Query-Based Image Retrieval Using SVM

Neha Janu[1(✉)], Sunita Gupta[2], Meenakshi Nawal[1], and Pooja Choudhary[3,4]

[1] Department of Computer Science and Engineering, Swami Keshvanand Institute of
Technology Management and Gramothan, Jaipur, India
{nehajanu,meenakshi.nawal}@skit.ac.in

[2] Department of Information Technology, Swami Keshvanand Institute of Technology
Management and Gramothan, Jaipur, India
sunita@skit.ac.in

[3] Department of Electronics and Communication Engineering, MNIT, Jaipur, India
pooja.choudhary@skit.ac.in

[4] Department of Electronics and Communication Engineering, Swami Keshvanand Institute of
Technology Management and Gramothan, Jaipur, India

Abstract. In today's world, images are used to extract information about objects in a variety of industries. To retrieve images, many traditional methods have been used. It determines the user's query interactively by asking the user whether the image is relevant or not. Graphics have become an important part of information processing in this day and age. The image is used in image registration processing to extract information about an item in a variety of fields such as tourism, medical and geological, and weather systems calling. There are numerous approaches that people use to recover images. It determines an individual's query interactively by asking users whether the image is relevant (similar) or not. In a content-based image retrieval (CBIR) system, effective management of this image database is used to improve the procedure's performance. The study of the content-based image retrieval (CBIR) technique 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 aforementioned industries. Several research papers examining color feature and texture feature extraction concluded 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's also known as Query by Image Content (QBIC), and it introduces technologies that allow you to organize digital pictures based on their visual characteristics. They are built around 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 fundamental task in image processing that is used for pattern matching of two or more images taken at different time intervals from different sensors and from different perspectives. Almost all large industries that value images require image registration as a steadfastly associated operation. Identical for objective reorganization real time images are targeted in literal guesses of anomalies where image registration is a major component.

On 3-D datasets, two types of image registration are performed: manually and automatically. Human operators are responsible for all of the process corresponding features of the images that are to be registered when manually registering. To obtain logically good registration results, a user must select a significantly large number of feature pairs across the entire image set. Manual processes are not only monotonous and exhausting, but they are also prone to inconsistency and limited accuracy. As a result, there is a high demand for developing automatic registration techniques that require less time or no human operator control [1].

In general, there are four steps for image registration for supporters, such as:

Detection of Features: It is a necessary step in the image registration process. It detects features such as closed-boundary fields, ranges, edges, outlines, crossing line points, angles, and so on.

Feature Matching: In this step, the process of matching the features extracted by Database Images and Query Images that result in a visually similar result is performed.

Estimation of the Transform Model: The sort and parameters of the so-called mapping purposes, uses, aligning the sensed image with the statement, and direction image are all assigned a value. The parameters of the mapping purposes and uses are worked out using the made specific point letters.

Image Re-sampling and Transformation: The sensed image is significantly altered as a result of the mapping purposes and uses. Image values in non-integer orders are calculated using the correct interpolation expert method of art, and so on.

1.2 Fundamentals of CBIR System

In typical CBIR systems, the visual content of the pictures is extracted from the database (Fig. 1). And it described by multi-dimensional feature vectors. These features vector of the images in the database forms a feature database. To retrieve the images, users offer the retrieval system with example images (Fig. 1).

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 then extracts three features from the picture and saves them as a feature vector. There are three distinct characteristics: texture, color, and form. The color moment might be included with a feature. Images may be translated and partitioned

into grids by extracting their color features. A few seconds have been shaved off the end result. Wavelet transforms with pyramidal and tree-structured structures are employed to extract texture information. It is converted to grayscale and the Daubechies wavelet is used. The image's feature vector is used to aggregate all of the image's features.

Database Unit: The graphics' feature vectors are gathered and stored in the database unit in the same way that a feature database is (FDB). In order to do a comparison query, both pictures and database images must have the same capabilities. As a result, features are shown.

Retrieval Unit: You may use this unit to store your database feature vector and query picture. An SVM classifier is tinkered with in this game. It is then sorted into one of these categories. To get the 20 most similar photos to the one being searched for, it compares the image being searched for to all of the others in the class.
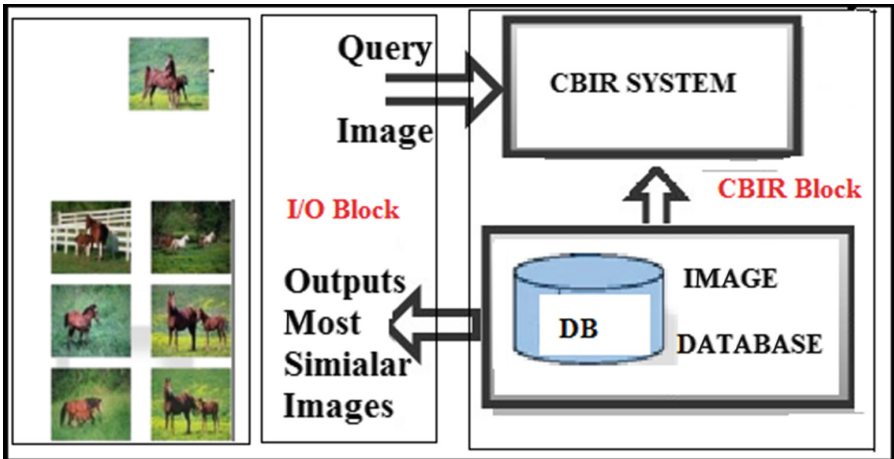

Fig. 1. Block diagram of CBIR

1.3 CBIR Applications

There are many applications for CBIR technology as below:

1. To identification of defect and fault in industrial automation.
2. It is used as face recognition & copyright on the Internet.
3. In medical plays a very important role in Tumors detection, Improve MRI and CT scan.
4. For the weather forecast, satellite images.
5. To map making from photographs.
6. For crime detection using fingerprints matching.
7. In the defense used for detection of targets.

1.4 Feature Extraction

When extracting the features, feature extraction techniques are applied. It has properties such as the colour, the texture, the form, and the ability to show vectors (Fig. 2). It doesn't matter what the size of a picture is, colour is employed extensively to convey an image's meaning. Color is a feature of the eye and the mind's ability to absorb information. Color helps us communicate the distinctions between different locations, things, and epochs. Colors are often specified in a colour space. Similarity dimension key components and colour space, colour quantization are used in colour feature extraction. RGB (red, green, and blue), HSV (hue, saturation, and value), or HSB (hue, saturation, and brightness) can be used (Hue, Saturation, and Brightness). These colour histogram moments are referred to as cases. The independent and picture representation of the magnitude of an image is commonly utilised for colour [2].

A surface's texture specifies the visual patterns it creates, and each pattern has a certain degree of uniformity. It contains crucial information on the structure of this exterior layer, such as; clouds, leaves, bricks, and cloth, for example. It explains how the top is connected to the surrounding environment. A thing's form might be characterised as its "outline or shape," or as its "characteristic configuration." In this context, "shape" does not refer to a geographic place, but rather to the shape of the picture being searched. It allows an object to stand out from the rest of the landscape. Both Boundary-based and Region-based shape representations exist. It may be necessary for shape descriptors to be scale, rotation, and translation invariant. For example, a feature vector is an n-dimensional array of numerical characteristics that describe an object in pattern recognition and machine learning. Since numerical representations of things make processing and statistical analysis easier, many machine learning methods rely on them. For pictures, feature values may correlate to the number of pixels in an image, while for texts, feature values might correspond to the number of times a phrase appears in a sentence. An n-dimensional vector of numerical characteristics that represent a few things is known as a feature vector in pattern recognition and machine learning applications. Many machine learning algorithms rely on numerical representations because they make it easier to conduct statistical analysis and processing. Textual representations may use pixels to indicate the value of a feature, while pictures may use feature values corresponding to the pixels.

1.4.1 Basic Architecture of Feature Extraction System

1.4.2 Feature Extraction Techniques

There are following techniques used as follows:

In terms of Color-Based Feature Extraction, there are two main options: Color Moments and Color Auto-Correlation. An image's colour distribution may be measured using a concept known as "colour moments." Image similarity scores are calculated for each comparison, with lower scores indicating a greater degree of resemblance between pictures. It is the primary use of colour moments that is colour indexing (CI). It is possible to index images such that the computed colour moments may be retrieved from the index. Any colour model may be used to compute colour moments. Per channel, three different colour moments are calculated (9 moments if the colour model is RGB and 12
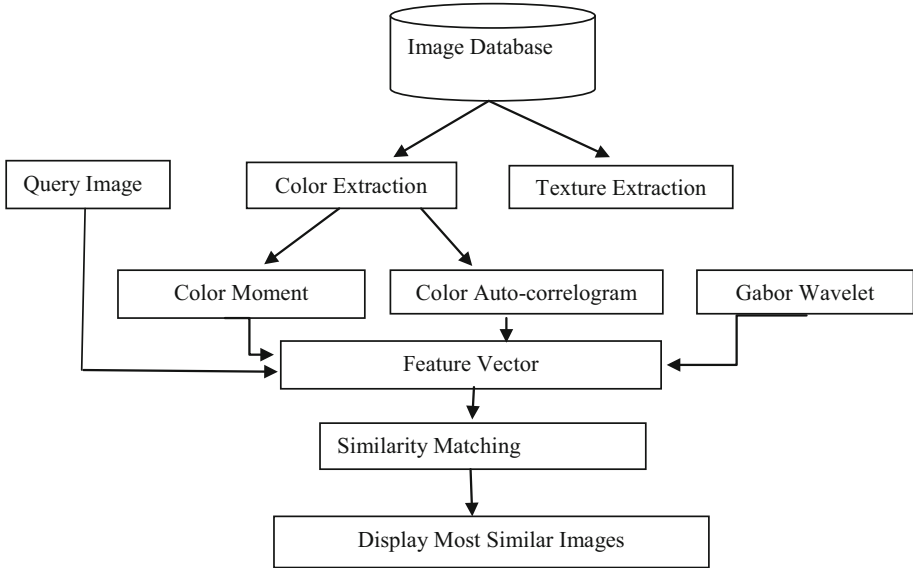
Fig. 2. Feature extraction system

moments if the colour model is CMYK). Using colour pairings (i,j) to index the table, a colour auto-correlogram may be stored as a correlogram, with each row indicating how often it is to be found between two adjacent pixels (i,j).

As opposed to storing an auto-correlogram in the form of a table where each dth item represents the likelihood of finding a pixel I from the same one at the given distance. Auto-correlograms therefore only display the spatial correlation between hues that are the same [2]. In this Texture-based approach, there are a number of Texture attributes, such as Coarseness, Directionality, Line likeness, Regularity and Roughness, that may be extracted from the image. Techniques like Wavelet Transform and Gabor Wavelet can be used to categories textures. You may extract information from any kind of data using a wavelet: this includes audio signals and pictures. Analysis allows us to use very large time intervals in which we are looking for specific information. Other techniques of signal analysis omit elements like self-similarity, breakdown points, discontinuities in higher derivatives, and patterns that this approach is capable of exposing about the data. When applied to a signal, wavelet analysis de-noises or compresses it without causing significant deterioration [3]. A filter named after Dennis Gabor; the Gabor filter is a wavelet. Also known as a Gabor wavelet filter, the Gabor filter. The extraction of texture information from pictures for image retrieval is generally accepted. Gabor filters may be thought of as a collection of wavelets, each of which directs energy in a certain direction and at a specific frequency. Like the human body, the Gabor filter has frequency and orientation representations. Extraction and harmonic functions are achieved via Gabor filters using Fourier transforms of the function's Fourier transforms. This set of energy

distributions may then be used to extract features. The Gabor filter's orientation and frequency adjustable characteristic makes texture assessment simple [4].

Performance Evaluation

The performance of a retrieval system is evaluated using a variety of criteria in system. Average accuracy recall and retrieval rate are two of the most often utilized performance metrics. Each query image's accuracy and recall values are used to construct these performance metrics. Percentage of findings that can be reliably deduced from recovered visuals is what we mean by precision. Recall, on the other hand, indicates the proportion of all database graphics results that are overall.

2 Literature Survey

In this review of the literature, we looked at a wide range of publications from diverse authors in the field of computer vision and its applications to discover the latest CBIR trends. In addition, the present state of content-based picture retrieval systems is also covered in this paper. In addition to the above, there are a number of existing technological aspects.

2.1 Support Vector Machine (SVM)

It was proposed by V. Vapnik in the mid-1990s that support vector machine (SVM) be implemented. Since the beginning of the millennium, it has been the most common machine learning algorithm. As a sophisticated machine learning technique, SVM is currently commonly employed in data mining projects like CRM. Computer vision, pattern recognition, information retrieval, and data mining all rely on it these days. A binary labelled dataset's linear separation hyperplane may be found using this technique. It's a classifier with a hyperplane for dividing classes (Fig. 3).

The SVM model depicts the Examples as objects in space, with a visible gap as broad as possible between the various kinds. Non-linear classification and classification are routine for it. High-dimensional feature spaces can be mapped from their inputs. The classifier accomplishes this by focusing on the linear combination of the characteristics while reaching a classification decision. SVM is a type of binary decision-making option. Classification process that uses labelled data from two groups as input and generates sparks in a file to maximize the number of new data points possible. Analyzing and training are the two main phases. In order to train an SVM, one must provide the SVM with a finite training set of previously known statistics and decision values. A two-class classification problem can be found as input data is mapped. Using the RBF kernel and hyper-plane linear interpolation Vectors closest to the decision border are placed in this space by the classifier.

Let m-dimensional inputs $=$ xi (i $=$ 1,2,3… ………M) belong to Class – "*" or Class– " +" and the associated labels be yi $=$ 1 for Class A and $-$ 1 for Class B.

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

Decision function for SVM is

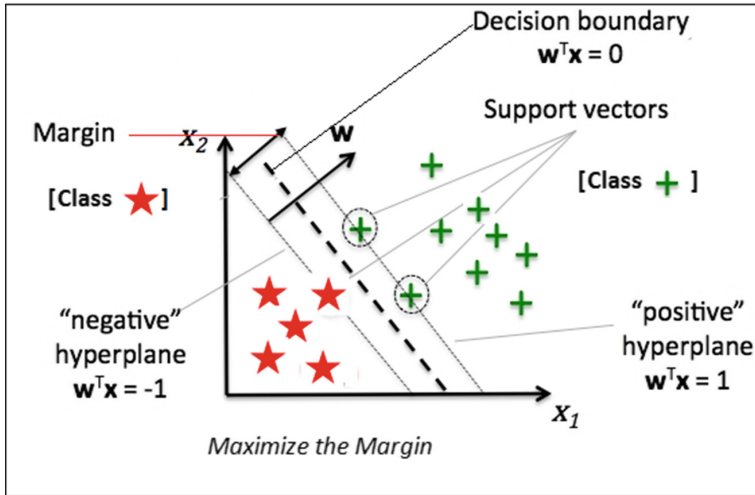$$D(x) = w^T x + b; \qquad\qquad (i)$$

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

where w is an m-dimensional vector, b is a scalar, and yi

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

The distance between the separating hyper plane $D(x) = 0$ and the training datum nearest to the hyper plane is called the margin. The hyper plane $D(x) = 0$ with the maximum margin is called the optimal separating hyper plane.

2.2 Surface Based Image Registration

For image registration in the medical business, the three-dimensional edge surface of an anatomical item or composition is intrinsic and provides substantial geometrical feature data. By combining various registration approaches, it is possible to create equivalent surfaces in many pictures and then compute the transformation between the images. On the surface, there is just one object that can be collected. An implicit or parametric surface, such as a B-spline surface, is revealed by a facial view. Skin and bone may be easily extracted from CT and MR images of the head [5].

2.2.1 Applications of Surface-Based Image Registration

Many fields, including remote sensing, medical imaging, computer vision, and others, make use of image registration in some capacity or another. Surface-based image registration may be broken down into a number of subcategories, some of which include:

1. Image acquisition is done at various times for the same view (Multiview registration). Changes in perspective can be found and estimated by using the mean.

2. Different Perspectives: Images of the same scene are taken from several angles. To get, a huge two-dimensional or three-dimensional perspective of the scene being photographed is used. It is possible to use remote sensing and computer vision to create mosaics of photos from a scanned region.
3. A picture of the same scene is obtained through the use of many sensors, which is known as multimodal modal image registration. More comprehensive scene representations can be achieved by integrating information from several sources.

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

Medical data in 3D may be evaluated using this approach. To categories the many variations of the ICP, several criteria may be used. Examples include: selecting portions of 3D data sets, locating correspondence points, weighting estimated correspondence pairings, rejecting false matches and assigning an error score. If true, then another is necessarily true and parametric curves and comes to the top of the ICP algorithm which is a general-purpose, representation-independent shape-based number on a list algorithm that can be used with a range of geometrical early people including point puts, line part puts, triangle puts (much-sided comes to the top).

$$\sqrt{\sum\nolimits_{j=1}^{N} W_j^2 d^2(T(xj),\ y} = \sqrt{\sum\nolimits_{j=1}^{N} W_j^2 ||T(xj) - yj||^2} \qquad \text{(iii)}$$

$$\text{Where}\ \ y_j = C(T(xj),\ y) \qquad \text{(iv)}$$

2.4 Academic Proposals

Understanding, extracting, and obtaining knowledge about a certain field of a topic is critical to doing a literature study. In this work, a number of existing approaches to picture retrieval are evaluated, including:

ElAlami, [6] who stated that the 3D colour histogram was a better option than the 2D colour histogram. Furthermore, the Gabor filter technique is capable of describing the image's attributes in a matter of seconds. Color coherence vectors and wavelets were used to improve recovery speed in a new version. Preliminary and reduction processes are used to extract the features from the collection. After then, the method is used to decrease the search area and the time spent retrieving information.

In this paper, [7] Darshana Mistry, et al. Detectors such as references and feel graphics can be used to compare two or more images of the same scene taken at different times, from different perspectives, or at different times of day. According on location and feature, they are grouped into categories. The picture registration process has four stages.

1. Feature detection
2. Feature matching
3. Alter model estimation
4. Image Re-sampling and transformation.

Data sets utilising point-cloud data organisation were used by Bohra, et al. [8] to minimise errors and time in surface-based image registration methods. To store CT, MRI, and Tumor pictures, as well as to construct 3D models of sets of these images. ICP method that registers two 3D data collections and finds the closest points into data collections based on the given tolerance distance for each data collection.

Enhances the overall efficacy of picture retrieval, according to N. Ali, et al. [9]. Images are divided into two halves by using the rule of thirds to identify focal points in the cortical lines that make up this grid. Color and texture are two features that provide the human visual system a suitable response.

Gabor feature extractions are more accurate than other wavelet transforms and discrete cosine transforms, according to Pratistha Mathuretl, Neha Janu, [10]. That Gabor's analytical features reveal that Gabor is a better method for extracting edge or shape features than DCT and DWT feature extraction, as shown in the findings. A percentage of the features are lost in the low-frequency feature sub-band (LL) and other frequency groups during the feature extraction process in DWT and DCT. This results in lesser accuracy as compared to Gabor. Gabor with a scale projection was more accurate than Gabor without a scale.

As a result of the incorporation of two new ways, such as graphic feel and spatial importance of pairs of colour within feature, G.S. Somnugpong, Kanokwan Khiewwan provided the high image that is shifting more robustly. In order to improve the precision of averaging, new strategies must be developed. It's possible to use colour correlograms to process information, whereas EDH gives geometric information when it comes to the image. Better than any one feature is a combination of low-end characteristics. Dimensioning can benefit from using Euclidean distance [11].

Shape descriptors, shape representations, and texture aspects were all introduced in this article. Using a CBIR method, information may be accessed both locally and globally. In order to combine colour and texture data, they proposed a novel CBIR approach. Extracting colour information may be done using the Color Histogram (CH). Discrete Wavelet Transform and Edge Histogram Descriptor are used to extract features. A lot of work has been done on the features. They've integrated a few elements to get greater results than a single feature could have on its own. As a result, the human visual system is well-served by texture and colour. Characteristic Extraction technique was introduced by Neha Janu, Pratistha Mathuretl., [12]. They include Gabor filter, DCT and DWT. J. Cook, et al. [13] developed a novel method of 3D face identification using the Iterative Closest Point (ICP) algorithm in this newspaper. Face and face modelling is effective for recording stiff parts when contrast is present in a certain area. This method employs 3D registration methods. It is used to correct for the surface's characteristics and build a correlation between target and assessment using the ICP algorithm.

BiaoLi [14] and YinghuiGao Wang [14] contributed the data. Measures and representations of attribute similarity. When used in conjunction with a convolution neural network (CNN), the SVM is used in this study in order to develop an image-discriminating hyperplane that can distinguish between visually similar and different picture pairings at a large level. The approach has been shown to improve the overall efficacy of CBIR in tests. SVM is used to learn the similarity measures, while CNN is used to extract the feature representations in this research work.

3   Conclusions and Research Motivation

In this study, a quick introduction to image retrieval and its structure are given. Researchers are using feature extraction algorithms to extract visuals from their training database. Textures, forms, and a variety of unique visuals are included in each of these databases. In order to retrieve pictures based on colour, Gabor filters and colour histograms (RGB and HSV) have been determined to be the most effective techniques for performing texture-based extraction. SVM classifiers and surface-based image registration approaches are frequently used in 3-D data sets image enrollment by a wide range of organisations. Finally. I.C.P (Iterative Closest Point) is the most common and efficient algorithm for surface-centered picture registration, and it is beneficial for improved image registration. A service vector system may be able to make this work easier and more efficient, according to a number of studies. The ICP algorithm provides excellent results and selects the images that best suit the needs of the user. ICP and SVM may be used in the future to classify images in datasets.

References

1. Mistry, D., Banerjee, A.: Review: image registration. Int. J. Graph. Image Process. **2**(1) 2012
2. Yu, J., Qin, Z., Wan, T., Zhang, X.: Feature integration analysis of bag-of-features model for image retrieval. Neurocomputing **120**, 355–364 (2013)
3. Hassan, H.A., Tahir, N.M., Yassin, I., Yahaya, C.H.C., Shafie, S.M.: International Conference on Computer Vision and Image Analysis Applications. IEEE (2015)
4. Yue, J., Li, Z., Liu, L., Fu, Z.: Content-based image retrieval using color and texture fused features. Math. Comput. Model. **54**(34), 1121–1127 (2011)
5. Amberg, B., Romdhani, S., Vetter, T.: Optimal Step Nonrigid ICP Algorithms for Surface Registration. This work was supported in part by Microsoft Research through the European PhD Scholarship Programme
6. Neha, J., Mathur, P.: Performance analysis of frequency domain based feature extraction techniques for facial expression recognition. In: 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence, pp. 591–594. IEEE (2017)
7. Elalami, M.E.: A novel image retrieval model based on the most relevant features. Knowl.-Based Syst. **24**(1), 23–32 (201l)
8. Bohra, B., Gupta, D., Gupta, S.: An Efficient Approach of Image Registration Using Point Cloud Datasets
9. Ali, N., Bajwa, K.B., Sablatnig, R., Mehmood, Z.: Image retrieval by addition of spatial information based on histograms of triangular regions
10. Jost, T., Hügli, H.: Fast ICP algorithms for shape registration. In: Van Gool, L. (ed.) DAGM 2002. LNCS, vol. 2449, pp. 91–99. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-45783-6_12
11. Nazir, A., Ashraf, R., Hamdani, T., Ali, N.: Content based image retrieval system by using HSV color histogram, discrete wavelet transform and edge histogram descriptor. In: International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Azad Kashmir, pp. 1–6 (2018)
12. Somnugpong, S., Khiewwan, K.: Content based image retrieval using a combination of color correlograms and edge direction histogram. In: 13th International Joint Conference on Computer Science and Software Engineering. IEEE (2016). 10.1109

13. Janu, N., Mathur, P.: Performance analysis of feature extraction techniques for facial expression recognition. Int. J. Comput. Appl. **166**(1) (2017). ISSN No. 0975-8887
14. Cook, J., Chandran, V., Sridharan, S., Fookes, C.: Face recognition from 3D data using iterative closest point algorithm and Gaussian mixture models, Greece, pp. 502–509 (2004)
15. Fu, R., Li, B., Gao, Y., Wang, P.: Content-based image retrieval based on CNN and SVM. In: 2016 2nd IEEE International Conference on Computer and Communications, pp. 638–642
16. Kostelec, P.J., Periaswamy, S.: Image registration for MRI. In: Modern Signal Processing, vol. 46. MSRI Publications (2003)
17. Won, C.S., Park, D.K., Jeon, Y.S.: An efficient use of MPEG-7 color layout and edge histogram descriptors. In: Proceedings of the ACM Workshop on Multimedia, pp. 51–54 (2000)
18. Kato, T.: Database architecture for content-based image retrieval. In: Image Storage and Retrieval Systems, Proceedings SPIE 1662, pp. 112–123 (1992)
19. Elseberg, J., Borrmann, D., Nüchter, A.: One billion points in the cloud – an octree for efficient processing of 3D laser scans. Proc. ISPRS J. Photogram. Remote Sens. **76**, 76–88 (2013)
20. Kumar, A., Sinha, M.: Overview on vehicular ad hoc network and its security issues. In: International Conference on Computing for Sustainable Global Development (INDIACom), pp. 792–797 (2014). https://doi.org/10.1109/IndiaCom.2014.6828071
21. Nankani, H., Mahrishi, M., Morwal, S., Hiran, K.K.: A Formal study of shot boundary detection approaches—Comparative analysis. In: Sharma, T.K., Ahn, C.W., Verma, O.P., Panigrahi, B.K. (eds.) Soft Computing: Theories and Applications. AISC, vol. 1380, pp. 311–320. Springer, Singapore (2022). https://doi.org/10.1007/978-981-16-1740-9_26
22. Dadheech, P., Goyal, D., Srivastava, S., Kumar, A.: A scalable data processing using Hadoop & MapReduce for big data. J. Adv. Res. Dyn. Control Syst. **10**(02-Special Issue), 2099–2109 (2018). ISSN 1943-023X