A Novel Approach for Object Detection in VHR Images

Ashwini Kunte[1], Siddhesh Shirodker[2], Rakesh Menaria[3] and Bhavesh Patel[4]

Department of Electronics & Tele-communication, Thadomal Shahani Engg. College, Mumbai, INDIA
[1]askunte@gmail.com, [2]siddhesh.shirodker@gmail.com, [3]menariarakesh@gmail.com, [4]bhaveshpatel1712@gmail.com

Abstract

This paper presents two methods for buildings extraction in Very High Resolution (VHR) remotely sensed multispectral (4 band) images based on supervised and unsupervised segmentation using different image properties. The proposed approach for unsupervised or automatic building detection involves four stages, primarily, filtering to smoothen and enhance the objects present and sharpen the details. Secondly, a binary mask creation over which edge detection is applied. Edge linking is done to preserve information about the object. Lastly region properties like area, perimeter, etc are applied on the prepared mask and buildings are detected. The semi-automated or supervised method uses advanced color based segmentation algorithm to extract the buildings tops. This technique creates a number of masks based on segmentation and uses region properties based on color. Experiments are made on VHR images captured from satellites of commercial companies like Digital Globe and Geo Eye. Results of both methods are compared with respect to various accuracy measures at the end. The results illustrate that supervised algorithm using color property produces more accurate object delineation.

Keywords: *VHR images, Building Detection, Edge based Segmentation, Color based Segmentation*

1. Introduction

Object detection is described as the process that subdivides image into its constituent parts and extract those parts of interest. Segmentation forms a section of computer vision and is used when we need to automate the particular activity. The ultimate aim in an automated system is to extract important features from the image data, from which description, interpretation and understanding of scene can be provided by machine.

Object Detection is one of the most important application in various fields namely Remote Sensing, Military, Medical Imaging, Machine vision and many more. With accelerated development of remote sensing technologies, we have proposed a novel approach for Building Detection from Aerial Imagery. The use of high resolution cameras installed on satellites like IKONOS, Worldview, QuickBird, GeoEye-1 and 2 provided by commercial companies like DigitalGlobe, GeoEye and RapidEye having resolution ranging from 0.6 meters to 4 meters. This has created the need to detect objects for variety of applications like as site selection, flood management, military, urban and rural mapping of natural resources and of natural disasters, tax mapping, agriculture and forestry analysis, mining, engineering, construction, and change detection [1]. These satellites transmit thousands of images per unit time, and there is a need to process these images automatically without human intervention.

Development has concentrated on two methods, Automatic (Unsupervised) and Semi-Automatic (Supervised). In both these approaches we first perform pre-processing techniques like filtering to reduce noise, then extract high level features and finally use region properties to detect the buildings. The whole process of Automatic Building Detection is divided into four small stages of pre-filtering, mask creation, further filtering and using region based properties. Preprocessing techniques to eliminate noise like trees, cars are done by Gaussian

and Order Statistics filters. Post filtering we create a Binary Mask which is further processed to eliminate noise. Then, we apply Edge Detection algorithm along with edge-linking to create another mask. Using region properties like area, perimeter, eccentricity, etc. this mask is superimposed on the original image and building detection is performed.

The work in this context were mostly preferred the data driven approach, which primarily relies on the extraction of high-level features like color. The extracted colors were the main source to a sequential grouping process that was used to achieve higher level features i.e. the entire structure. Filtering is done at every level to reduce noise. The grouping process was concluded with the generation of the highest level features, the rectangles or polygons, which were further processed using region properties to detect the object [2-5].

Due to complexity of object characteristic, the extraction of the field boundaries in high-resolution imagery requires sophisticated methods that are able to identify each areas of interest with exact boundaries as a single unit in spite of superfluous details and possibly low contrast to neighboring regions(for example: roads).The performance of each method is evaluated with four data sets. The results of performance assessment support the robustness of the approach developed.

2. Methodology

The main steps followed in the proposed methodology are given in Figure 1. As different images are fed to the system, by using the best suitable component (R, G or B) of image, usually R component, pre-processing filtering is done. Segmentation based analysis is used as pixel by pixel processing techniques like Classifiers give results with less accuracy. Algorithms mentioned use canny edge detection or L*a*b segmentation and accordingly objects are detected.

2.1. Unsupervised Method

Images have been taken from different satellites like IKONOS, Worldview, Quick Bird and Geo Eye. Images are filtered by various filters to smooth and enhance the objects present and sharpen the details. Unwanted objects such as trees are removed using filtering operation from binary image and edges are calculated using canny operator. As there are different classes of image so we can apply canny operator to each image by changing parameter of canny method for best suitable result. Canny parameter includes threshold and sigma (deviation) [3-7]. Edges obtained from canny method are then smoothed by edge linking process. After edges are obtained by canny method and linking buildings are located using geometrical properties of objects such as area, shape *etc*. Once object is detected boundary of object is traced and mask is created. Mask obtained is thus superimposed on original image and object is highlighted. According to the result number of buildings is calculated and compared with original no of buildings present in image. This method uses canny operator as edge calculator and geometrical property of an object to detect buildings.

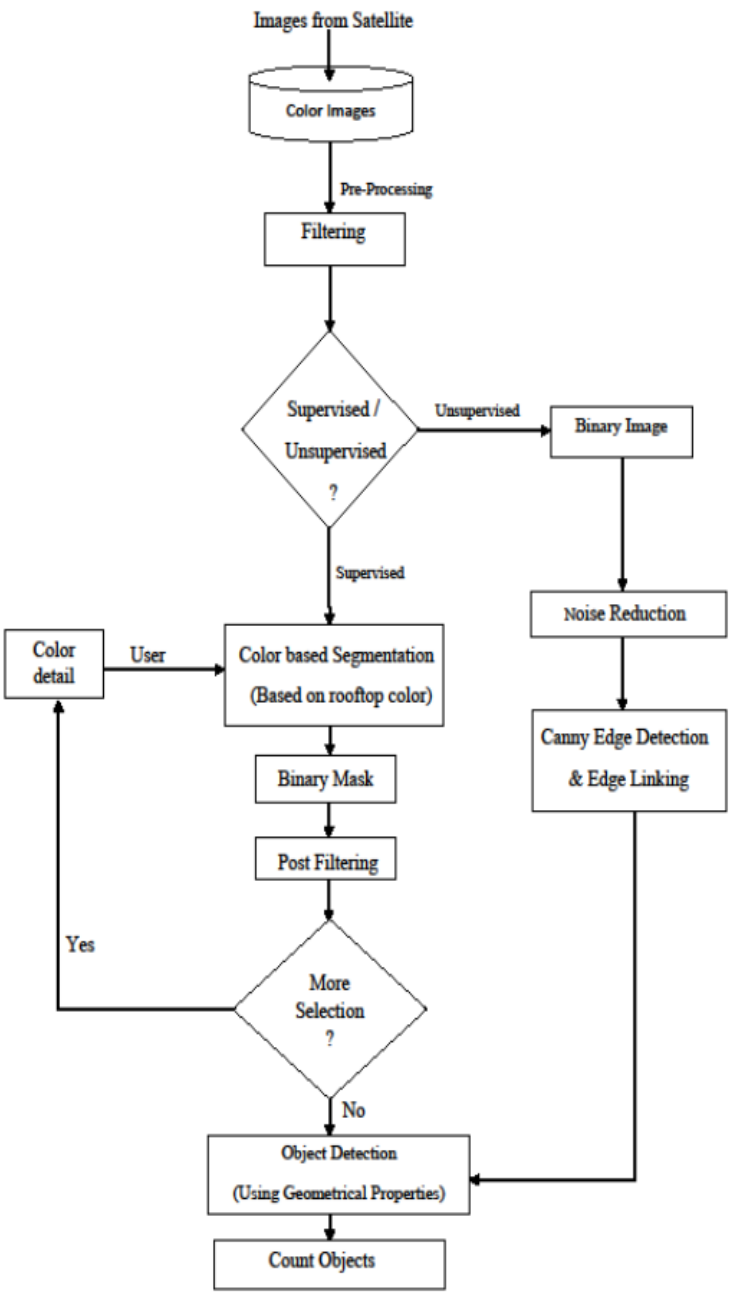


Figure 1. Flow Chart of Building Detection

2.2. Supervised Method

This method is initialized with the color segmentation algorithm in which irrelevant objects such as cars, small trees, chimneys, solar energy panels *etc.*, are removed. Color segmentation using the basic distance property like Euclidean distance (color difference dE) is used as there are different colors of roof tops. Then for each color, separate mask is created, following which they are combined and final mask is obtained. And hence building detection is done [9-12].

3. Implementation

This part consists of implementation of the two different approaches to detect objects (building) considering automatic and semi-automatic detection process using Image Processing Toolbox in Matlab. In the Automatic process, satellite images are fed to the system

and it gives the required result without any user's input while semiautomatic process deals with user's intervention to give some required parameter and then based on this information system calculates the objects. Stepwise implementation of both the algorithms is given below.

3.1. Algorithm 1: Edge Based Segmentation

Step 1: Feed image to the system.

Figure 2. Original Image

Step 2: Pre-processing of image: Images are filtered by various filters to smooth and enhance the objects present and sharpen the details.

Figure 3. Pre-filtered Image

Step 3: Binary image is formed and noise present in object is removed by filling holes.

Figure 4. Intermediate Filtered Image

Step 4: Unwanted objects such as trees are removed using filtering operation from binary image and edges are calculated.

Figure 5. Final Filtered Image

Step 5: Objects *i.e.*, buildings are located using geometrical properties of objects such as area, shape *etc*.

Figure 6. Edge Detected Image

Step 6: Objects are calculated and results are compared.

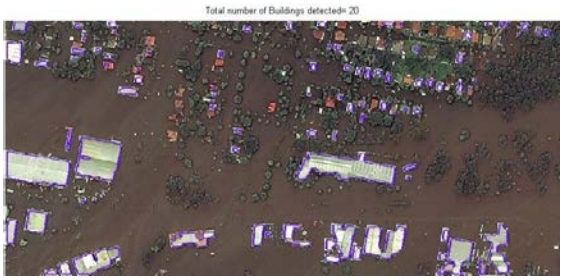


Figure 7. Final Result of Algorithm 1

3.1. Algorithm 2: Color based Segmentation

Step 1: Feed image to the system.

Figure 8. Original Image

Step 2: Image is segmented using color based segmentation and buildings are located based on color of rooftop.

Figure 9(a). Colour Based Segmented Mask 1

Figure 9(b). Colour Based Segmented Mask 2

Step 3: Post processing: Binary masks are created which are further filtered to reduce noise.

Figure 10(a). Post-filtered Mask 1

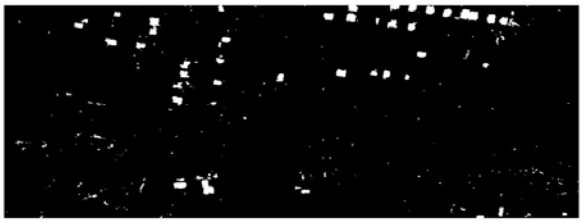


Figure 10(b). Post-filtered Mask 2

Step 4: The masks are logically added depending on the user input initially. Boundary of building is detected using region properties and is highlighted.

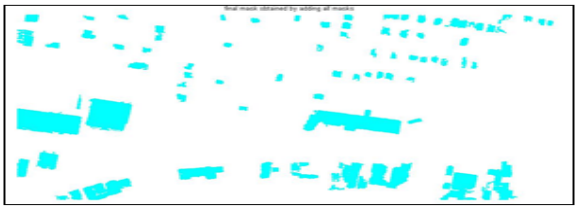


Figure 11. Addition of All Masks

Step 5: Count the number of objects and compare the results with actual number of objects present in image and result is compared with that of algorithm 1.

Figure 12. Final Result of Algorithm 2

Result of both algorithms is shown in Figures 13 and 14 and results of various data sets are compared in Tables 1 and 2.

4. Result and Analysis

The following accuracy measures proposed are used to test the quality of the results of the proposed methodology [1, 2]. Five different measures are computed:

- Detection Percentage = 100 x TP / (TP + TN)
- Branch Factor = 100 x FP / (TP + FP)
- Correct Building Pixels Percentage
- Incorrect Building Pixels Percentage
- Correct Non-building Pixels Percentage

In the first two measures, TP (True Positive) denotes the number of buildings that exist in the reference vector dataset and detected by the method, TN (True Negative) denotes the number of buildings that only exist in the reference vector dataset but not detected by the method, and

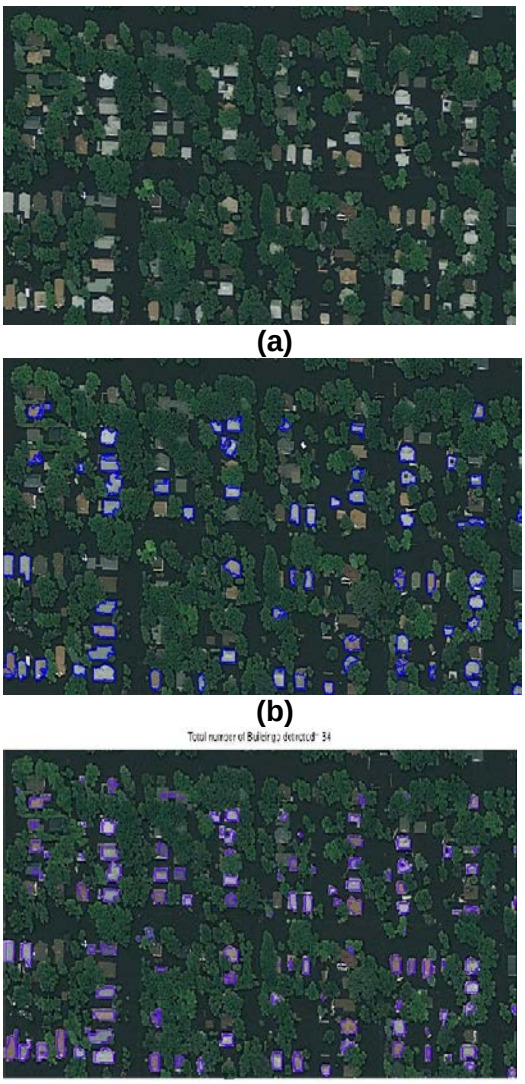


Figure 13. Results of Test Image2. (a) Image2 (b) Result of Algorithm1 (c) Result of Algorithm2

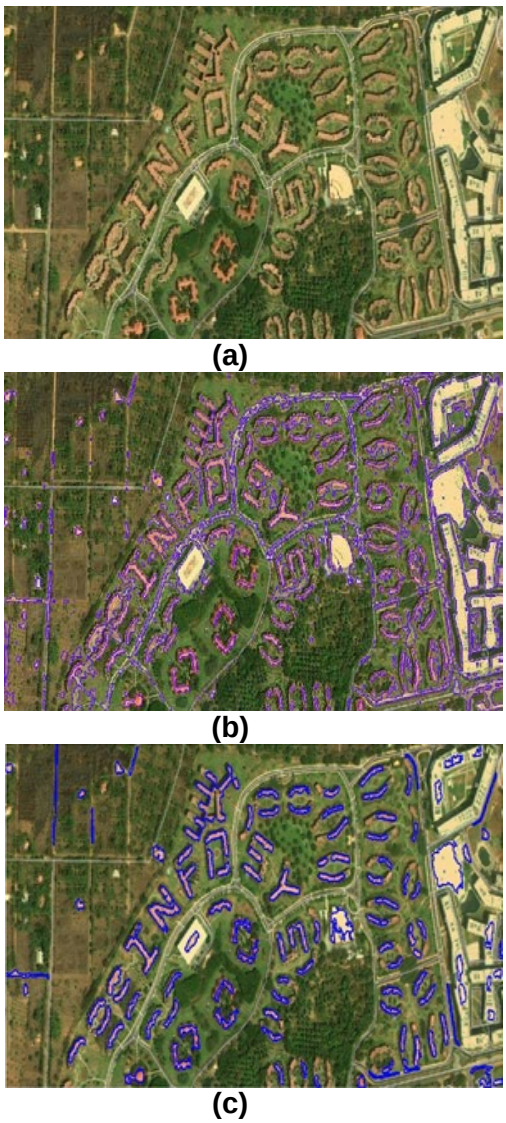


Figure 14. Results of Test Image3. (a) image 3 (b) Result of Algorithm1 (c) Result of Algorithm2

(a)

(b)

(c)

Figure 15. Results of Test Image 4. (a) Image 4 (b) Result of Algorithm1 (c) Result of Algorithm2

FP (False Positive) denotes the number of buildings detected by the method but not exist in the reference vector dataset. The last three measures are determined by counting the correct building and non-building pixels [13, 14].

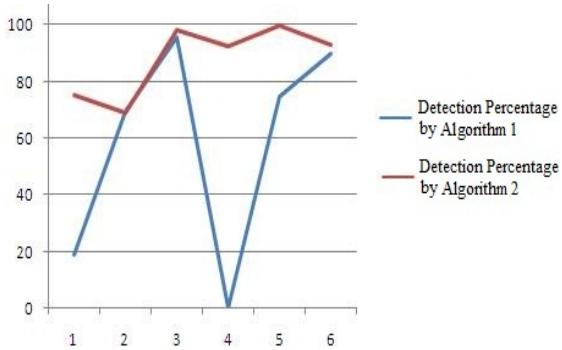


Figure 16. Comparison of Detection Rate of Algorithm 1 and 2

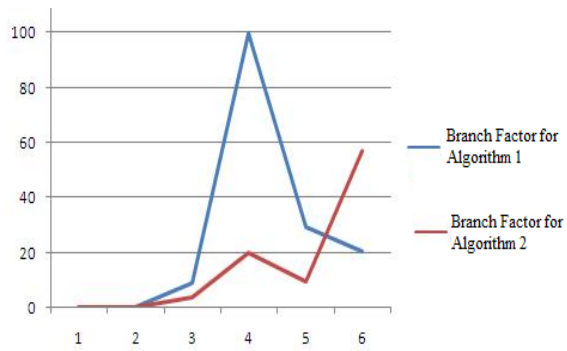


Figure 17. Comparison of Branch Factor of Algorithm 1 and 2

Table 1. Evaluation of the Edge Based Detection Algorithm 1

Test Image	Original No of Building	Result	TP	TN	FP	Detection Percentage	Branch Factor
1	105	20	20	85	0	19.05	0
2	96	67	67	29	0	69.79	0
3	115	127	116	5	11	95.87	8.66
4	233	28	0	233	28	0	100

Table 2. Evaluation of the Colour Based Segmentation Algorithm 2

Test Image	Original No of Building	Result	TP	TN	FP	Detection Percentage	Branch Factor
1	105	79	79	26	0	75.24	0
2	96	66	66	30	0	68.75	0
3	115	116	112	2	4	98.25	3.45
4	233	193	155	12	38	92.81	19.69

From Table 1 and Table 2, we observe that color based algorithm performs better then edge based detection algorithm. Detection percentage signifies the accuracy of system which should be high in order to yield good response of system. Though canny operator performs very well in certain cases, more processing is required to eliminate noise.

We have seen that for test image 4, edge based algorithm gave 0% detection whereas color based algorithm gave far better result. This is due to fact that in algorithm 2 we have considered color property and it is unique for a particular object. Algorithm 1 does not incorporate color property so it considers a group of different connected object as one object and gives the count and hence the error. But if a selected area contains some noise, then mask so obtained will also be having same noise which is present all over the area e.g. if roof top of a particular building contains green color objects like plants, then mask generated will also be containing that noise. So user must be very careful while selecting an area of interest. Test image 3 is an example of such case. Selection of the threshold point (Tolerance) in algorithm 2 is the key factor to obtain good result.

There are two types of histogram. One of the original image and second of the mask in which objects of interest are included. We should select a point in which more amount of area of mask is there instead of whole image. If two areas are sufficiently distinct then we get good mask of interested object and less noise. If two areas overlap to certain extent then we must select a point such that there are fewer amounts of pixels corresponding to original histogram and more pixels corresponding to selected mask's histogram. Extreme case is when both areas are merging in each other and there is no optimum point available to select a point. In such cases we better not select any of the point and reconsider our point of selection.

Results may vary slightly if selection of mask is different from the previous selection. So in brief we can summarize the selection criteria as follows:

1. Area of interest must be as large as possible.
2. Area selection must be done carefully as selection of unwanted area results into large variation (standard deviation varies) of final output.
3. More than one different sample of a particular object must be chosen.
4. Selection should not include too many numbers of samples though it increases precision but it also increases processing time.

We can observe that from Figures 16 and 17 that color based segmentation improves edge based detection to a greater extent. Also detection of false positive is also reduced tremendously. But at the end there is still small chance of getting false object detection due to presence of noise. In short we can conclude that color based segmentation performs better in many cases.

5. Conclusion and Future Scope

In this paper two methods are introduced, supervised and unsupervised method to detect buildings in aerial and satellite images. Algorithm 1 uses the individual spectral component of the image for detection. Algorithm 2 uses color based segmentation algorithm is used to segment the entire image. A pre-selection reduces the feature extraction to limit noise. The clustering is based on robust features like color. To evaluate the building detection approach, it is tested on a high resolution images captured by sensors like GeoEye2.

As seen in the supervised method, we have to input the choice of segmentation based on color by first examining the image and provide a threshold to create the masks. On comparing both these methods, algorithm 2 yields better result compared to the algorithm1 even in the presence of noise. Color based segmentation algorithm consumes more time than the edge based method and can prove a costly affair in terms of time to segment images.

This paper can be further enhanced by using more complex techniques like using more number of features like texture along with the existing framework. Shadow analysis has always been considered to be one of the most important clues to find elevated structures, especially buildings. By a combination of these two methods along with more complicated processes in a dynamic way can result more robust building detection.

References

[1] C. Özgün Ok, “Robust Detection of Buildings from a Single Color Aerial Image”, Dept. of Geodetic and Geographic Information Technologies, Middle East Technical University, Turkey. Commission VI, WG VI/4.
[2] S. Muller and D. Wilhelm Zaum, “Robust Building Detection in Aerial Images”, U. Stilla, F. Rottensteiner, S. Hinz (Eds) CMRT05. IAPRS, Vienna, Austria, vol. 36, Part 3/W24, **(2005)** August 29-30.
[3] E. Nadernejad, S. Sharifzadeh and H. Hassanpour, “Edge Detection Techniques: Evaluations and Comparisons”, Applied Mathematical Sciences, vol. 2, no. 31, **(2008)**, pp. 1507-1520.
[4] J. Canny, “A computational approach to edge detection: Pattern Analysis and Machine Intelligence”, IEEE Transactions on, PAMI, vol. 8, no. 6, **(1986)** November, pp. 679-698.
[5] R. Gonzales and R. Woods, “Digital Image Processing”, 2nd Edition, Pearson Education, **(1984)**.
[6] L. Ma, X.-Ping Zhang, J. Si and G. P. Abousleman, “Bidirectional Labeling and Registration Scheme for Grayscale Image Segmentation”, IEEE Transc., vol. 14, no.12, **(2005)** December.
[7] T. Pavlidis and Y. T. Liow, “Integrating Region Growing and edge Detection”, IEEE Transc. On Pattern Analysis and Machine Intelligence”, vol. 12, no. 3, **(1990)** March, pp. 225-233.
[8] D. Pham, C. Xu and J. Prince, “A Survey of current Methods in Medical Image Segmentation”, Annual Review of Biomedical Engineering, **(1998)** January 19.
[9] B. Diaz Acosta, “Experiments in Image Segmentation for Automatic US Licence Plate Recognition”, Virginia Polytechnic Institute and State University, **(2004)** June 18.
[10] F. W. Billmeyer and M. Saltzman, “Principles of Color Technology”, 2nd ed., New York: Wiley.
[11] B. Amanpreet, S. Neeru and K. Maninder, “Color Image Segmentation Based on Color and Texture properties”, International Journal of Advanced Engineering Sciences and Technologies, vol. 8, no. 2, pp. 152-156.
[12] M. A. Hunt, J. S. Goddard and K. W. Hylton, “Imaging tristimulus colorimeter for the evaluation of color in printed textiles”, Oak Ridge National Lab, Oak Ridge, TN, 37831.

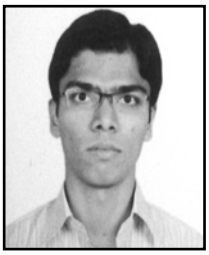
[13] C. Lin and R. Nevatia, **(1998)**, "Building Detection and Description from a Single Intensity Image", Computer Vision and Image Understanding, 72(2), pp. 101-121.
[14] B. Sırmaçek and C. Ünsalan, "A Probabilistic Satellite Images", IEEE Transaction on Geoscience and Remote Sensing, vol. 49, no. 1, **(2011)** January.

Authors

Ashwini Kunte received PhD degree from Government College of Engineering, Aurangabad, India. She is Associate Professor and Head of Department of Electronics and Telecommunication in Thadomal Shahni Engineering College, Mumbai, India. Her research interests are Signal and Images processing, Wireless networks. She has 10 publications in International conferences and journals to her credit.

Siddhesh Shirodker received the B.E. degree in Electronics and Tele-Communication from University of Mumbai, India in 2012. He is currently working in Accenture India as a Software programmer. His area of interest lies in Signal and Image Processing and intends to do his Masters in that field.

Rakesh Menaria received the B.E. degree in Electronics and Tele-Communication from University of Mumbai, India in 2012. He is a web application developer in IBM India Pvt. Ltd. His area of interest lies in application development and embedded programming.



Bhavesh Patel received the B.E. degree in Electronics and Tele-Communication from University of Mumbai, India in 2012. He is a Java developer in Accenture India. His area of interest lies in application development and Image processing.