A survey on depth based 3D image Modeling

Dr. Dayanand G Savakar [1], Ravi Hosur [2]

[1]Associate Professor, Department of Computer Science
PG Centre, Rani Chennamma University, Torvi, Vijayapur(Karnataka), India
[2]Assistant Professor, MCA Department
BLDEA's V. P. Dr. P. G. Halakatti college of Engg. & Tech., Vijayapur(Karnataka), India
[1]dgsavakar@gmail.com, [2] hosuravi@gmail.com

***Abstract*— Depth based 3D image modeling is a biggest challenge that persists in the digital image processing depending on different depths. The method of recognizing in dynamic areas is an unanswered challenge facing the community of robotics. The answer can be framed by including the methods like segmentation, tracking, and classification components, etc. This paper reveals the survey of the works carried on the construction of 3D model with their own method and technologies applied in the process. The paper mainly goals the Reality systems to enhance the techniques used for the 3D image construction and also to segment the blurry areas in the image captured from different depths. Also the image is desirable to recognize 3D-objects of the system in the user's environment, in order to avoid manual based method. Many existing systems are surveyed which perform 3D-object model construction.**

Keyword- Depth, 3D image, modeling, Reconstruction, Acquisition

I. INTRODUCTION

Depth based 3D-object construction and its recognition is the most challenging area of research in the DIP (Digital Image Processing). The goal of 3D modeling is to correctly identify objects of different pose when given to a system for identification that are present in a 3D scene of a depth present with range image, and also to be used in the security systems and the military systems, and also intensity data systems, range data systems, and both range and intensity systems (sometimes including color). In cases like the intensity-image systems, points and straight line segments are the considered common features. Recently, the alignment method blindly matches the points or line segments thrice from the image with the similar of the model, using little or no contextual information. These all inventions lead to the assumptions that the points or line segments are the reliable features of the class of the identified object to be recognized or located and (b) the view of the object is determined can be unique using these features. They lead to the view of polyhedral objects or those modelled images with sharp and straight edges. Hence they are categorized under curved-surface objects.

Three-dimensional reconstruction of objects has become an important technique in the context of inspection: Industry quality inspection – for observing the quality of a product from different views, Archaeology analysis – reconstruction of an object from raw data obtained from muddy area, Medical science - for conduction sessions of teaching, and for reconstruction body check up from collected data like details obtained from body check up, Forensic analysis – reconstruction of a scene from raw data for identifying unseen actions**.**

II. LITERATURE SURVEY

(Jose L. Herrera Carlos R, et al., (2015)) proposes a method for constructing 3D using 2D by using a technique centroid clustering process by combining all the depth images in the cluster.

(Zhongjun Wu, et al., (2015) proposes a Multi-Depth Generic Elastic Models (MDGEMs) for the construction of 3D face modeling by using varying multiple depth maps of an image.

(Yan Zhou, Huiwen Guo, et al., (2015)) proposes an algorithm for reconstructing a 3D view using a wavelet transform and Support Vector Machine (SVM) models to get the image focus quality and Mean shift algorithm for the depth scene. The results prove to be improving the efficiency better compared to previous techniques.

(Masaya Kaneko, et al., (2015)) proposes a method for detecting the edge by generating a decision tree, a solution 25 times better to a conventional method existing by training the decision tree using supervised learning method in the process

(Xuyuan Xu, et al., (2014)) recovers the object boundaries of captured depth images with sharp that refine the use of adaptive block by reallocation of the same position and expand to increase the depth accuracy by avoiding false depth boundary refinement.

(Er. Mahendra Singh, et al., (2012)) proposes a technique of Steganography using JPEG images based on 2D block DCT that transform cover image blocks from spatial domain to frequency domain. The method brought up with 54 bits secret block embedding as information bits and hence a cover image of 417×417 pixels can embed 146718 secret bits into it. The experimentation resulted in acceptable image quality and a large message capacity.

(Shahram Izadi, et al., (2011)) proposes a model for 3D reconstruction of an indoor scene using a movable standard Kinect system. The scenes are captured with the help of low-cost scanning, and geometry augmentation and related based interactions. Using the techniques point-plane data association TSDF integration access the scenes are projected with assumed threshold and then truncated with maximum and minimum thresholds. The extraction is done using raycasting and the overlapping is tested using a touch-map technique.

(M. Alptekin Engin, et al., (2011)) proposes a method for JPEG based on 3D formation of original image by spiral order scanning and the application of 3D discrete cosine transformation. In the process 3D image is represented in 1D as a spiral ladder for a better efficiency in coding. The technique on experimental results proved that the method has better performance than JPEG compression in low and high bit rates of background images. The image is deployed in a cube of 8×8×8.

(Prabin Bariya, et al., (2010)) proposed a model for identifying 3D images in concern to scale variability among existing geometric structures. The method adds extra information to enrich the variabilities that occur in the structures so that the inferior elements are being hierarchical coarse to fine manner. After a series process of geometric space evaluation, interpretation tree; the fine tuning of few adjustable nodes is carried to evaluate occlusion and clutter factors. This resulted in 97.5% of recognition rate with upto 84% occlusion.

(Kenneth Weiss, et al., (2010)) proposes a method for extracting crack free approximations of 3D image at any uniform or variable resolution and results in reduction of size of data set without fall of accuracy. In the process the image is discreted into mesh resolution and using cumulative distribution function the vertices labelled of each are distracted to approximate intensity. The method proves that it is possible to analyze the image using much less samples then full image.

(Federico Tombari, et al., (2010)) proposes a Hough voting technique for detection of free-form shapes in a 3D space for object recognition, filled with occlusion and clutter. In the process interest points are extracted, circled as identification for storing description and then model the points. After neighbourhood computation, transformation with rotation and transformations are also calculated to highlight the presence of objects by thresholding the peaks of Hough space which increases the accuracy of recognizing 3D objects in a very noisy 3D data in a real-time stereo setup.

(Doaa Hegazy, et al., (2009)) has proposed a technique for recognizing images from range distances using Time of Flight cameras where the images can be of low resolution or of noise content. The technique has been built with shape-specific local feature histograms because of the direct information provided by range images that added features like pixel depth, surface normal and curvature in turn helped for 3D recognition. The work resulted in a database for generic 3D object and a model for its recognition where the learning process is carried by using RealAdaBoost algorithm.

(Min-Wen Chao, et al., (2009)) proposes a 3D steganography scheme with multilayered embedding scheme to hide secrete messages in the vertices of 3D polygon models. The experimentation leads to the optimization of partitioning intervals to obtain a good balance between distortion and capacity, also it resulted in the method can provide much higher capacity for medium/big sized models which produced almost similar the cover and sego images.

(Chen Tsung Kuo, et al.,, (2009)) proposes a watermarking method on 3D mesh models in which the watermark is hidden within the 3D graphical object by modifying the subset of carefully selected edge vertices using Moment-Preserving principle. The extraction process is done depending on the strength of the edge being hashed into buckets of a hash table where with the closest edge strength are mapped to the bucket. The algorithm has been tested with scaling down X-axis, up Y-axis, along Z-axis and mixed scaling in X, Y and Z-axes of different scaling factors and resulted in robust nature with transparency against a wide variety of attacks.

(Jiang Ze-tao, et al., (2009)) proposes an automatic 3D reconstruction technique based on images. With the extraction of feature points and accordingly extract the coordinate points in 3D, using which fundamental and essential matrix using extrinsic and intrinsic parameters of camera is calculated. Reconstruction is done by Delauny triangulation, and texture mapping and rendering methods during late stages. Experimentation proved that the technique consumes fewer tines for reconstruction of large-scale virtual scene in concern to efficiency, stability and robustness.

(Hui Chen, et al., (2009)) proposes a method for rapid indexing and recognition of highly similar objects from a database by combining the feature embedding on surface descriptors. The method is developed with a support vector machine-based learning technique for ranking the hypotheses by using FastMap embedding algorithm. The geometric hashing technique helps in comparing the data from the database collection to demonstrate the

efficiency and effectiveness of the method whereas SVM method is used training and testing the dataset values. The experimental results proved that the method is fast and robust in indexing with performance being efficient for searching the nearest neighbours.

(Xinju Li, et al., (2007)) proposes a method to recognize 3D range images by using local surface descriptors without using surface modeling. The data is converted initially into patches according to detected features like salient points and local signatures for analysis using pyramid kernel function and the classification is done using SVM for comparing with new data for recognition. The experimentation resulted that the method is invariant to 3D transformation and robust to occlusion.

(Hui Chen, et al., (2007)) proposes a model for an integrated local surface descriptor by identifying the points with their features of large shape variations and hence, avoids the calculations of all points. After calculation on every local surface patch, they are indexed into a hash table, on the match similarity of patches of surfaces the performance verification is done by the Iterative Closest Point (ICP) algorithm. Experimental results measured efficiency and effectiveness with spin and spherical image representation by a factor of 3.79 to 4.31 a model-test pair.

(J.L.Toutant, et al., (2006)) proposes an embedding technique in an image invisibly and provides an investigation about the noise induced in images and minimize the same. With the implementation of JPEG and data hiding methods with adapted quantization, resulted in 100% and 50% quality factor.

(Rajasvaran Logeswaran, et al., (2005)) proposes a method using scale space for analyzing the characteristics like multi-resolution within 3D images, and constructing a 3D image from a series of 2D images. By applying the regular image processing steps like analysis, segmentation, verification, etc. are followed for constructing or reconstruction of 3D images from 2D images where initially the image taken is blurred using a filter with a decreased coefficient value at each coarseness. The experimental results proved that the reconstruction loses the actual originality due to noise and partial effects, reflects in the quality of image.

(Massimiliano Corsini, et al., (2004)) proposes an idea of extending perceptual image watermarking to 3D watermarking where the analysis leads to reliable data as a result. In the process 3D impairment predictions have been done by combining roughness estimation with subjective data.

(Guoping Qiu, (2004)) proposes an image coding method for image based content retrieval with very little computation by segmentation, and channel decomposition to represent spatial patterns as chromatic and achromatic. The image descriptor indexes the region size and pattern. The experimentation resulted that in content based image retrieval it is easy to compare the states like color correlogram and latest MPEG7 color structure, and also it is the possibility of developing compressed image coding techniques.

(Christian Cachin, (2004)) proposes a passive adversary model to yield into a secure steganographic scheme through conditional satisfaction between the cover text and modified text. The stegosystem has been designed by using their own model using secret-key system for embedding the key with message at the sender and then during extraction at the receiver. The experimentation proved that the steganography results are highlighted if-and-only if it is treated to be a problem of hypothesis testing so that the receiver detects the presence of hidden message.

(Delia Soimu, et al., (2004)) proposes a solution for the problems affected to the quality of 3D images during reconstruction from projections by using Feldkamp method. The images taken on cone beam scheme for larger cone angles, approximations have been taken with severe artifacts where the geometry of scan arc length and number of projection acquired influences the quality of reconstructed images. Back propagation cannot be applied for such cases with missing data in projection.

(Remo Ziegler, et al., (2003)) proposes a method for reconstructing 3D scenes from a set of images using 2D photo-editing tools. The results shown that the Cell Carving algorithm computed the 3D model with maximum volume, and is faster in using 2D intersection operations for computing a polygon module. The user interfaced experiments were carried in visual basic and the algorithm in C++, rendered the scenes from six viewpoints using 3D StudioMax.

(Zhiqiang Yu, et al., (2003)) proposes a robust 3D watermarking scheme for the triangle meshes for perturbing the distance between the vertices and the centre of the model that preserves the visuality of the models and intern the strength of the embedded watermark signal. The mesh has been checked with all parameters like weight, controlling features like extraction and embedding, addition of noise to check for resistance, and also the combination of attacks like simplification, noise and cropping attacks to check the scheme for robustness. In experimentation due to the use of weighting scheme, the watermarking scheme is able to withstand common attacks on 3D models with all possible attacks.

(Gunter Hetzel, et al., (2001)) proposed a view based approach to recognize free-form objects in range images. The recognition has been made by using Histogram matching or probabilistic recognition algorithm and, has resulted in 93% accuracy on ideal images and 89% with 20% occlusion.

(Jessica Fridrich, et al., (2001)) proposes a method for detecting the least significant bit non-sequential embedding in digital images. The experimental results proved that 0.005 bits/pixel is safe for least significant bit embedding even for high quality images from scanners and digital cameras. If the bit rates are more than 0.005 then they are detectable.

(Mauro S. Costa, et al., (2000)) proposes a technique for 3D image recognition with a 2D intensity image with an edge comprising intensity based images pair constructed with a camera by relational indexing of new high-level features like points, line-segments, and ellipse-circles from each edge. A series of pose estimation can be framed for transformation from a 3D mesh object to the image is verified using 3D model with its wireframe. The effectiveness (distance error) is computed by Hausdorff-like between the obtained model and the edge image. Experimental results proved that the improvements can be made to the occluded images by masking process to make the verification model reliable.

(David G. Lowe, (1999)) proposed an object recognition system from scale invariant features like scaling, translation, and rotation. The recognition include generation of image key used as input to a nearest neighbor indexing method for identifying candidate object matches and then matching is done with the low residual least squares solution for unknown parameters being added by a SIFT key. The experimentation proved that the object recognition can be even achieved in case of cluttered or occluded images.

(Oliver Benedens, (1999)) proposes a watermarking algorithm for storing information in the model's geometry. Initially the properties like capacity and robustness of 3D watermarking are collected with the geometric operations that recommend speed of processing, embedding, and knowledge of prior data to be preprocessed. The carrying of transformations through core watermarking algorithm, consistent surface normal pitches are calculated with different degrees and radians of identified vertices in the geometry. These are implemented in C++ using the graphics package Modeling Animation Machine /Virtual Renderings System. The technique proves that it is tolerable to private watermarks and the model achieved 36% resistance to robustness to simplification.

(D.M.Gravila, et al., (1992)) proposed a general technique for identifying an object in the 3D scene depending on its position and orientation, called Geometric Hashing. The technique focuses on invariant features that include occlusions in a scene, where 3D object recognition is done from 2D images. The recognition included with the processing steps like affine transformation in 2D (that include rotation, translation, scaling and shearing) for object recognition, Geometric Hashing for feature extraction, Complexity analysis for recognizing the number of hashing's required for each point when mapped, Error analysis for analysing for detection of errors due to quantization with hash table and measurement errors, then is the conversion from 2D to 3D. The results were compared with different techniques which lead to efficient results in identification after conversion from 2D to 3D.

(David Forsyth, et al., (1991)) proposes a visual technique for constructing the projection invariant descriptors by posing the images in 3-dimension. It is developed with the inclusion of model-based vision system to recognize curved plane as objects and also stabilization to measurement errors. On experimentation it is proved that the cross ratio is not only stable but also accurate, whereas in five and two coplanar conics invariants are stable and accurate.

Methodology

From the research carried on construction of a 3D object from 2D or from captured sequence of images reveals that the techniques make a common strategy of integrating the pixels of their nearness behaviour and, also mesh modeling, polygon construction, etc. by keeping a feature in consideration in concern to the construction of 3D image. But the works are showing it par in case of depth based captured images and also the 3D object rendering.

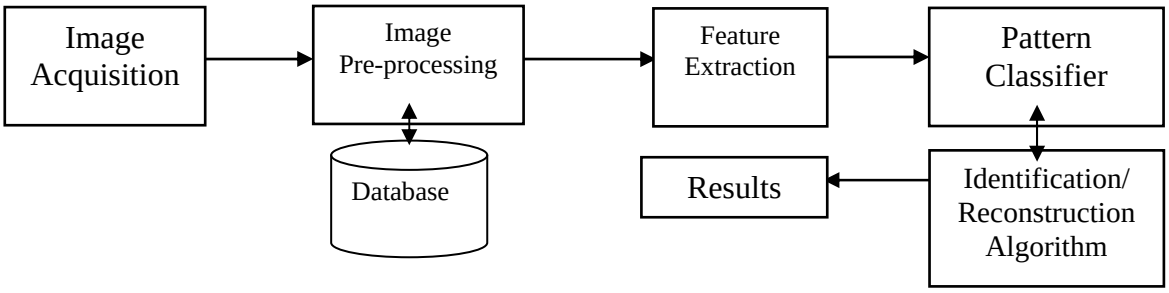

Figure 1: Process of 3D object Construction

Hence, there is a need of a method in reconstruction of a 3D object for the images captured with different depths so that the image identification leads in better results

III. CONCLUSION

Numbers of methods are proposed in the reconstruction of 3D images but still there exists a scope also for also constructing 3D images or objects that are with faster rendering type and of different depths with improved efficiency. So, introducing the method for construction with real time images of different pose can be made with different depth of 3D image is vital.

References

[1] Jose L. Herrera Carlos R. del-Blanco Narciso Garcia, (2015), "Edge based depth gradient refinement for 2D to 3D Learnt Prior Conversion", IEEE 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2015.
[2] Zhongjun Wu, Jiayu Li, Jiani Hu, and Weihong Deng, (2015), "Pose – invariant face recognition using 3D multi-depth generic elastic models", 2015 IEEE
[3] Yan Zhou, Huiwen Guo, Ruiqing Fu, Guoyuan Liang, Can Wang, Xinyu Wu, (2015), "3D reconstruction based on light field information", Proceeding of the 2015 IEEE International Conference on Information and Automation Lijiang, held on August 2015
[4] Masaya Kaneko, Takahiro Hasegawa, Yuji Yamauchi, Takayoshi Yamashita, Hironobu Fujiyoshi, and Hiroshi Murase, (2015), "Fast 3D edge detection by using decision tree from depth image", 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) Congress Center Hamburg Sept 28 - Oct 2, 2015. Hamburg, Germany, pp: 1314-1319
[5] Xuyuan Xu, Lai-Man Po, Chun-Ho Cheung, Litong Feng, Kwok-Wai Cheung, Chi-Wang Ting, Ka-Ho Ng, (2014), "Adaptive block truncation filter for MVA depth image enhancement", 2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)
[6] Er. Mahendra Singh, Er. Rohini Sharma, Er. Dinesh Garg, (2012), "A New Proposed Issue for Secure Image Steganography Technique Based on 2D block DCT and DCT ",International Journal of Advanced Research in Computer Science and Software Engineering, Vol. No. 2, pp: 29-33
[7] Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew, Davison, Andrew Fitzgibbon, (2011), "Kinectfusion: real-time 3D reconstruction and Interaction Using a Moving Depth Camera", UIST'11
[8] M. Alptekin Engin, Bulent Cavsoglu, (2011), "New Approach in Image compression:3D spiral JPEG", IEEE communications Letters, Vol. No. 15, Issue No. 11, pp:1234-1236
[9] Akhil Khare, Meenu Kumari, Pallavi Khare, (2010) "Efficient Algorithm For Digital Image Steganography", Journal of Information, Knowledge and Research In Computer Science and Applications, Vol. 1, Issue No. 1, pp: 1-5
[10] Prabin Bariya, Ko Nishino, (2010) "Scale-Hierarchical 3D Object Recognition in cluttered scenes", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp: 1657-1664
[11] Kenneth Weiss, Leila De Floriani, Mohammed Mostefa Mesmoudi, (2010), "Multiresolution analysis of 3D images based on Discrete Distortion", 2010 International Conference on Pattern Recognition, pp: 4093-4096
[12] Federico Tombari, Luigi Di Stefano, (2010), "Object recognition in 3D scenes with occlusions and clutter by Hough voting", 2010 Fourth Pacific-Rim Symposium on Image and Video technology, pp:349-355
[13] Doaa Hegazy, Joachim Denzler, (2009) "Generic 3D object recognition from time-of-flight images using boosted combined shape features", Proceedings of VISSAPP (2), pp: 321-326.
[14] Min-Wen Chao, Chao-Hung Lin, Cheng-Wei Yu, Tong-Yee Lee, (2009), "A high capacity 3D Steganography Algorithm", IEEE transactions on Visualization and Computer Graphics, Vol. No. 2, pp:274-284
[15] Chen Tsung Kuo, Shyi-Chyi Cheng, Da-Chun Wu, Chin-Chun Chang, (2009), "A blind robust Watermarking Scheme for 3D Triangular Mesh Models Using 3D edge vertex detection", Asian Journal of Health and Information Sciences, Vol. No. 4, Issue No. 1, pp:36-63
[16] Jiang Ze-tao, Zheng Bi-na, Wu Min, Wu Wen-huan, (2009), "A fully automatic 3D reconstruction Method based on Images", 2009 World Congress on Computer Science and Information Engineering, pp: 327-331
[17] Hui Chen, Bir Bhanu, (2009), "Efficient recognition of Highly Similar 3D objects in range Images", IEEE transactions on pattern analysis and machine intelligence, Vol. No. 31, Issue No. 1, pp:172-179
[18] Xinju Li, Igor Guskov, (2007), "3D object recognition from range images using pyramid matching", IEEE
[19] Hui Chen, Bir Bhanu, (2007), "3D free-form object recognition in range images using local surface patches", Pattern recognition letters, Elsevier pp:1252-1262
[20] J.L.Toutant, W.Puech, C.Fiorio,(2006), "Minimizing data-hiding Noise in Color JPEG Images by Adapting the Quantization", European Conference on Colour in Graphics, Imaging, and Vision, pp. (5 pages)
[21] Rajasvaran Logeswaran, Chikannan Eswaran, (2005), "3D scale space for structure analysis of 2D image series", 2005 IEEE International Conference on Industrial Informatics (INDIN), pp:502-505
[22] Massimiliano Corsini, Elisa Drelie Gelasca, Touradj Ebrahimi, (2004), "Watermarked 3D object quality Assessment", ITS technical report No. 29 , pp: 1-28
[23] Guoping Qiu, (2004), "Embedded colour image coding for content-based retrieval", Elsevier, pp: 507-521
[24] Christian Cachin,(2004),"An Information-theoretic Model for Steganography", Springer, pp:1-14
[25] Delia Soimu, Cristian BADEA, Nicolas PALLIKARAKIS, (2004) "Image Quality in Cone-Beam 3D reconstructions using Feldkamp algorithm: Acquisition Problems", proceedings of "Medicon and Health Telematics 2004" conference
[26] Zhiqiang Yu,Horace H.S.Ip,L.F.Kwok,(2003),"A robust watermarking scheme for 3D triangular mesh models",The journal of pattern recognition society, pp:2603-2614
[27] Gunter Hetzel, Bastian Leibe, Paul Levi, Bernt Schiele,(2001) "3D object recognition from range images using Local Feature Histograms", Proceedings of the 2001 IEEE Computer Society Conference on CVPR 2001, vol. No. II, pp:394-399
[28] Jessica Fridrich, Miroslav Goljan, Rui Du, (2001), "Reliable detection of LSB Steganography in color and Grayscale Images", MM&Sec '01 Proceedings of the 2001 workshop on Multimedia and security: new challenges, pp: 27 – 30
[29] Mauro S. Costa, Linda G. Shapiro, (2000), "3D object recognition and pose with relational indexing", Computer Vision and Image Understanding, Vol. No. 79, pp:364-407
[30] David G. Lowe, (1999) "Object recognition from Local Scale-Invariant features", proc. Of the International conference on computer vision, Corfu, pp: 1-8
[31] Oliver Benedens, (1999), "Geometry based Watermarking of 3D Models", IEEE Journal of Image Security, pp:2-11
[32] D.M.Gravila, F.C.A. Groen, (1992) "3D object recognition from 2D images using geometric hashing", Elsevier Science Publishers, pp: 263-278

[33] David Forsyth, Joseph L. Mundy, Andrew Zisserman, Chris Coelho, Aaron Heller, Charles Rothwell, (1991), "Invariant descriptors for 3D object recognition and pose", IEEE transactions on pattern analysis and machine intelligence, Vol. No. 13, Issue No. 10, pp:971-991

AUTHORS PROFILE

Dr. Dayanand G. Savakar (Date of birth 02-10-1968), having 26 years of teaching and 12years of research experience, he has completed his B.E degree from Karnataka University Dharwad, Post Graduation from Birla Institute of Technology and Ph.D from Visvesvaraya Technological University., Belgaum, India. He is now working as Associate Professor, in the Department of Computer Science, Post Graduate Centre, Rani Chennamma University, Vijayapur-Karnataka, India. He has published more than 35 research articles in international journals/conferences. Currently he is guiding eight Ph.D candidates. He served/serving as a member of various boards of several universities. His areas of interests are – Image Processing, Pattern recognition and Information security.

Mr. Ravi Hosur (Date of birth 22-08-1983), having 10 years of teaching and 3 years of research experience, he has completed his B.E degree from Visvesvaraya Technological University Belagavi, Post Graduation (M.Tech) from Visvesvaraya Technological University Belagavi and pursuing Ph.D from Visvesvaraya Technological University, Belagavi, Karnataka(India). He is now working as Assistant Professor, in the Department of MCA(Master of Computer Applications), BLDEA's Vachana Pitamaha Dr. P.G.Halakatti College of Engg. & Tech., Vijayapur-Karnataka, India. He has published more than 3 research articles in international journals/conferences. His areas of interests are – Image Processing, Pattern recognition and Information security.