



Query-Based Image Retrieval using Support Vector Machine (SVM)

Ajay G.¹, Abhishek Kumar^{2*}, Venkatesan R.³

¹Advanced Engineering Specialist, Accenture, Ireland, Europe

²Professor, Department of Computer Science and Engineering, JAIN University, Bangalore, 560069
India

³Associate Professor, Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Coimbatore, 641114, India

Emails: a.govindasamy@accenture.com; abhishek.maacindia@gmail.com; rlvenkei_2000@karunya.edu

Abstract

In many industries, the picture of the day now plays an important role in extracting information about the item. Many traditional image retrieval techniques have been used. It answers the user's question interactively by asking whether the image is relevant. In this digital age, graphics have become an important part of information processing. The image plays an important role in extracting information about the object in a variety of areas, including weather systems, tourism, medicine, and geology, in the processing of image registrations. There are several methods for retrieving images. It determines a person's inquiry interactively by asking users whether the image is relevant (similar). The efficient image database business has improved the process's functioning in the content-based image recovery system (CBIR). Content-based image recovery (CBIR) research has grown in importance. As individuals, we have studied and investigated various features in this manner or in combinations. We discovered that image Registration Processing (IRP) is a critical area in the industries. Several research papers on color feature and texture feature extraction were reviewed, and it was determined that point cloud data structure is best for image registration using the Iterative Closest Point (ICP) algorithm.

Keywords: CBIR, SVM, Feature Extraction, Point Cloud.

1. Introduction

Today's image processing plays an important role in image registration. In the 1990s, a new research field emerged: Content-based Image Retrieval, which aims to index and retrieve images based on their visual contents. It is also known as Query by Image Content (QBIC) and it presents methods for the visual organization of digital images. They are dedicated to solving the image retrieval problem in data bases. CBIR includes retrieving images from a graphics database to a query image. Similarity comparison is one of CBIR's tasks.

1.1 Image Registration Process

Image registration is a critical task in image processing that involves matching two or more images acquired at different time intervals from different sensors and perspectives. Almost all large industries that value images require image registration as a steadfastly associated operation. Identical for objective reorganisation real time images are targeted in literal guesses of anomalies where image registration is a major component.

There are two varieties of Image registration on 3-D datasets is done manually first, then automatically. Human operators are in charge of the entire process when manually registering. The relevant image characteristics should be documented. To obtain logically acceptable registration results, a user must select a significant number of feature pairs across the entire image set in succession. The manual procedure is not only uniform and laborious, but it also has discrepancy and limited precision. This is why there is a stringent standard requiring less time or no human operator supervision when developing automated registration methods [1].

Generally, there are four supporters' steps for image registration such as:

Feature Detection: It is an essential step for Image registration. It detect the features like closed-boundary fields, ranges, edges, outlines, line points going across, angles, and so on.

Feature Matching: In this step, Process of matching the features extracted by Database Images and Query image that yield to a result that is visually similar.

Transform model Estimation: A value is placed on the kind and parameters of the so-called mapping purposes, uses and matches the detected picture with the statement, direction image. The parameters of mapping purposes are determined using the indicated letters.

Image re-sampling and transformation: With the assistance of the mapping purposes, the detected picture is significantly altered. Image values in non-integer orders are developed via the appropriate art method for interpolation, etc.

1.2 Fundamentals of CBIR System

In traditional CBIR systems, the visual information of the images is retrieved from a database. It was also distinguished by multidimensional vectors. These attributes are a vector of the images in the database. Users provide an example image recovery system in order to obtain the images.

There are three main components used in this CBIR system such as:

1. User's Query, 2. Database Unit, 3. Retrieval system.

- **Query unit:**

The Query Unit selects a photograph from the database and extracts and stores three features as a function vector. Texture, colour, and form are the three characteristics. The colour moment may occur concurrently with the feature. By extracting the colour feature, the image can be transformed and partitioned into the grid. Some seconds have been removed. To extract the texture, the structured pyramid and tree-structured wavelet transformations are used. The image is grayscaled, and the wavelet is also used. Finally, the characteristics are merged from the image's vector.

- **Database unit:**

The graph vectors are extracted and saved as a feature data base in the database unit (FDB). Pictures and database images must be capable and thus have the same length for the comparison query. As a result, features are stated.

- **Retrieval unit:**

The database vector and query image are both contained in this unit. It employs an SVM ranking. The query image is then classified into various classes. It finds the 20 best images based on the similarity between the query image and the photos in the image class.

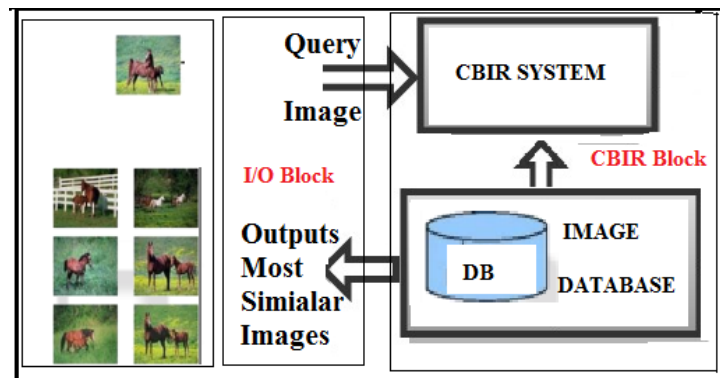


Figure 1: Block Diagram of CBIR

1.3 CBIR Applications

There are many applications for CBIR technology as below:

- To identify defects and defects in industrial automation.
- It is utilised on the Internet as a facial recognition & copyright.
- In the medical field, tumour identification, improving MRI and CT scans are extremely essential.
- For the weather forecast, satellite images.
- To map making from photographs.
- For crime detection using fingerprints matching.
- In the defense used for detection of targets.

1.4 Feature Extraction

To extract the characteristics, extraction methods are used. Color, texture, form, and vectors are some of the characteristics.

Color: Color is widely used for image display and is independent of image size. Color is a characteristic that is determined by the eye and how the input is processed. Color is used to distinguish between locations, things, and times. Colors are typically specified using colour spaces. Color quantization is the extraction of colour features using similarity dimension key components and colour space. The RGBs (Red, Green, Blue), HSVs (Hue, Saturation, and Value), or HSBs (Hue, Saturation, and Brightness) can be either (Hue, Saturation, and Brightness). Color histogram moments are examples of a block-based histogram. Color is frequently used to convey the size of a picture on its own.

Texture: Texture is an inherent feature of all surfaces that define visual patterns with their own uniformities. It includes important details about the outside lining's structural layout, such as clouds, leaves, bricks, and textiles, among other things. It depicts the relationship between the top and the surrounding environment.

Shape: A shape is a feature of something; a contour or form. Shape does not describe the shape of an area, but it does refer to the shape of an image. It allows an item to stand out from its surroundings. The form's representations can be divided into two categories: limited and regional. Scaling, rotation, and translation may not affect shape descriptors.

Feature Vectors: A vector is an n-dimensional numerical vector that represents an item in pattern recognition and machine learning. Many machine-learning methods require a numerical representation of objects in order to simplify statistical analysis and processing. When representing text, the feature values may be matched with the pixels of an image, possibly by term frequencies. A feature vector is an n-dimensional vector of numeric characteristics used in pattern recognition and machine learning. Because such representations enable the statistical examination and processing of many machine learning methods, numerical depiction of objects is required. When images represent words, the value may be equal to the pixels of a picture.

1.4.1 Basic architecture of Feature Extraction System

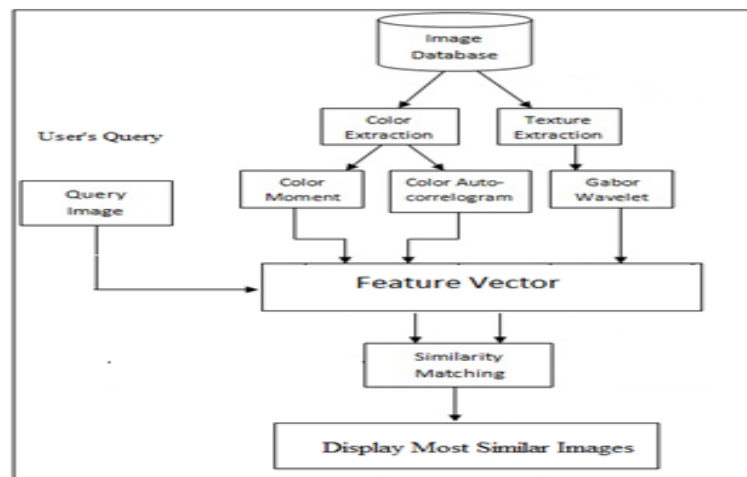


Figure 2: Feature extraction System

1.4.2 Feature Extraction Techniques

There are following techniques used as follows:

- **Color-Based Feature Extraction:** There are following Color Feature Extraction namely first one is Color moments and another one is Color Auto-Correlogram.

Color moments are measures that characterise the colour distribution in an image. Each image comparison generates a similarity score, and the lower the score, the more similar the two images are supposed to be. Color indexing (CI) is the most common use of colour moments. Images can be indexed, and the indexed colour moments will be included. For any colour model, colour moments can be computed. Per channel, three colour moments are computed (9 moments if the colour model is RGB and 12 moments if the colour model is CMYK).

Color Auto-Correlogram: A correlogram can be saved as a table of pairs of colours (i, j) , where the d -th entry indicates the likelihood that pixel j will be found at distance d from pixel i . An auto-correlogram, on the other hand, can be saved as an indexed colour table, with a d -th entry indicating the likelihood of finding a pixel i at distance d from the same pixel. As a result, the autocorrelogram only shows the spatial correlation of the same colours. [2].

- **Texture-Based Feature Extraction:** The text-based method includes several textual characteristics such as coarseness, contrast, directionality, linearity, regularity, and ruggedness of the image. There are some texture classification techniques available, such as Wavelet Transformation and Gabor Wavelet.

Wavelet: Wavelets can be used as a tool to extract data from a variety of sources, including audio and images. Analysis allows for the use of extremely long periods of time when specifics are required. It may reveal data characteristics that other signal research techniques miss, such as autonomy, breakdown points, interruptions in higher derivatives, and patterns. Wavelet noise analysis or can usually compress a signal without causing significant degradation. [3].

Gabor Wavelet: The Gabor filter, named after Dennis Gabor, is a filter. The Gabor filter is also known as the Gabor wavelet. It is widely used to extract texture characteristics from images in order to recover lost images. Gabor filters are actually a collection of wavelets, each firing energy in a different direction and frequency. The Gabor filter's frequency and orientation are very similar to the anatomy. Gabor filters use the Fourier transformation of the Fourier function for extraction and harmonic function. This collection of energy distributions can then be used to extract features. The Gabor filter's orientation and frequency adjustable properties make texture assessment simple [4].

1.5 Performance Evaluation

The performance of a retrieval system is evaluated in system using several criteria. As a result, some of the most commonly used performance measures are average accuracy recall and average retrieval rate. All of these performance measures parameters are calculated for each picture query using calculated accuracy and recall levels. Precision denotes the proportion of one's relevant results retrieved from absolute recovered graphics. On the other hand, recall denotes the percentage of total results from database graphics that are overall correct.

2 Literature Survey

We examined many research articles in this study that described different authors of different fields of computer vision and its applications' new developments of CBIR methods. In addition, some technical aspects of modern content-based image retrieval systems are discussed. Many technical elements are discussed further below:

2.1 Overview of Support Vector Machine (SVM)

In the mid-1990s, V. Vapnik proposed the support vector machine (SVM). In the last decade, it has been the most widely used algorithm for machine learning. SVM is currently being used extensively as a powerful machine learning technique for data mining in a CRM project. It is now widely used as the foundation for computer vision, pattern recognition, information retrieval, and data mining applications. A binary labelled dataset's linear separating hyperplane is used. It is a hyperplane-defined classifier.

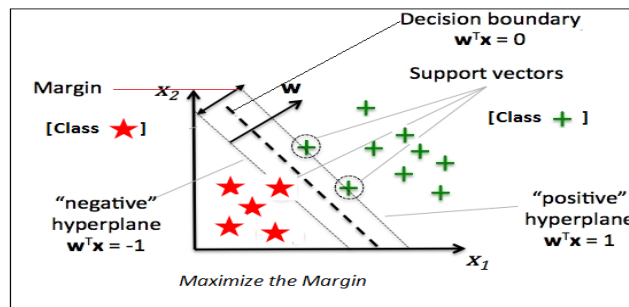


Figure 3: SVM maximize the Hyperplane and Margin two different class Red Star and Green plus (under 2D)

The SVM model depicts the Examples as space-based objects, with a clear and feasible gap separating the various categories. It is used for classification and non-linear classification. It converts the inputs into high-dimensional feature areas.

This is achieved by making a choice on categorization focusing on the significance of the linear combination of characteristics.

SVM is a binary choice Classification process that accepts data from two groups as the input label after it generates sparks to maximize fresh data. There are two major stages: analysis and training. SVM training entails providing the SVM with statistics and a previously established decision worth in order to create a final training set.

Two categorization problems may be discovered as input data is mapped. Using the RBF kernel and hyper-plane linear classification, this area contains vectors that are closest to the border of choice.

The class $-$ or class $+$ and the accompanying labels are $y_i=1$ for class A, and -1 for class B, let m -dimensional inputs = x_i ($i=1,2,3, \dots, M$).

Here Class -A is Red Star and Class -B is Green Plus.

SVM Decision Function

$$D(x) = w^T x + b; \quad (i)$$

Where w is an m -dimensional vector, b is a scalar, and y_i

$$D(x_i) \geq 1 \text{ for } i=1, 2, 3 \dots, M. \quad (ii)$$

The distance between the hyper separation plane $D(x) = 0$ and the training data closest to the hyperplane is known as the margin. The $D(x) = 0$ hyperplane is termed the optimum hyper plane with the greatest margin.

2.2 Surface Based Image Registration

The 3-D edge of an anatomic item or composition is inherent and provides valuable information on geometric features that may be pushed for image registration in the medical sector. These recording techniques combine

similar surfaces into different images and calculate the process that could be used to align these surfaces. Surface representation of surface points is just one example of what can be collected. An implicit surface or a parametric surface, such as a B-spline surface. Surfaces such as skin and bone can be easily removed, and CT and MR images are automatically removed. [5].

2.2.1 Applications of Surface-Based Image Registration

Remote sensing, medical imaging, computer vision and many more are extensively utilised as image registration. Surface-based image registration is split into major categories, including as:

- **Different Times (Multiview registration):** Pictures from the same perspective are obtained at various periods. The mean is to detect and quantify visual changes.
- **Different Viewpoints:** Images from the same scene from several angles are obtained. The objective is to obtain a big 2-D picture or a 3-D representation of the photographed scene. Applications such as remote sensing are available: for mosaicing pictures from the surveyed region and computer view for form recovery.
- **Multimodal modal image registration (Different Sensors):** The same scene image is captured from a separate sensor source. The primary goal is to combine information collected from many sources to create a more comprehensive portrayal of the scene.

2.1 Iterative Closest Point Algorithm (I.C.P)

This method is used to evaluate 3D medical data. There are several ICP variations that can be classified based on several criteria, including the selection of subsets from the supplied 3D data sets, the discovery of correspondence points, the weighting of estimated match pairs, the refusal of fake matches, the assignment of error metrics, and the minimization of error metrics. The ICP method is the list algorithm's broad purpose; the figure is independent of the shape and can be used with a variety of early geometric people such as pointing, line-part puts, and triangle puts (a great deal of the side is at the top).

$$\sqrt{\sum_{j=1}^N W_j^2} d^2(T(x_j), y) = \sqrt{\sum_{j=1}^N W_j^2} \|T(x_j) - y_j\|^2 \quad (\text{iii})$$

$$\text{Where } y_j = C(T(x_j), y) \quad (\text{iv})$$

2.4 Academic proposals

The literature survey is essential for understanding, extracting, and acquiring information about specific subject fields. Various current ideas have been evaluated in this article to form a picture:

M.E. ElAlami,[6] who stated that the 3-dimensional colour histogram And the Gabor filter method is a simple way to explain the characteristics of this image. A version with the colour consistency vector and wavelets was included to improve recovery performance. To extract the features set, there are two preparatory and reduction sequential methods. This method is then used to reduce the search space and time required for the search procedure.

Darshana Mistry et al.,[7] which is presented in this publication. At any given time, only two or more images of the same subject may be taken from different perspectives or detectors, such as references and feel graphics. They are classified according to their area of application and function. The registration process is divided into four stages.

Inch. Function detection

2. Matching feature

3. Model Alter estimate

4. Image Re-sampling and processing.

B. Bohra et al. [8] described in this paper how to reduce errors and time when registering data sets on the surface using point-cloud data arrangements. Normally used in the business to save CT scans, MRI images, and tumour photos, as well as to build collections of 3D models. I.C.P method for recording two 3D data collections and determining the closest points in the data collection based on the tolerance distance.

N. Ali et al. [9], which demonstrated the overall efficiency of image recovery to Improve Standard graphics are created using the rule of thirds, which divides a picture into two halves by placing regions or attention items in the grid's cortical lines. Thus, the texture and colour characteristics provide a suitable and efficient outcome to the human visual system.

Pratistha Mathuretl., Neha Janu, [10] In comparison to various transformations on waves and discrete transformations of cosine, presenting Gabor feature extractions has more precision. Gabor is studied for the extraction of edge or shape compared to the extraction of DCT and DWT as a result of its analytic characteristics. Returning to DWT and DCT, the feature is extracted in the low frequency subband (LL) along with other frequency groups, resulting in a loss of many characteristics and lower accuracy than Gabor. Gabor with scale projection achieved greater accuracy than Gabor without scale projection.

Kanokwan Khiewwan, G.S.Somnugpong, [1-1] which presented the combination of two new methods such as graphic feel and spatial significance of pairs of colour inside feature adds the high image that is shifting with greater robustness To provide averaging precision, the techniques must also be improved. By utilising a colour correlogram, information can be treated, whereas EDH provides geometry information when it comes to the image but in a different coloration. It is a combination of low-cost features that produces a better result than a single feature. The Euclidean distance is useful for dimensioning purposes.

AtifNazir and colleagues[1 2]. This newspaper's feature introduces form descriptors, shape displays, and texture characteristics. The CBIR methodology combines local and global information characteristics. They proposed a novel CBIR method for combining colour and texture properties. Color Histograms (CH) can be used to extract colour information. The features Edge Histogram Descriptor and Discrete Wavelet Transform are extracted. Characteristics are created. They combine superior outputs with a few characteristics rather than just one. As a result, the texture and colour characteristics provide an efficient and appropriate finish to the human visual system.

Neha Janu and Pratistha Mathuretl. [1 3] pioneered the frequency domain characteristic extraction procedure. The Gabor filter, Discrete Cosine Transform, and Discrete Wavelet Transform are the three feature extraction procedures that are used.

J.Cook et al. [14] presented a novel 3D face identification method in this paper using the Iterative Closest Point (ICP) method. When there is contrast, it is useful for recording stiff facial and facial parts of the registration error across an area. This method makes use of 3D registration methods. The ICP method is used to compensate for the properties of such surfaces and to match the target with assessment.

YinghuiGao Wang, RuigangFu, BiaoLi, and RuigangFu [15]. By means of similarity measures and, in particular, representations. This article discusses CBIR with CNN and uses the linear vector support machine (SVM) to train a hyper-plane that can separate comparable image pairings and different graphic combinations at a large scale. According to the results of the tests, the method may improve the overall efficiency of CBIR. This SVM and CNN-based research paper is used to learn similarity measures, and CNN can be used to extract feature representations.

3. Conclusions and Research Motivation

This article gives a brief overview of image recovery and its structure. To obtain visuals from the training database, many academics use extraction methods. Each of these databases contains textures, forms, and a variety of other colorful images. Color histograms (RGB and HSV) are used to extract color images, and Gabor filters have all been shown to be the most effective when using textured extraction. Finally, in 3 Data Sets image registration, the classification vector support machine (SVM) and surface-based image registration methods are widely used in many businesses. For the entire surface-centered image registration process, the point-based image registration method is most efficient and popular, with the renowned I.C.P (Iterative Closest Point) algorithm. According to the findings of several studies, the service vector system may make this job easier and more efficient. The ICP algorithm produces excellent results and selects images that are relevant to a person's needs. ICP and SVM may be used in future approaches to record images in datasets.

References

- [1] Darshana Mistry, Asim Banerjee, "Review: Image Registration", *International Journal of Graphics & Image Processing* [Vol 2|issue 1|February 2012].
- [2] J. Yu, Z. Qin, T Wan, and X. Zhang, Feature integration analysis of bag-of-features model for image retrieval, *Neurocomputing*, vol. 120, pp. 355-364, 2013.
- [3] H. A. Hassan ; N. M. Tahir ; I. Yassin ; C.H.C Yahaya ; S. M. Shafie *International Conference on Computer Vision and Image Analysis Applications Year: 2015 | Conference Paper | Publisher: IEEE.*
- [4] J. Yue, Z. Li, L. Liu, and Z. Fu, "Content-based image retrieval using color and texture fused features, *Math. Comput. Model.* vol. 54, no. 34, pp. 1121-1127, 2011.
- [5] Brian Amberg, Sami Romdhani and Thomas Vetter "Optimal Step Nonrigid ICP Algorithms for Surface Registration". This work was supported in part by Microsoft Research through the European PhD Scholarship Programme.
- [6] Janu Neha, PratisthaMathur, "Performance analysis of frequency domain based feature extraction techniques for facial expression recognition." In *7th International Conference on Cloud Computing, Data Science & Engineering-Confluence*, 2017, pp. 591-594. IEEE, 2017.
- [7] M. E. Elalami, A novel image retrieval model based on the most relevant features, *Knowledge-Based Syst.*, vol. 24, no. 1, pp. 23-32, 2011.
- [8] Brahmdukt Bohra1# Deepak Gupta2* Shikha Gupta3## "An Efficient Approach of Image Registration Using Point Cloud Datasets".
- [9] Nouman Ali, Khalid Bashir Bajwa , Robert Sablatnig , ZahidMehmood " Image retrieval by addition of spatial information based on histograms of triangular regions".
- [10] Jost T., Hügli H. (2002) Fast ICP Algorithms for Shape Registration. In: Van Gool L. (eds) *Pattern Recognition. DAGM 2002. Lecture Notes in Computer Science*, vol 2449. Springer, Berlin, Heidelberg.
- [11] Janu Neha, PratisthaMathur, "Performance analysis of frequency domain based feature extraction techniques for facial expression recognition." In *7th International Conference on Cloud Computing, Data Science & Engineering-Confluence*, 2017, pp. 591-594. IEEE, 2017.
- [12] AtifNazir, Rehan Ashraf, TalhaHamdani, NoumanAli,"Content based image retrieval system by using HSV color histogram, discrete wavelet transform and edge histogram descriptor",*International Conference on Computing, Mathematics and Engineering Technologies (iCoMET),Azad Kashmir,(2018)1-6.*
- [13] S. Somnugpong and K. Khiewwan, "Content Based Image Retrieval using a combination of Color Correlograms and Edge Direction Histogram", *13th International Joint Conference on Computer Science and Software Engineering*,DOI:10.1109, IEEE, (2016).
- [14] Neha Janu, PratisthaMathur, "Performance Analysis of Feature Extraction Techniques for Facial Expression Recognition", *International journal on Computer Applications*, ISSN No. 0975 –8887, Volume-166, Issue-1, May 2017.
- [15] Jamie Cook,Vinod Chandran, SridhaSridharan and Clinton Fookes, "Face Recognition From 3d Data Using Iterative Closest Point Algorithm And Gaussian Mixture Models",*Greece*, 2004, pp. 502-509.
- [16] Ruigang Fu, Biao Li, Yinghui Gao, Ping Wang, "Content-Based Image Retrieval Based on CNN and SVM", *2016 2nd IEEE International Conference on Computer and Communications*, pages (638-642).
- [17] Peter J. Kostelec and SenthilPeriaswamy, "Image Registration for MRI". *Modern Signal Processing MSRI Publications Volume 46*, 2003.
- [18] C. S.Won, D. K. Park and Y. S. Jeon, "an efficient use of MPEG-7 Color Layout and Edge Histogram Descriptors", *proceeding of the ACM workshop on multimedia*, (2000), pp. 51-54.
- [19] T. Kato, "Database architecture for content-based image retrieval", in *Image Storage and Retrieval Systems*, *Proc SPIE 1662*, (1992) pp112-123.
- [20] Jan Elseberg, DoritBorrmann and Andreas Nüchter, "One billion points in the cloud – an octree for efficient processing of 3D laser scans". In *Proc. ISPRS Journal of Photogrammetry and Remote Sensing 76 (2013) 76–88.*
- [21] Kumar and M. Sinha (2014), "Overview on vehicular ad hoc network and its security issues," *International Conference on Computing for Sustainable Global Development (INDIACom)*, pp. 792-797. doi: 10.1109/IndiaCom.2014.6828071.

- [22] Mr. Ankit Kumar, Dr. Dinesh Goyal, Mr. Pankaj Dadheech, (2018), “A Novel Framework for Performance Optimization of Routing Protocol in VANET Network”, *Journal of Advanced Research in Dynamical & Control Systems*, Vol. 10, 02-Special Issue, 2018, pp-2110-2121, ISSN: 1943-023X.
- [23] Mr. Pankaj Dadheech, Dr. Dinesh Goyal, Dr. Sumit Srivastava, Mr. Ankit Kumar, (2018), “A Scalable Data Processing Using Hadoop & MapReduce for Big Data”, *Journal of Advanced Research in Dynamical & Control Systems*, Vol. 10, 02-Special Issue, 2018, pp-2099-2109, ISSN: 1943-023X.