

Content-based Medical Image Annotation and Retrieval using Perceptual Hashing Algorithm

Suresh Kumar Nagarajan¹, Shanmugam Saravanan²

¹Senior Assistant Professor, School of Computing Science and Engineering, VIT University, Vellore, India

²School of Computing Science and Engineering, VIT University, Vellore, India

ABSTRACT

Nowadays, Medical Imaging is emerged well with computer applications. The diagnosis process is fully automated; based on the medical image, cases are reasoned to examine the causes such as Cancer Prediction, Age Estimation, and Injuries and so on. Automatic Image Annotation (AIA) is widely used in Medical field. Generally, the images are the retrieved using CBIR (Content-based Image Retrieval) technique. In this paper, we are proposing a novel method for the Image Retrieval. This paper uses Perceptual Hash (P-Hash) Algorithm for the similar image retrieval based on the Query-Image. The hash value of the Query image is matched with the medical training-set images and the corresponding keywords are automatically tagged. In this methodology, the system computes the similarity measure or score based on the Hamming Distance of two Hash strings. This methodology is faster and efficient approach for web-based application. In addition to this, robustness of the algorithm could be examined by altering (change in the image orientation) the query-image may not affect the hash value.

Keywords – Automatic Annotation, CBIR, Hamming Distance, Medical Image, Perceptual Hash

I. INTRODUCTION

In recent years, Content-Based Image Retrieval System is mostly used in all sort of domain. CBIR is a method of retrieving the image based on the input image. In this CBIR system, the content of the image is being analyzed attributes such as color, shapes and texture of an image. This technique is done with various algorithms and methods for matching the images from database. Nowadays, CBIR system is a powerful resource used widely in Medical. CBIR has the possibility to give medical doctors with accurate result in diagnoses. For example, CBIR system is used to diagnosis a patient result based on the existing patient disease details. A research says that 93% accuracy of a diagnosis based on content-based image retrieval. The conventional goal of CBIR for medical is to retrieve the closest images to the query-image from the dataset. This concept is also called as Query by Image content (QBIC). In case of reasoning in medical images, the diseases are labeled with image which could be used for disease based classification with respect to the input image. Some of the CBIR medical applications are Cancer Detection, Bone Age Estimation. The large database CBIR systems are namely PACS, DICOM which had 12,000 medical^[1].

Automatic Image Annotation (AIA) is a process assigning the information or data to an image. It is done with the caption or the keyword of the image. Machine learning techniques are commonly used for the image classification and image feature analysis such as segmentation and so on. In this image annotation method, a general term is usually quoted called 'blob'. A blob is a part of an image with a vocabulary meaning. The image is separated into blobs based on the region or cluster and the corresponding blobs are labeled with a vocabulary. In the medical imaging, image annotation helps the doctor to find out the description of an image. For example, in the case of spine

x-ray image each section could be annotated and it would be useful for the diagnostic to examine the case based on the image description^[2].

Phash (Perceptual Hash) is a fingerprint of an audio and video file formed from several characters from its content. Phash is a robust algorithm mostly used for similarity identification. Phash is a hashing function which could be used in Crypto-hashing, Digital Watermarking and Digital Signal Processing. There are four types of phash algorithm used: 1. DCT (Discrete Cosine Transform) based Hash, 2. Marr-Hildreth Operator based Hash, 3. Radial Variance Based Hash, 4. Block Mean Value based Hash. In this paper, we discuss on Average Perceptual Hashing Algorithm which is similar like Block Mean Value based Hashing Algorithm^[3].

The content similarity is identified using Hamming Distance technique. Hamming Distance is a method two find out the difference between two strings, discussed in below sections.

IRMA (Image Retrieval in Medical Applications) is a Diagnostic Radiology Centre. The dataset is collected from IRMA which contains all sorts of Radiological images in digital format. Various categories of images are seen in the dataset. All the images are given in PNG format^[4].

II. RELATED WORKS

Content Based Audio Retrieval using Perceptual Hash algorithm was proposed by Qiong Li, Jing.Wu, Xin He. There are two stages exist: first, the indexing the perceptual hash values and removing the immaterial data is done using n-gram algorithm. Secondly, the matching audio hash value is retrieved correctly using Bit Error Rate (BER) technique. N-gram Algorithm is a novel one and quite commonly used in speech recognition and particularly meant for indexing.

The n-gram contains n units of audio frame in a five bits size. The n-gram indexes of audio are stored in a database with a perceptual hash values. The n-gram and perceptual hash value of a query signal is matched with candidate selection and also with n-gram in audio database which results to similar audio by calculating the BER (Bit Error Rate). The paper shows that, similar audio retrieval using perceptual hash gives the 100.0% precision and 98.33% recall where as Bit Error Rate of 0.15. Fig. 1 shows the architecture diagram of Content-based audio retrieval [5].

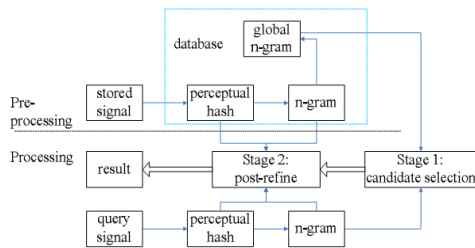


Fig.1. Architecture of Content-base Audio Retrieval

Automatic medical image categorization is also done using CBIR and data mining. The research article says that, the accuracy of the image categorization examined with best ten matches is 97.7%. Texture measure is done using image feature techniques. Image similarity is evaluated based on the distance or score compared with images in the database. The automatic classifier process is processed using k-nearest-neighbor which fixes the distance measures and also the classifier prefers the exact category for an input image based on votes over the k-references near to featured distance measures. The best recognition rate is 43.9% for the DCT-based features. The system categorized with 6335 images in which there are 80 categories and also evaluations are done for 54, 57 and 70 categories of radiology images. Within in ten good matches the accuracy of the system is noted as 98.0 percentages [6].

The research introduced a new algorithm for medical image annotation which is called learning-based algorithm. The proposed algorithm is robust and achieved the high accuracy rate while performing the image classification using spatial configuration models technique. Chest radiograph images were compared and results to the best accuracy for all kind of classes. The sparse spatial configuration (SSC) algorithm is logical and authentic to remove the deviations and locals are held back. The reasoning module finds the orientation and content of an image and it is been classified by classification task. The result is compared with various methods namely Lehmann, Boone, Improved Projection Profile, Subimage pixel Intensity, PatchBOF and SIFTBOF. Most of the comparison result shows the difference the accuracy range between 96.05% - 99.93%. In case of PA-AP/(Posteroanterior/Anteroposterior) LAT based radiograph test shows the result with $\pm 0.40\%$ when compared with above mentioned systems [7].

III. PROPOSED WORK

In this paper, we implemented the system using an open-source platform. The system is implemented in PHP

scripting language with Apache Server and MySQL database. Fig. 2 depicts the overall view of this system. As we proposed a web based application, the user interfaces are rich web elements which make the system user-friendly. In addition to this, our system could be treated as client-server model so that it could be used in diagnosis centre.

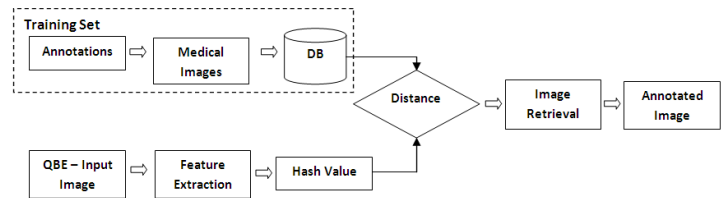


Fig.2. Overview of Our Proposal System

3.1. Web Server

In this methodology, we used WAMP Application which consists of Apache Server, MySQL and PHP. The system is tested in Windows and Linux (Ubuntu) operating system. We used Apache 2.2.21 version server which supports 5.3.9 PHP version scripting language and also we used MySQL as a backend of 5.0.8 version.

3.2. Training Sets

The training sets consist of predefined medical keyword and sets of images. In our study, we taken 1926 images provided by IRMA medical research centre. All the images are in PNG format of various dimensions. The images are classified into several categories such as Skull, Bone Joint, Hand-Wrist, Chest-Abdomen, Breast, Foot and Spine. The images are categorized based on a unique id as an image file name. Table 1 shows the total number of image and categorization used in this system.

Table 1. Image Categories

S.No	Category	Total Images	File Format
1	Bone Joint	118	PNG
2	Breast	302	PNG
3	Chest-Abdomen	440	PNG
4	Hand-Wrist	303	PNG
5	Foot	120	PNG
6	Pelvis	98	PNG
7	Skull	499	PNG
8	Spine	60	PNG

3.3. Architecture

Fig. 2 shows the architecture of our system. The system accepted only PNG images for the processing. The system contains different process in which input image is examined and result is obtained. The system processes are explained in below sections.

3.3.1. Input Image

The user inputs the image via web rich file upload element. Users are allowed to select the file from the local drive. The system is validated such way that it accepts only PNG images. If the users selects or uploads an invalid image

format then the system alerts the users through notifications. Fig. 3 shows the sample screenshot of our system.

3.3.2. Feature Extraction

In this process the input PNG image is stored in a temporary directory of the web server and image processing techniques are applied to the user query-image namely grayscale conversion and resizing. Grayscale conversion is done at this process using in-built function of PHP. After the color conversion, the image is resized to a define size.

3.3.3. Hash value

Based on the color pixels of a feature extracted image hash value is generated. Perceptual hashing generation method is used and the hash value is represents in binary and hexadecimal format. Perceptual hashing algorithm and methods are discussed in the following sections.



Fig.3. Screenshot of Hash generation

3.3.4. Similarity Measure

The query-image hash string is compared with the training set images. Similarity is measure in the basis of hamming distance score. Based on the score the similar images are retrieved. If the hamming distance score is too far when compared to query-image hash string then it is neglected.

3.3.5. Image Retrieval and Annotation

In the webpage, the results are shown to user which contains the similar images and the corresponding hamming distance score. The system gives an option to examine the precision and recall by selecting the matched and unmatched image for all retrieved images via select box so that query result is evaluated.

3.4. Perceptual Hashing

Perceptual hashing is a robust algorithm widely used for content identification. In our system, we used block-mean perceptual hashing algorithm where the algorithm takes the color pixels of an image to generate the hash value. The study says P-hash is reliable and fastest algorithm. In this methodology, similar like Block-Mean value Perceptual hashing image color mean value hashing technique is used. Where the hash string is generated using the color of an images. Pixel color average is taken as mean value based on the average rate the hash string is formed. The image is scaled to defined small size. We generate the hash value by observing the color pixel and its average color

pixel value. Hash bit is generated, with respect to the average of total color pixels and each color pixels of the image^[3].

3.5. Algorithm for Hash Generation

The pseudocode of the hash generation is given below

Algorithm of Image P-hash generation

```

I ← input Image
I ← convertGrayscale(I)
I ← scale to a defined size of w and h
for i ← 0 to w
  for j ← 0 to h
    sum ← sum + colorpixel(i,j)
  colors[] ← sum
  avg ← sum / (w * h)
foreach pixval in the list of colors
  if pixval > avg then
    hashbit ← 1
  else
    hashbit ← 0
  loop ← loop + 1

if not loop % 4 then
  hash ← hash ++ hexadecimal(hashbit)
    
```

In the above mentioned algorithm colorpixel is a function used to get the RGB color values of the current pixel position. We implemented the above algorithm in PHP scripting language.

3.6. Hamming Distance

Hamming distance is a measurement of two different strings. The strings can be collection of characters, numbers or binary coded numbers. Let A denote an alphabet of finite length. P and Q are the finite string which belongs to A. where $P = (p_1, p_2, p_3, \dots, p_n)$, $Q = (q_1, q_2, q_3, \dots, q_n)$ The Hamming distance (Δ) of two String P and Q is defined as

$$\Delta(p, q) = \sum_{p_i \neq q_i} 1, i = 1, 2, 3, \dots, n$$

In other words, if the bit of two strings are not same then record as 1 else record as 0 and at last sum the number of one's and zero's to find out the hamming distance score.^{[3][8]}

3.7. Annotated Keywords

The original annotated keywords are used for automatic annotation. The keywords are classified and categorized based on the image id. Each category has a unique id and in the case of images, it is categorized based on the id. The automatic annotation works depend upon the retrieved images and its id with corresponding categorized keyword.

IV. RESULT AND DISCUSSION





The result of automatic annotation with precision and recall shows in Table 3. The average execution time of the proposed system in web-based environment is 41.17 seconds where database contains 1926 images. The precision and recall are depicts in Fig. 6 and Fig. 7. The proposed system has the mean precision of 0.62 and the mean recall of 0.31. We examined that, as per the query image and retrieved result are closely matched with corresponding annotation to 80%. Fig. 4 shows the screenshot of web-based CBMIR and Automatic Annotation. The similarity hamming score of an image lies between the ranges 1 - 11. Beyond the defined ranges are falls under the dissimilar images.



Fig.4. Screenshot of CBMIR System




The system robustness is tested such way that the input image is tested in different orientation and a sample result of oriented Skull image is shown in Table 2.




Table 2 Result of Skull Image with different orientation

Input Image				
Orientation	0°	90°	180°	270°
Precision	0.83	0.89	0.81	1.00
Recall	0.0508	0.00882	0.0186	0.00311
Number of Retrieved image	112	19	44	6
Matched Images	98	17	36	6

The below table describes the sample images and the automatic annotated keyword with precision and recall.

Table 3 Sample Result of Automatic Annotation

Input Image			
Original Annotation	Chest-Abdomen	Pelvis	Hand-Wrist

Automatic Annotation	Chest-Abdomen	Pelvis	Hand-Wrist
Precision	0.94	1.00	0.75
Recall	0.008	0.006	0.0015
Image			
Original Annotation	Skull	Bone Joint	Foot
Automatic Annotation	Hand Leg Skull	Bone Joint Hand -Wrist	Foot Skull
Precision	0.837	0.8	0.0927
Recall	0.050	0.0020	0.0046

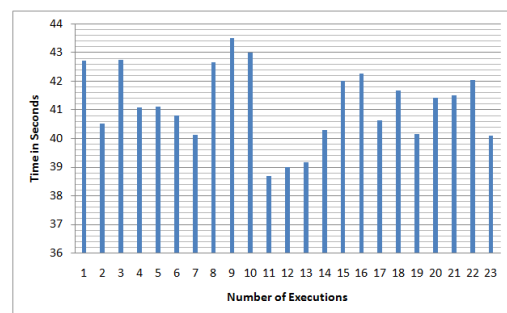


Fig.5. Execution Time given in Seconds

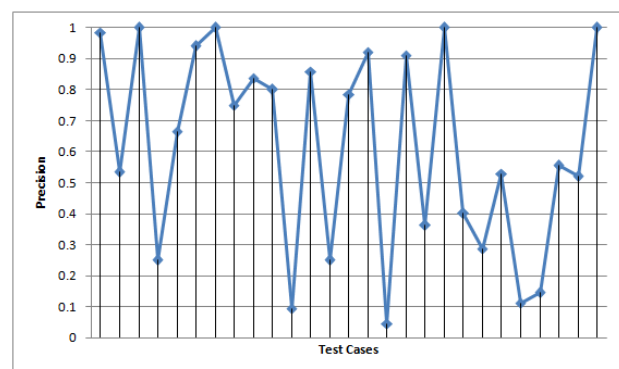


Fig.6. Precision Result of the proposed system

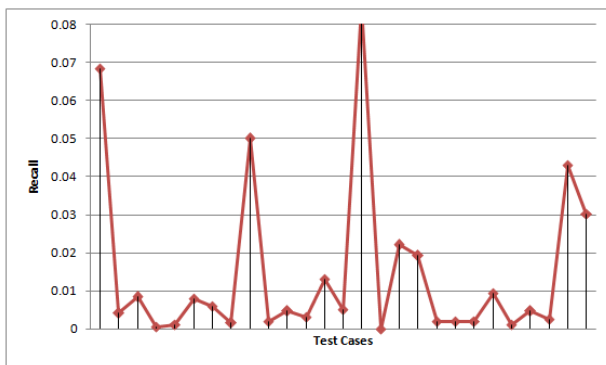


Fig.7. Recall Result of the proposed system

The result reveals that the proposed system is accuracy and the robustness of the perceptual hashing algorithm is proved. The robustness of the system has been examined with the query image, by changing the orientation of images in different angles result to similar outcome.

V. CONCLUSION

We conclude that Perceptual hashing algorithm has very high robustness and in the CBIR it would be useful. This novel approach could be used in Medical diagnostic centre or Research Medical Hospitals for case-based reasoning. Web Application plays a vital role in all domains as a matter of fact it is a portable application used by all the actors via web. Content-based Image Retrieval is an emerged technique but the application of CBIR is yet to attain advancement and it is an open research topic in the medical area. Day by day the image dataset in the medical data centre is increasing drastically so the amount of time taken to compare whole database will be raised. To overcome such kind of problem an effective algorithm is perceptual hashing.

ACKNOWLEDGMENT

We would like to thank IRMA (Image Retrieval in Medical Application) for providing the needed radiological image dataset and VIT University.

References

- [1]. Müller, H., Michoux, N., Bandon, D. and Geissbuhler, A. "A Review of Content-Based Image Retrieval Systems in Medical Applications – Clinical Benefits and Future Directions".
- [2]. Shao, W., Naghdy, G. and Phung, S. L. "Automatic image annotation for semantic image retrieval" Proceedings of the 9th international conference on Advances in visual information systems', Springer-Verlag, Berlin, Heidelberg, 2007, pp. 369--378.
- [3]. Zauner, C. "Implementation and Benchmarking of Perceptual Image Hash Functions", Master's Thesis, University of Applied Science, Upper Austria, 2010.
- [4]. <http://irma-project.org/>
- [5]. Li, Q., Wu, J. and He, X. "Content-Based Audio Retrieval Using Perceptual Hash" Intelligent Information Hiding and Multimedia Signal Processing, 2008. IHHMSP '08 International Conference on', 2008, pp. 791 -794.

- [6]. Lehmann, T. M., Güld, M. O., Deselaers, T., Keysers, D., Schubert, H., Spitzer, K., Ney, H. and Wein, B. B. "Automatic categorization of medical images for content-based retrieval and data mining," Computerized Medical Imaging and Graphics (29:2-3), 2005, pp. 143 - 155.
- [7]. Tao, Y., Peng, Z., Krishnan, A. and Zhou, X. "Robust Learning-Based Parsing and Annotation of Medical Radiographs," *Medical Imaging, IEEE Transactions on* (30:2), 2011, pp. 338 -350.
- [8]. Hamming, R. "Error Detecting and Error Correcting Codes," *Bell System Technical Journal* (26:2), 1950, pp. 147--160.