision, 2008, pp. 2-15.

Contributors

Mr Gagandeep Singh received his BTech Computer Sci. Engg.), Indian Institute of Technology, Patna in 2012. He is currently pursuing his master's in computer science in Software Reliability Lab at ETH Zurich. He received President of India Gold Medal for securing highest CPI in BTech programme at IIT Patna.

Dr Maheshkumar H. Kolekar received his PhD (Electronics and Electrical Communication Engg.) from Indian Institute of Technology Kharagpur in 2007. Presently working as Assistant Professor in Electrical Engineering, Indian Institute of Technology Patna. His research interests are in digital image and video processing, video surveillance and medical image processing.