Journal of Computer Science 10 (8): 1488-1496, 2014
ISSN: 1549-3636

doi:10.3844/jcssp.2014.1488.1496 Published Online 10 (8) 2014 (http://www.thescipub.com/jcs.toc)

FAST IMAGE MATCHING IN HUGE DATABASE

[1]Nidhal K. El Abbadi and [2]Lamis Al Saadi

[1]Computer Science Department, Education College, University of Kufa, Najaf, Iraq
[2]Computer Science Department, Education College, University of Babylon, Babylon, Iraq

Received 2014-01-28; Revised 2014-01-29; Accepted 2014-04-02

ABSTRACT

In today's multimedia age, images are becoming more and more common in various fields. Database is an area that has found extensive use of images in past few years. With the increased usage of image data, the need to search and matching images efficiently and accurately from a large image database or a cluster of images databases has increased. Image search has become a popular feature in many search engines, including Google, Yahoo!, MSN, etc., majority of which use very little, if any, image information. In this study we introduced a novel algorithm for searching for image in huge database based on SVD, the proposed algorithm solved some of the other searching algorithms problems, like problems of image scaling, image rotation, search time. The contribution of this algorithm is to retrieve the exact image and/or the image with different view or captured in different time with fixed time regardless the database size, database can be indexed according to the Suggested Values extracted from SVD. The proposed algorithm gives high performance and solved many issues related to other image search algorithms.

Keywords: SVD, Image Processing, Image Registration, Search Engine

1. INTRODUCTION

Image database is increasing day by day, because searching images from large and diversified collection using image features as information is difficult and imperative problem. Most of search engines today used the feature of image search and almost these search engines based on text for image search. Retrieval process results in commercial image search engines based on text. Text search based is more popular than using image features in the image retrieval process; this is due to difficulties in extracting and analysis image features, which need high cost and time consuming. However, it frequently finds irrelevant results, because the search engines use the insufficient, indefinite and irrelevant textual description of database images (Pushpanjali and Prashant, 2013).

Image matching and retrieval has been adopted in most of the major search engines, including Google, Yahoo!, Bing, etc. Most of image search engines index the images mainly on the image name and texts around the images, the text places in the image are only in two main places which are in the name of image (Title) and in the tags which are proposed and implemented using web 2.0 concepts? Most of the time user make query in the text format for search contents over any search engine (Gupta *et al*., 2012).

The process of matching two or more images captured from different viewpoints, at different times and under different environment conditions (lighting, different sensors) called image registration. Once feature correspondences have been established and the geometric alignment has been performed, the images are combined to provide a representation of the scene that is both geometrically and photo-metrically consistent. Image registration is a key problem in computer vision that shows up in a wide variety of applications such as image mosaicing, medical image analysis, face tracking, handwriting recognition, stereo matching and motion analysis (Taylor and Bhusnurmath, 2008).

Recent advances in computing and communication technology are taking the actual information processing tools to their limits. Great accumulation of digital data was grown in the last years such as images, video and audio. A large of images databases used for many applications such as satellite and medical imagery, where it is often difficult to describe or to annotate the image content.

Corresponding Author: Nidhal K. El Abbadi, Computer Science Department, Education College, University of Kufa, Najaf, Iraq

The Internet is good example of distributed databases containing many millions of images.

It is easy to ordering, indexing and retrieve the image. While information contents of images in the scientific database applications is not explicit and it is not easy to classification, indexing and retrieving. In particular, the huge image databases regards as the most challenging problem in the field of scientific databases. In this field there are many research, some of them are:

Chouragade and Prashant (2013) present a new method called visual reranking. This approach depend on analyzing the distribution of visual similarities among the images and image ranking system that finds the multiple visual themes and their relative strengths in a large set of images. In this method the retrieved images resulted from searching will be mostly the relevant images, this approach will reduce the number of irrelevant images. Also, the ranked images resulted by database searching based on text search, images features will be extracted to help rerank the images.

Jiang *et al*. (2013) Presents new method that enables ranking the retrieved images that has equal hamming distance to the queries. This is implemented by firstly offline learning bitwise weights of the hash codes for a diverse set of predefined semantic concept classes. They formulate the weight learning process as a quadratic programming problem that minimizes intra-class distance while preserving inter-class relationship captured by original raw image features. Online computing of query adaptive weights for both query and semantic concept classes are measured.

It is easy to order the retrieved images by weighted Hamming distance when using the query-adaptive bitwise weights.

Matungka *et al*. (2009) Introduced new technique called Adaptive Polar Transform (APT) that samples the image in the Cartesian coordinates. Combining (APT) with an innovative projection transform along with a matching mechanism, the suggested approach enhanced the LPT method to reduce the computational load and increased the accurate registration. This method presents retrieving search scheme based on Gabor feature to minimize the search space for recovering the translation of the model image in the target image.

1.1. Singular Value Decomposition

Singular Value Decomposition (SVD) method can transform matrix (A) into product USV^T, this produced to refactoring a digital image into three matrices and allows us to represent the image with a smaller set of values, which can preserve useful features of the original image, but use less storage space in the memory (Lijie, 2007).

The SVD is one of the strongest mathematical tools that can be used to decompose any square or non-square matrix (A) into the multiplication of two unitary matrices U and V and one Diagonal matrix S as shown in **Fig. 1**. SVD can be used to apply on any images to extract its useful features by decomposing the digital image matrix into three orthogonal matrices, which reduces the high dimensional matrix into small dimensional matrix (Malakooti *et al.*, 2013).

Singular Value Decomposition (SVD) is said to be a significant topic in linear algebra by many renowned mathematicians. SVD has many practical and theoretical values; Special feature of SVD is that it can be performed on any real (m, n) matrix. Let's say we have a matrix A with m rows and n columns, with rank r and $r \leq n \leq m$. Then A can be factorized into three matrices:

$$A = USVT$$

where, Matrix U is an m×m orthogonal matrix:

- U = [u1 , u2 ,...ur , ur+1 ,..., um]. Column vectors ui for i = 1, 2, …, m, form an orthonormal set
- And matrix V is an n×n orthogonal matrix
- V=[v1 , v2 ,...vr , vr+1 ,..., vn]. Column vectors vi for i = 1, 2, …, n, form an orthogormal set

Here, S is an m×n diagonal matrix with Singular Values (SV) on the diagonal (Lijie, 2007).

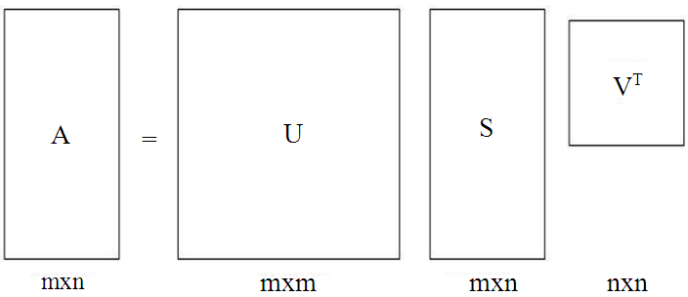

Fig. 1. Illustration of Factoring A to USV T

2. MATERIALS AND METHODS

Searching for image by Google image search engine face some of issues when retrieve image, the main problem with Google image search is retrieve different images sets before and after the image rotated. To explain this issue let us take image1 as shown in **Fig. 2** and searching the Google to find the same image or the similarity, the results shown in **Fig. 3.** Now if the image1 rotated with 90° and searching the Google for it, the results shown in **Fig. 4** which is totally different from the results of image1 before rotated. Group of images (visually similar images) which displayed with Google search will be also different from the same group for the images before rotated. Same image (origin image before rotation) may retrieve in group of (visually similar image). Even the best guess (which display with Google search) for this image is different. Same thing happened when searching for image1 after rotated 180°, the results different from the results when searching image1 and results when searching for image1 rotated with 90° as shown in **Fig. 5**. Also the results differ when image1 rotated 270° from the previous results as shown in **Fig. 6**.

Another issue is when the input image not exists in images database, at this case the retrieved images will be set of images not related with origin image and may be totally different as shown in **Fig. 7**, also we noted that searching time will be very long compared with searching time for image exist in images database, it may be more than 20 times more than time needs to find exist image in the same images database.

Also when we search for exist image but with different file extension like (*.bmp, *.png, *.tif) at this case more time need to match image in Google search.

The suggested algorithm will be solve all these issues and it is applied for both grayscale image and colored images. The color images either converts to grayscale image or take each color band separately (Red, Green and Blue). The input image will be transformed by using SVD transformation. As we mentioned above the result of SVD transformation will be three matrices (U, S and V).

The proposed algorithm work on the diagonal matrix (S) resulted from the SVD, we counts the norm of this matrix; the norm is the square root for the summation of square of each diagonal value in the S matrix, which we called in this study “registration value”. Always we get one unique number for each image. We suggest in this study to included the registration value for the images store in the database with the name of image as image index (either part of name or the registration value represent the image name) and then the images in the database will be sorted in ascending order according to registration value. To search for an image we first determine the registration value for input image and then compared this value with the registration values of images in database, at this case the direct access to any image in the database will be implemented according to the registration value and then fix time need to access the image in database if it is exist regardless the location of image in the database.

Registration value is fixed for image when its rotated, scaling and changing its extension and the same time for matching exist image or related images and not exist image.

Fig. 2. Image1

Fig. 3. Result of searching for image1 in Google search

Fig. 5. Result for searching rotated image1 with 180° in Google search

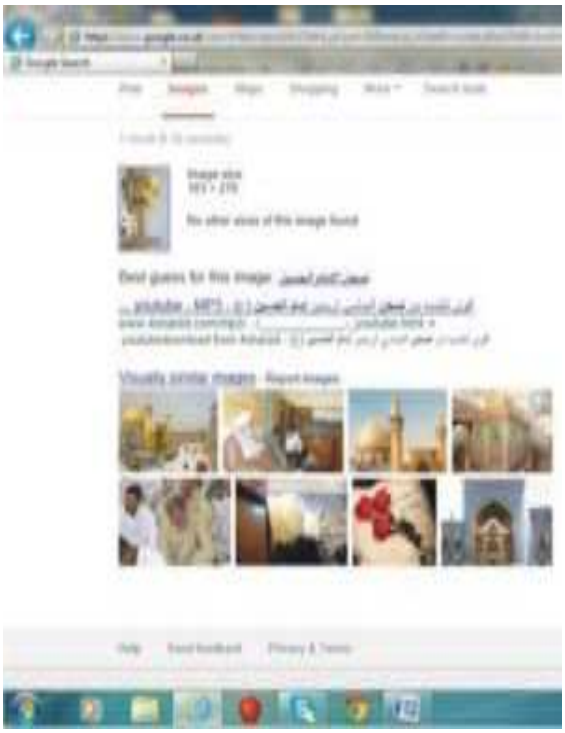

Fig. 4. Result for searching rotated image with 90° in Google search

Fig. 6. Result for searching rotated image1 with 270° in Google search

Fig. 7. Result for searching for image not exists in images database by Google search.

3. RESULTS

To test the proposed algorithm we choose the image2 in **Fig. 8** and determine the registration value for it. Same image will store with different file extension and then determine the registration value the result shows in **Table 1**.

Image2 will be rotated with 90o and 180o as shown in **Fig. 9**. After rotation we determine the registration value for each image and the results shows in **Table 2**.

The other experiment is searching for an image, the input image was the image3 in **Fig. 10** and the retrieved images are the images in **Fig. 11, Fig. 12, Fig. 13 and Fig. 15**. The searching processes to find the matched image start with the determining the registration value for image3 and comparing it with the registration values which embedded in the images names for the images in databases. The registration values shows in **Table 3**.

4. DISCUSSION

From the results, it is clear that registration value almost the same for any image independent on the image extension as shown in **Table 1**.

Table 1. Images with various file extensions, and their corresponding registration values

Image extension	Registration value
*.jpg	57.59
*.png	57.59
*.bmp	57.59
*.tif	57.59

Table 2. Rotate images and their corresponding registration values

Images	Registration value
Image2	57.59
Image2 rotated with 90°	57.58
Image2 rotated with 180°	57.58

Tabel 3. Registration value for image3 and retrieved images

Image number	Registration values
3	652.9931
4	652.9929
5	652.9931
6	652.9931
7	652.9931
8	651.7668

Fig. 8. Image2 with extension (*.jpg)

Also the algorithm solves the Google issue for the image rotation and we proved that as shown in **Table 2**. Also if we compare the results of our algorithms we discover that in our algorithm the correctness of matching image when it is rotated was 100%, while this ratio will differ according to the algorithm used as shown in **Fig. 16** (Manzar *et al*., 2008).

(a) (b)

Fig. 9. Image in Figure 8 after rotated: (A) 90°. (B) 180°

Fig. 10. Image3

Fig. 11. Image4

Fig. 12. Image5

Fig. 13. Image6

Fig. 14. Image7

Fig. 15. Image8

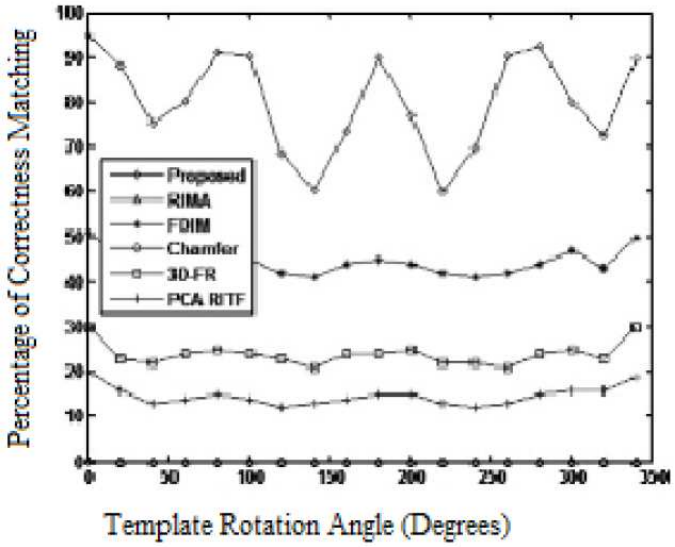

Fig. 16. Image correctness matching when the image rotated with different angles

Another issue with Google search is the increasing search time when searching for image with different file extension, or when searching for image does not exist in database. In all these cases the access time in the proposed algorithm is fixed regardless the database size.
One of the important results in this research is the ability to retrieve images with different views or captured in different times as shown in **Fig. 10-15** and **Table 3**.

5. CONCLUSION

In this study we suggested novel search algorithm for image in images database based on SVD, which solving some issues in some of the searching engine such as Google image search. The algorithm depend on direct access to any image in database regardless the location of image in the database and then the search time will be fixed whether the image in database or none and this solved one of the Google search image problem which take very long time if the image not exists in database or image with different extension. The algorithm solved the Google search image problem when rotate the image.

The proposed algorithm suggested indexing the images in database according to registration value which embedded in the image name or may be represent the image name.

Although the algorithm work very well, but we may be adopted some limitations which is related to retrieve more images related to image looking for, for that we suggest for future work to develop this algorithm by combining the registration value with some features extracted from the image, this can help to retrieve more images related with the input image.

6. REFERENCES

Chouragade, P.M. and P.N. Chatur, 2013. Visual re-rank: An approach for image retrieval from large-scale image database. Int. J. Adv. Res. Comput. Sci. Software Eng., 3: 471-476.

Gupta, N. D. Vijay and A. Athavale, 2012. Image retrieval using collaborative approach based on random walk and similarity matching by hybrid graph representation. Int. J. Emerg. Technol. Adv. Eng.

Jiang, Y.G., J. Wang, X. Xue and S.F. Chang, 2013. Query-adaptive image search with hash codes. multimedia. IEEE Trans. Multimedia, 15: 442-453, DOI: 10.1109/TMM.2012.2231061

Lijie, C., 2007. Singular value decomposition applied to digital image processing. Division Computing Studies, Arizona state University Polytechnic Campus.

Malakooti, M., F.P. Zahed and S.M. Hashmi, 2013. Image recognition method based on Discrete Wavelet Transform (DWT) and Singular Value Decomposition (SVD). Proceedings of the 3rd International Conference on Digital Information Processing and Communications, (IPC '13), SDIWC Digital Library, United Arab Emirates, pp: 42-47.

Manzar, M.A., T.A. Cheema, A. Jalil, I.M. Qureshi, 2008. New image matching technique based on hyper-vectorisation of grey level sliced binary images. IET Image Process., 2: 337-351, DOI: 10.1049/iet-ipr:20080029

Matungka, R., Y.F. Zheng and R.L. Ewing, 2009. Image registration using adaptive polar transform. Proceedings of the 15th IEEE International Conference on Image Processing, Oct. 12-15, IEEE Xplore Press, San Diego, CA., pp: 2416-2419. DOI: 10.1109/ICIP.2008.4712280

Taylor, C.J. and A. Bhusnurmath, 2008. Solving image registration problems using interior point methods. Proceedings of the 10th European Conference on Computer Vision, Oct. 12-18, Springer Berlin Heidelberg, Marseille, France, pp: 638-651.