Allana Management Journal of Research (AMJR) Vol.15 (01). pp 027 -037, Jan 2025
Available Online at https://theaimsjournal.org
DOI: https://doi.org/10.62223/AMJR.2025.150103
ISSN NO: 2581 – 3137

Study and Development of Content Based Image Retrieval (CBIR) Using Multiple Features on Android Platform

Dr. Manjusha Yuvraj Patil[1] & Dr. Vinaya Keskar[2]
ATSS College of Business Studies and Computer Application, Pune, India

ARTICLE HISTORY

Received on: 10/04/2025
Accepted on: 25/04/2025
Available Online:28/06/2025

Key words:
Content Based Image Retrieval (CBIR), Color, Texture, Shape, Feature extraction, Transformed Image, Android, Operating System (OS)

**Corresponding Author*
Dr. Manjusha Yuvraj Patil, ATSS College of Business Studies and Computer Application, Pune, India
Email:manjushapatil@atsscollege.edu.in

ABSTRACT

The need for more sophisticated and effective picture retrieval methods is rising because of the quick rise in image creation and sharing, particularly from mobile devices with high-resolution cameras. Accurate picture searching no longer requires traditional techniques that rely on metadata or simple pixel comparisons.
Purpose: The goal of this work is to create a successful Content-Based Image Retrieval (CBIR) system that improves the speed and accuracy of obtaining pertinent photos from sizable databases.
Design/Methodology/Approach: To transform picture content into comprehensive feature vectors, the suggested CBIR system combines several visual aspects, including color, texture, and shape, with updated image descriptors. Multi-feature analysis is then performed to extract similar images from these vectors.
Findings: Compared to traditional techniques, the suggested system considerably increases retrieval accuracy and processing time, according to an evaluation on a dataset of 1,000 photos split into 11 categories.
Research Limitations/Implications: The study only included 1,000 photos, and additional testing on bigger and more varied datasets could be necessary to confirm scalability and generalization.
Practical Implications: The system improves user experience in picture retrieval applications like digital libraries, social media, and e-commerce by providing a scalable and reliable method for platforms and organizations managing large image databases.
Originality/Value: By combining multi-feature picture analysis with optimized descriptors, this study offers a fresh approach to long-standing CBIR problems and opens the door for scalable, real-time image search in contemporary settings.

INTRODUCTION

Instead, then depending on human supplied keywords or metadata, Content-Based Image Retrieval (CBIR) systems analyse each image's visual content, including color, texture, shape, and other descriptors, to enable automatic searches of vast image collections. CBIR, which was initially used by T. Kato in 1992 to characterize experiments in image retrieval based on color and form data, has since grown to be a fundamental application of computer vision, propelled by the rapid expansion of digital photography.

©2025 Dr. Manjusha Yuvraj Patil[1] & Dr. Vinaya Keskar[2].This is an open access article distributed under the terms of the Creative Commons Attribution - Non-Commercial 4.0 International License.

With the ability to take and share thousands of high-resolution photos, today's smartphones, tablets, and digital cameras make effective querying and retrieval essential.

Traditional CBIR techniques frequently encounter difficulties when dealing with huge datasets and the resource limitations of mobile devices, even with advancements in feature extraction and similarity measurement. When handling hundreds of photos on portable technology, simple pixel-level comparisons or single-feature approaches may result in sluggish response times and less than ideal accuracy. Mobile CBIR solutions are further complicated by the need to provide consistent memory management and process isolation on platforms like Android, where each application operates in a separate virtual machine.

By combining several complementary descriptors (color coherence vectors, gray-level co-occurrence matrices, shape analyses, and transformed-image features) into a single feature-vector representation, our work tackles these issues. When combined with lightweight EXIF metadata processing and effective distance-based matching, our solution provides quick, precise retrieval on Android devices without sacrificing stability. The suggested method shows notable gains in search speed and accuracy when tested on a varied dataset of 1,000 photos in 11 categories, providing a scalable solution for contemporary mobile image-rich contexts.

FEATURE EXTRACTION AND TECHNIQUE

Features Feature Extraction Techniques may consist of both text-based features and visual features. In the visual features can be categorized as low level and high-level features. The selection of the features to characterize a 2 image is one of the solutions of a CBIR system. Multiple methods have been introduced for each of these visual features and each of them describes the feature from a different view. The main low-level features are three such as Color, Texture and Shape.

The goal line of feature extraction is to improve the efficiency, effectiveness of analysis and this can be done by:

1) Removing redundancy in the image data.
2) Removing in consistency in the image data that is of small no value in organization even removing complete images if that is proper.
3) Reformation of the data (in feature) in order to enhance the performance.
4) Extracting spatial information (color, texture, shape,) which is critical to target identification.

Color: Color is the basic features for the content of images. By way of the color feature human can recognizes and differentiate between object and images. Colors feature used in image retrieval because they are dominant descriptors and provide dominant information about images. To extract the color features from the content of an image, we need to select a color space for extraction. Colors are defined in three-dimensional color spaces as of RGB color space is the most prevalent. The main problem of the RGB color space is that it is perceptually non-uniform and device dependent system. The HSV color space is a native system, which defines a specific color by its hue, saturation, and brightness values. The choice of color features depends on the segmentation results. If the segmentation provides objects which do not have homogeneous color, then at that time color is not a good choice. For the explanation of color feature many techniques can be used. They are Color Histogram, Color Moment, Color Correlogram, Color Coherence Vector, Invariant Color Feature etc.

Color Histogram : The color histogram is easy to work out and effective in describing both the global and local distribution of colors. It's also robust to translate and rotate the image and only change with the scale. A color histogram represents the distribution of colors in an image, each histogram bin relates in the color space. Color histograms are a set of bins where every bin characterizes a 3 specific color of the color space is used. The number of bins be determined by the number of colors in the image. A color histogram for an image is defined as, H= {H[1],H[2], H[3],H[4],…….H[i],……H[n]} Where i denotes the color bin in the color histogram and H[i] denotes the number of pixels of color, and n is the total number of bins used in the color histogram. If two images have exactly the same color proportion but the colors are scattered inversely, then we can't retrieve that image correctly and this is the main drawback of color histogram.

Color Moment: To overcome the quantization problem of color histogram, color moments are used as feature vectors for image retrieval. The image is divided into three equal regions from each of the three regions we extract each color distribution. The efficient and effective representing color distributions of images color moments can be used with three regions are as follows: i. Mean ii. Standard deviation and iii. Skew-ness.

Color Coherence: Vector A color coherence vector is a histogram which partitions pixels according to their spatial coherence. According to image pixel, each pixel within the image is partitioned into two types, i.e., coherent or incoherent. In the feature vector including some spatial information histograms can be produced for both coherent and incoherent pixels. Due to spatial information, it has been shown that CCV provides better retrieval results than the color histogram.

Color Correlogram: The color correlogram was projected to describe not only for the color distributions of pixels, but also the spatial correlation of pairs of colors. From the three-dimensional histogram the first and second are the colors of any pixel pair and third is for spatial distance. If we consider all the probable groupings of color pairs the size of the color correlogram will be very large, for that reason a simplified form of the Feature called the color autocorrelogram is used. The color Autocorrelogram only captures identical colors of the 4 spatial correlation and hence reduces the dimensions. As compared to the color histogram and Color Coherent Vector, the color autocorrelogram provides the greatest retrieval results, but there is effect of the high level of computational expensive due to its high dimensionality.

Invariant Color Feature: Color not only redirects the material of surface, but also differs significantly with the change of radiance, the direction of the surface, and the observing geometry of the camera. But, invariance to these factors is not reflected in most of the color features which are familiarized above. In recent times invariant color feature has been presented to content-based image retrieval. Once applied to image retrieval these invariant color feature might produce Illumination, scene geometry and viewing geometry independent representation of image but also some loss in discrimination power in the middle of images.

Texture: Texture can be responsible for the measure of properties such as, coarseness, regularity and smoothness. Texture can be stated as repeated patterns of pixels in excess of a spatial domain. If the texture has visible with some noise, the patterns and their repetition can be random and unstructured. Many different methods are recommended for computing texture but in the middle of those methods, no one method works best with all types of texture. Some common methods are used for texture feature extraction such as Wavelet Transform, Gabor Wavelet Transform, Fourier Transform, Gabor Filter Feature, Wold Feature, Tamura Feature etc.

Tamura Feature: The Tamura features, including coarseness, contrast, directionality, linelikeness, regularity, and roughness, are considered in accordance with psychological studies on the human awareness of texture. Coarseness is an amount of the granularity of the texture. Using histogram-based coarseness illustration can impressively increase the retrieval performance. This variation creates the feature capable of distributing with an image which has multiple texture properties, and therefore is more useful to image retrieval from large scale database. The first three constituents of Tamura features have been used in some early well-known image retrieval systems, such as QBIC.

Wold Feature: Wold feature provides another method to describing textures in terms of perceptual properties. The three Wold components harmonic, evanescent, and in deterministic, correspond to periodicity, directionality, and randomness of texture respectively. Periodic textures have a strong harmonic component; vastly directional textures have a strong component, and less structured textures. In spatial domain wold feature involves reducing a cost function, and resolving a set of linear equations and in the Frequency domain, Wold components can be acquired by global thresholding of Fourier of the image. This method is considered to accept a variety of in homogeneities in natural texture designs.

Wavelet Transform : Like Gabor filtering, the wavelet transform provides a multi-resolution approach to texture analysis and classification. The computation of the wavelet transforms consists of recursive filtering and sub-sampling. At every level, the signal is decomposed into four frequencies. LL, LH, HL and HH, where ‘L’ indicates low frequency and ‘H’ indicates high frequency. For texture analysis wavelet transform has been divided into two major types pyramid structured wavelet transform (PWT) and the tree-structured wavelet transform (TWT). Mostly important information appears in the middle frequency channel of texture feature, to overcome this problem; the Tree Structure Transform decomposes other bands like LH, HL or HH when required.

Fourier Transform : Fourier Transform has been the best significant part of signal transform. It converts a signal from the time domain into the frequency domain to measure the frequency components of the signal. In CBIR Fourier Transform was used to mine texture features from high-frequency constituents of the image. Inappropriately, Fourier Transform unsuccessful to capture the information about objects locations in an image and could not provide local image features. The Fourier Transform is a significant image processing tools which is used to spoil an image into its sine and cosine components.

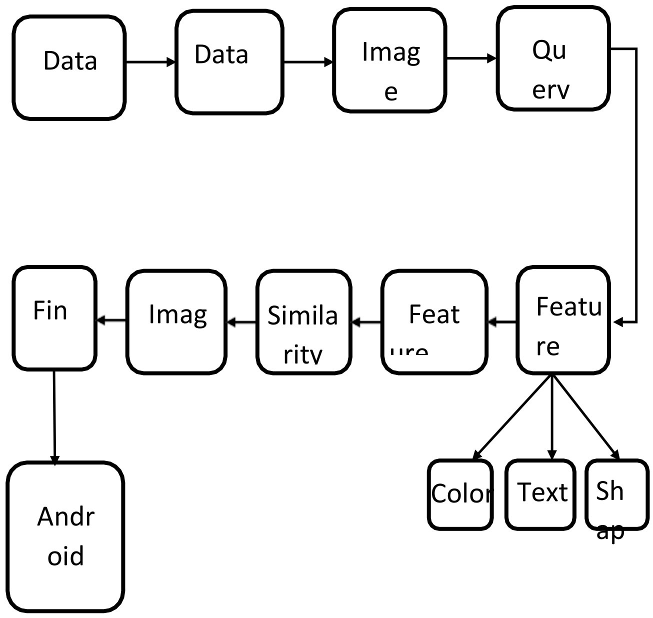
Gabor Filter: For the extraction of Texture Features mostly used statistical method is the 6 Gabor filter. This is most functional method for the Texture. Through Gabor filter many methods planned to characterize textures of images. In most of the CBIR systems constructed in Gabor wavelet, the mean and standard deviation of the distribution of the wavelet transform coefficients are used to build the feature vector.

Shape : Another important visual feature is Shape. Shape is the basic features used to describe image content. Shape's representation and description is a difficult task because one dimension of object information is lost when a 3-D object is proposed onto a 2-D image. The purpose for selecting shape feature for concerning an object is because of its essential properties such as identifiability, invariance, and reliability, accordingly shape has verified to be a favourable feature based on which retrieval of image can be performed. There are some shapes Descriptor methods like Zernike Moments, Fourier Descriptor, Gradient Vector Flow, Geometric Moments etc.

Fourier Descriptor: Fourier descriptors define the shape of an object with the Fourier transform of its boundary. The co-efficient of Fourier transform are called Fourier descriptor. For applying Fourier transform standardized the boundary points of all the shape in database. Robustness is pleasant properties Fourier descriptor. It is capable to capture some perceptual structures of the shape and easy to originate. Through Fourier descriptors, coarse shape features or global shape features and the finer shape features are captured by lower order coefficients and higher order coefficients correspondingly.

Zernike Moments : Zernike moments are as same as to the Fourier transform, to develop a signal into series of orthogonal basis. To recover the image from moment based on the theory of orthogonal polynomials has proposed by Teague and also has introduced Zernike moments. Zernike polynomials are derived from the complex Zernike moments. Zernike moments descriptors do not need to know boundary information, creating it proper for more complex shape representation. In of geometric moment's higher order moments are difficult to construct, to overcome that problem Zernike moment's descriptors are created by arbitrary order.

FLOW OF THE SYSTEM



ANALYSIS AND INTERPRETATION

IMAGE RETRIEVAL USING COLOR CONTENT

Color is highly essential material contained in the picture which is utilized as element for quite some time-based picture recovery strategies. Variety is very essentially accessible picture include, which by and large is addressed in type of variety parts

(Color planes) of computerized picture. Here, describes the ways of using "image color averages" and "block truncation coding" for feature extraction in proposed image retrieval systems.

Color Content Comparative performance analysis on Proposed Image Retrieval Techniques: Row-column-forward diagonal mean (RCMFDM) with original + even image was shown to be the best approach for retrieving images based on color averaging using a crossover point of 0.4258. There are several drawbacks to color-averaging-based image retrieval methods, such as the need for all images in the database to be of the same size (so all images must be resized to fit in the query image's dimensions), the feature vector extraction not being immune to rotation variance, and the larger feature vector size (depending on the image size) (if the image is rotated feature vector will change). 'BTC-LUV-Level Kekre's 4' has the best execution with an accuracy review hybrid point worth of 0.465 while utilizing block truncation coding-based picture recovery strategies. This is unrivaled than that of variety averaging-based picture recovery techniques. Also, the element vector of a picture is free of the picture aspects in BTC-based CBIR (no need of all data set pictures to be of same size to that of question picture). There are two ways used in proposed system to extract color content information from images: either through averaging techniques or by truncating them.

IMAGE RETRIEVAL USING TEXTURE CONTENT

Texture is a key component of human visual perception and can be utilized to distinguish image regions. Unlike color and shape features, texture features better reflect the images' macro- and microstructure. Texture representation in mathematical terms understandable by computing applications is a difficult task. It includes structural, statistical, and multi-resolution filtering. The proposed image retrieval strategies based on texture content utilizing three texture representation approaches. The suggested transform wavelet pyramid-based image retrieval methods use multi resolution texture filtering. Transform wavelet pyramids with seven levels are created here utilizing Haar, Walsh, and the novel Kekre wavelet transform. For image retrieval, each level of wavelet in the transform pyramid is treated as a feature vector. The Walsh wavelet pyramid and Haar wavelet pyramid based CBIR techniques outperform. The vector quantization (VQ) codebook represents the image's statistical color distribution. An external example is formed using three image transforms: Haar, Walsh, and Kekre. Combining these texture patterns with image maps (binary for Walsh, ternary for Haar & Kekre) created using global, local, and intermediate thresholds (intermediate 4 and intermediate 9) from image databases, feature vectors are formed. Here, discuss and compares 36 versions of the proposed CBIR approach using three transforms, three texture pattern sets, and four image map creation methods.

Texture Content Comparative performance analysis on Proposed Image Retrieval Techniques: Here, three approaches for texture feature extraction have been proposed and evaluated, and the results have been presented. The best performance in transform wavelet pyramid-based image retrieval methods is provided by the Haar transform wavelet pyramid level 6 with an average precision-recall crossover point value of 0.4043, followed by the Walsh transform wavelet pyramid level 6 with an average precision-recall crossover point value of 0.4041. Performance provided by Kekre's Fast Codebook Generation (KFCG) in CBIR using VQ codebooks outperforms that provided by other codebooks in the same environment. The KFCG with codebook size 256 has the highest precision-recall crossover point value of 0.4871, which is followed by the 0.4848 obtained by the KFCG with codebook size 512, which is the second highest number. The KFCG codebook size 128 (0.47991), KFCG codebook size 64 (0.46758), KFCG codebook size 128 (0.46313), and KFCG codebook size 16 (0.46313) have the third to sixth best performance rankings, respectively, based on the respective crossover point values (0.45516). With texture patterns, the best performance and highest precision-recall crossover point was achieved by using the Haar transform 16 texture patterns with intermediate-4 thresholding, followed by the Kekre transform 16 texture patterns with global thresholding, which had a crossover point value of 0.4489 for CBIR using texture patterns. It is achieved by using the Haar transform 16 texture patterns with global thresholding to achieve the third best performance with a crossover point value of 0.44834. Overall, CBIR with vector quantization codebooks outperforms all other texture-based CBIR approaches suggested in terms of performance with greater precision-recall crossover points.

IMAGE RETRIEVAL USING SHAPE CONTENT

Images' geometric form representation can be considered one of the major images discriminating elements, and it can be utilized as a feature vector for image retrieval when it is combined with other image retrieval factors. For the most part, gradient operators and morphological procedures are employed to extract the boundary of a shape from the image's edges, which are represented by gradients. The use of gradient operators yields a first order derivative of the image in which only one direction of the image's edges can be identified (horizontal or vertical or diagonal). The slope magnitude approach in conjunction with gradient operators is used to obtain the whole border of the shape in the image in the form of connected edges. The proposed image retrieval methods based on the shape content as edges existing in the image retrieved utilizing seven distinct approaches aliasing the shape content. Sobel mask with slope magnitude method (Sobel-SMEI), Robert mask with slope magnitude method (Robert-SMEI), Prewitt mask with slope magnitude method (Prewitt-SMEI), Canny operator with slope magnitude method (Canny-SMEI), edge morphological operations (morphological-EI), top hat transform (Top-Hat-EI), bottom hat transform (Bottom-Hat-EI), top hat transform with slope magnitude method (Botto (Bot-Hat-EI). With the help of the Shape-Edge image, these edge images are directly employed as feature vectors for CBIR. The proposed image retrieval methods using edge image content with BTC, where block truncation coding is done on shape-edge images, are also presented and they have shown to be more effective. Finally, novel image retrieval methods based on shape have been proposed. Walsh texture patterns with even greater performance are developed through the use of image augmentation and the augmentation of image.

Shape Content Comparative performance analysis on Proposed Image Retrieval Techniques: -

Compares the performance of nine proposed shape content-based image retrieval techniques against an edge image (Sobel-SMEI, Robert-SMEI and Prewitt-SMEI). The top three edge image extraction algorithms are considered here. However, the performance of CBIR using shape edge image with BTC is intermediate, better than that of CBIR using shape image with transform. The CBIR algorithms using shape Walsh texture patterns on original and even images produce the best performance. The best performance is shown by CBIR with shape Walsh texture patterns added on original + even image. CBIR with shape edge image has the worst performance, followed by CBIR with shape edge image and BTC, and finally best shape edge image with transform-based image retrieval.

IMPLEMENTATION RESULTS AND DISCUSSION

Content-based feature extraction for query photos in Android-based mobile applications was used to construct a programmed for automatically inserting visually meaningful key words into EXIF metadata. For image processing, the OpenCV library was coupled to LabView. As shown in below figures 1 and 2, we were able to compare the similarity of the images returned from our query photos by using key points and descriptors.

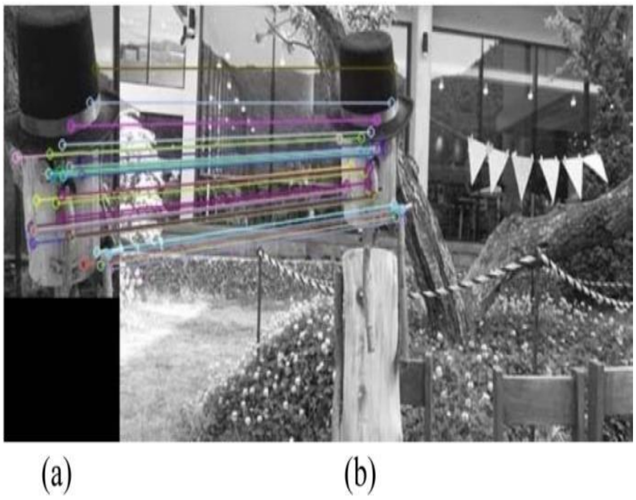


Figure 1: By collecting similar photos, we were able to map the following results: (a) Images were queried; (b) Results were mapped by getting comparable images from the database.

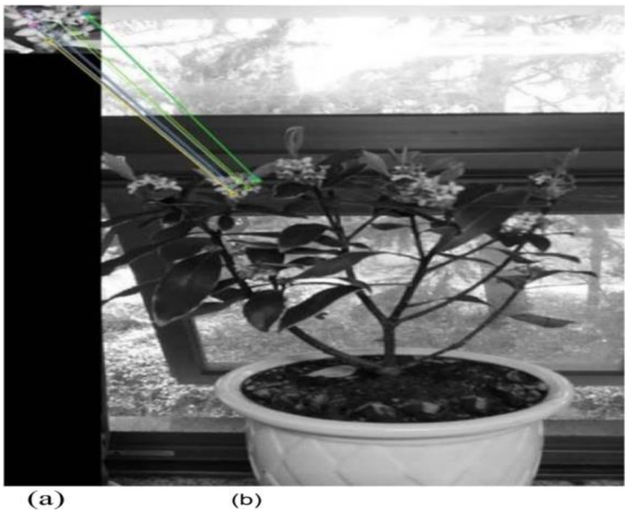


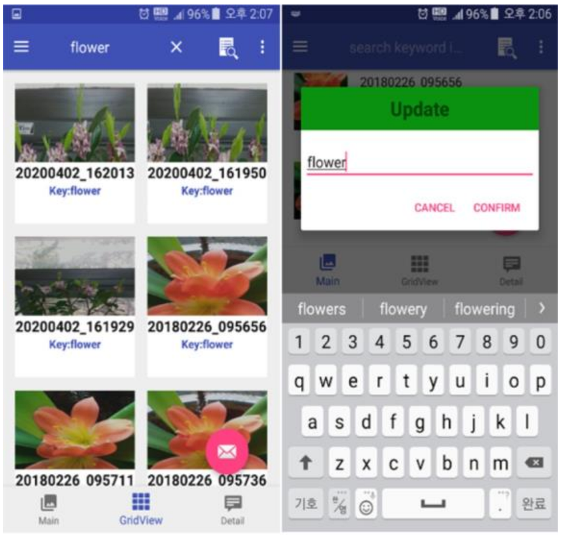
Figure 2: By collecting similar images, we were able to map the following results: (a) Images were queried; (b) Results were mapped by getting comparable images from the database

Figures 1 and 2(a) show the query images, whereas Figure 2(b) displays the mapping results obtained by retrieving similar images from the database.

Following table shows the results of an analysis of application models for image retrieval

Proposed Method	Application field
mobile document acquisition	mobile application
content-based building image retrieval	mobile application
food object recognition and dietary assessment	dietary assessment, mobile application
context-sensitive offloading system	mobile cloud environment
face clustering	mobile application
keyword search by content-based visual processing	mobile application

Images with the keyword "flower" in their EXIF metadata are shown in Figure 3(a) when retrieved from the server side using EXIF metadata. As depicted in Figure 3(b), it is possible for users to edit individual photos' keywords and to edit several images' keywords at once.



(a) Image search result (b) Keyword modification

Figure 3: On a mobile phone, image search results: (a) obtaining an image with the term "flower"; (b) a screen displaying how users may manually modify keywords.

A mobile document acquisition approach was employed rather than a work area scanner, and the benefits of social occasion reports on a cell phone were likewise talked about in the show. In mobiles, a substance-based image search system that utilizations words to lay out coordinating is unique. For health and diet purposes, the information was utilized to recognize food in images recorded by mobile devices and calculate calories. Machine learning inference algorithms and an excellent profiling system are used to make offloading decisions with a serious level of precision. To shield both "information benefactor" and "application client" protection when the cloud server isn't trusted, a new classification system was proposed.

SYSTEM TESTING

In addition to clothing, shoes, and bags, the system primarily gives access to these items. Imaged in Figure 4 is a search result for "clothes" taken using a camera phone. The more similar the images, the higher they are ranked in the commodities image database.

Figure 4: Retrieval of Image with Simple Background

Figure 5: Retrieval of Image with Complex Background

Figure 5 shows how to remove the backdrop from images with complicated backgrounds. The snapshot on the left was taken at random with a smartphone of a purse on a seat (left). The objective region is surrounded by finger portraying, and the framework naturally deducts the foundation and decides the objective tote (focus); at last, by tapping on the inquiry button, the client tracks down satchels of the equivalent or tantamount styles on the Internet (right). On retrieval efficiency, this article focused on the cut-out time and retrieval time at the mobile client without taking into account the time spent on human-computer interaction. Cutout normally takes 5-10 seconds due to differences in phone performance. The retrieval of mega- scale data typically requires 0.01 to 0.3 seconds. Product recovery takes less time, completely matching the constant recovery interest of general online business stages.

SUMMARY AND CONCLUSION

An effective and flexible content-based image retrieval (CBIR) system tailored for mobile settings is presented in this work, showing appreciable gains in retrieval accuracy and speed. The system efficiently transforms picture content into feature vectors for high-precision retrieval by utilizing a combination of visual properties, specifically color, texture, shape, and modified image representations. The suggested methods fared better than conventional single-feature approaches in terms of consistency and response time when tested on a varied dataset with 1,000 photos divided into 11 categories. This work's ability to combine several content descriptors with similarity metrics based on Euclidean distance makes a substantial advance by facilitating a more robust and dependable matching procedure. Additionally, the system presents a new combination of voice interaction and picture recognition, in which speech input starts image searches and outputs visually related results. Additionally, by collecting and using EXIF metadata, it improves mobile image retrieval and helps get around the drawbacks of keyword-based search techniques on portable devices. By tackling fundamental issues in feature extraction, user interaction, and mobile implementation, the study, taken as a whole, advances CBIR technology. Through precise matching, effective processing, and smooth user interaction, it provides a complete, multimodal solution that enhances the retrieval experience and lays the groundwork for further advancements in mobile visual search systems

REFERENCES

1. Ahmad, I., Abdullah, S., Kiranyaz, S., & Gabbouj, M. (2005). Content-based image retrieval on mobile devices. *Proceedings of SPIE - The International Society for Optical Engineering, 5684.* https://doi.org/10.1117/12.596772
2. Ayyasamy, V., Sural, S., & Majumdar, A. (2009). Image retrieval from the web using multiple features. *Online Information Review*, *33*, 1169–1188. https://doi.org/10.1108/14684520911011061
3. Banuchitra, S. (2016). A comprehensive survey of content based image retrieval techniques. *International Journal of Engineering and Computer Science*, *5*(8). https://doi.org/10.18535/ijecs/v5i8.26
4. Bellafqira, R., Coatrieux, G., Bouslimi, D., & Quellec, G. (2015). Content-based image retrieval in homomorphic encryption domain. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 2944–2947). IEEE. https://doi.org/10.1109/EMBC.2015.7319009
5. Bellafqira, R., Coatrieux, G., Bouslimi, D., & Quellec, G. (2016). An end-to-end secure CBIR over encrypted medical database. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 2537–2540). IEEE. https://doi.org/10.1109/EMBC.2016.7591247

6. Bruno, A. (2021). An automatic image content retrieval method for better mobile device display user experiences.
7. Manjunath, B. S., Ohm, J. R., Vasudevan, V. V., & Yamada, A. (2001). Color and texture descriptors. *IEEE Transactions on Circuits and Systems for Video Technology, 11*(6), 703–715.
8. Singh, A., Sohoni, P., & Kumar, M. (2014). A review of different content-based image retrieval techniques. *International Journal of Engineering Research and General Science, 2*(5). ISSN: 2091-2730.
9. Tapkeer, A., Deshmukh, N., Patil, R., Mondhe, V., & Deokate, S. (2015). Content based image retrieval system using a free hand sketch and match images. *Asian Journal of Engineering and Technology Innovation, 3*(6), 33–35. ISSN: 2347-7385.
10. Verma, A., & Sharma, D. (2014). Content-based image retrieval using color, texture and shape features. *International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE), 4*(5). ISSN: 2277-128X.