



The Republic of Iraq Ministry of Higher Education and Scientific Research University of Al-Qadisiyah College of Computer Science and Information Technology Multimedia department

Retrieving images based on their content analysis

A Graduation project is submitted to the Multimedia department in partial fulfillment of the requirements for the degree of Bachelor in Information Technology in Multimedia

BY:

Haneen Saad

Ahmed Jaseem

Fatima Ali

SUPERVISOR

Dr. Lamia Abed Noor 2018-2017

بسم الله الرحمن الرحيم

(قَالُواْ سُبْحَانَكَ لاَ عِلْمَ لَنَا إِلاَّ مَا عَلَّمْتَنَا إِنَّكَ أَنتَ الْعَلِيمُ الْحَكِيمُ)

صدق الله العلي العضيم سورة البقرة (32)

	1
4	



الشكر لله أولا على نعائمه كلها وعطفه ورحمته ، فلله الفضل في مانحن عليه اليوم. كما أن الشكر موصول إلى والدي ووالدتي لمساندتهم اياي طيلة ايام دراستي. ولابد من كلمه شكر اقولها بحق استاذتي ومعلمتي الاستاذه لمياء عبد نور التي وجدت فيها شخصية الكرم صدرا رحبا وقلبا مفتوحا ، ويدا كريمه كانها انسانه عالمه ، فتحت امامي نورا اضاء طريقي.

اليها اقدم خالص شكري وتقديري ،فهي التي منحتني فرصة الخوض في هذا العنوان المهم ، فلها كل الشكر والعرفان لاشرافها النبيل بنصائحها المتواصله ، ومتابعتها القيمه ، لها الشكر بعد الله تعالى وجزاها الله خير الجزاء .

واخر دعوانا الحمدلله رب العالمين و الصلاة والسلام على سيد المرسلين وعلى اله اجمعين

Contents

page	Subject	Numb
Α	الاية القرانية	1
В	الاهداء	2
С	الشكر و التقدير	3
D	Contents	4
1	Abstract	5
2	Chapter one	6
3	Introduction	7
3	Research problems	8
3	Research goals	9
4	Chapter two	10
5	Introduction	11
5	Digital image	12
5-6	Image processing	13
7	Statistical ladder (Histogram)	14
7	The method of assessing the similarity of images based on histogram analysis	15
7-8	Manhattan distance	16
8-9	noise	17
10	Chapter three	18
11	Introduction	19
11-13	Algorithms	20
14-16	Simplified explanation of user interfaces	21

page	Subject	Numb
17	Chapter four	22
18	Discussion	23
18-19	Conclusion	24
19	Feature work	25
20-21	Recourses	26

Abstract

Retrieving of images is a fruit full field in image processing application ,however there is need for retrieving the original image that would be erupted according to different reason. In this project algorithms were developed to evaluate the similarities between two or more images and to reduce the time spent in the search process based on the analysis of color intensity variation and the histogram of the color of the image these algorithms can be used to search for images that are stored in databases system or computer networks .and retrieve non-distorting images by searching for distorted images. This experiment was implemented using a computer program that was programmed using the MATLAB 2015 The results have been analyzed and illustrated by multiple example.

Chapter one

Introduction

• Introduction :

:

- Research problems :
- Research goals:

1-1 Introduction

At present the process of searching for images in databases and search engines is widely required such as searching the databases for images ,similar images ,original image ,clips ,etc. there are databases dealing with pictures works of art images from satellites and photography as well a collection of international photographs.

The databases used for these images can be very large in size and contain hundreds of thousands even millions of images . in most cases such databases are indexed by keywords these images are stored in the databases with the words they describe and the basic words in the databases are stored by people or user who classify images by specific categories.

1-2 Research problems

The process of searching for images and retrieve non-distorting images based on the index of words and stored in the databases requires a long time and therefore does not provide the necessary flexibility for the process of searching for the desired image where there may be similar images with one name, also in some of situation that distorted images have no indexed words.

1-3 Research goals

The purpose of the research is to reduce the time taken for the process of searching for similar images in the databases is to search for images on the basis of (search by model) when the user select full images that and affected by noise and search by similar images in database and put them as a search guide in the database .computer data and face similarities with this image so as to ensure accuracy in search results .

Chapter two

Research Methodology

- Introduction :
- Digital image:
- Statistical ladder (Histogram):
- The method of assessing the similarity of images based on histogram analysis:
- Manhattan distance:
- noise:
- Retrieval:

2-1 Introduction

Mathematical modeling methods and programming techniques were used for digital signal processing and imaging.

2-2 Digital image

Digital images are made of picture element called pixels. Typically pixels are organized in an order rectangular array. The size of an image is determined by the dimensions of this pixel array. The image width is the number of columns and the image height is the number of rows in the array. thus the pixel array is a matrix of b columns x n rows, to refer to specific pixel within the image matrix. We define its coordinate at x and y . the coordinate system of image matrices defines x as increasing from left to right and y as increasing from top to bottom .compared to normal mathematic convention . The organ is in the top left corner and the y coordinate is flipped .why is coordinate system flipped vertically? Digital images were defined in terms of the electron beam scanning to pattern of television. The beam scanned from left to right and top to bottom other than this historical reason , there is no purpose served by this inversion of the y coordinate .

Any image can be defined as a two dimensional dimension (D,X)where D, x are two level coordinates and denote the value of the function (d) in any pair of coordinates at the gray level or the intensity at these points the image is called a digital image when the values (D,X,d) all belong to a set of specified values called these elements "picture element" or "pixels" here we can clarify the meaning of both "image analysis" and "vision using the computer [21]

2-3 Image processing

Image processing is a method to convert an image into digital form and perform some operations on it in order to get an enhanced image or to extract some useful information form it, it is a type of signal dispensation in which input is image like video frame or photograph and output may be image or characteristics associated with that image, usually image processing system includes treating image as two dimensional signals while applying already set signal processing methods to them image processing basically includes the following three steps

- 1- image with optical scanner or by digital photography.
- 2- analyzing and manipulating the image which includes data compression and image enhancement and spotting patterns that are not to human eyes like satellite photographs.
- 3- output is the last stage in which result can be altered image or report that is based on image analysis .

Purpose of image processing is divided into five groups they are :

- 1- visualization-observe the objects that are not visible
- 2- image sharpening and restration to create a better image.
- 3- image retrieval seek for the image of interest
- 4- measurement of pattern measures various objects in an image.
- 5- image recognition –distinguish the objects in an image.

- Vision using computer: is interested in simulating the ability of human vision including the ability to learn and draw conclusions and take reactions based on visual input .

- Analyze the image: it is a medium area between the vision of the computer and the processing of images difficult to find what separates the separation between these subject(processing of images ,analysis and vision computer) but the processes that use the computer in this area can be divided into three levels:

- 1. low _level operations that include deformation ,improved contrast, and increased image size.
- 2. medium level operations which include dividing the image into regions or elements and then describing these elements to reduce them to a representation of the proper processing of the computer also includes the process of identifying elements of a specific image.
- 3. high level processes these include the process of understanding (making sense) of a group of elements that have been identified and at the top of the processes of this level are the processes of learning and gaining knowledge associated with the vision of the computer [1][2].

2-3 Statistical ladder (Histogram)

It is a descriptive pattern (graph) that shows the distribution of chromatic levels and its proportion in relation to its image the application of the statistical ladder(normalized histogram function) gives a quantitative measure of intensity of each pixel in the picture .The method of statistical analysis is one of the most effective methods for comparing and indexing images so each model has a statistical scale of its own .if the images match their statistical ladders match[3].

2-4 The method of assessing the similarity of images based on histogram analysis:

Most researched area in in the field of object oriented image processing, procedure is efficient and effective image segmentation , segmentation is a process of partitioning a digital image into multiple regions (sets of pixels). According to some homogeneity criterion, in this paper we introduce a spatial domain segmentation framework based on the histogram analysis and soft threshold . the histogram analysis uses discontinuity and similarity properties of image statistics in tandem with distribution of pixels to define the binary label for a homogenous region, the soft threshold used for classification is determined based on the localized statistics of the image in consideration for merging of the regions . simulation results and analysis would verify that the proposed algorithm shows good performance in image segmentation without choosing the region of interest. one of the methods that makes the search process more flexible is that it is possible to introduce histogram of the image and search databases for image similar to this histogram .

2-5 Manhattan distance

the distance between two points measured along axes at right angles. In a plane with p1 at(x1,y1) and p2 at(x2,y2), it is |x1-x2|+|y1-y2|.

The Manhattan distance is the simple sum of the horizontal and vertical components, whereas the diagonal distance might be computed by applying the Pythagorean Theorem

the Manhattan distance between two vectors(or points) a and b is defined as |aibi| over the dimensions of the vectors. this is known as Manhattan distance because all paths from the bottom left to top right of this idealized city have the same distance, the Manhattan distance function computes the distance that would be traveled to get from one data point to other if a grid-like path is followed. the Manhattan distance between two items is the sum of the difference of their corresponding components. [18]

2-6 Image noise

Image noise is random variation of brightness or color information in images, and is usually an aspect of electronic noise. it can be produced by the sensor and circuitry of a scanner or digital camera. image noise can also originate in film grain and in the unavoidable shot noise of an ideal photon detector .image noise is an undesirable by-product of image capture the obscures the desired information.

The original m meaning of "noise" was "unwanted signal "; unwanted electrical fluctuations in signals received by AM radios caused audible acoustic noise ("static"). By analogy , unwanted electrical fluctuations are also called "noise [1][2]".

Image noise can range from almost imperceptible specks on a digital photograph taken in good light, to optical and radio astronomical images that are almost entirely noise, from which a small amount of information can be derived by sophisticated processing. such a noise level would be unacceptable in a photograph since it would be connected to the sensor inject their own share of electronic circuit noise.[3].

Atypical model of image noise is Gaussian , additive , independent at each pixel, and independent of the signal intensity , caused primarily by Johnson- Nyquist noise (thermal noise), including that which comes from the reset noise of capacitors ("kTC noise ").[4] amplifier noise is a major part of the "read noise " of an image sensor , that is of the constant noise level in dark areas of the image[5] .in color cameras where more amplification is used in the blue color channel than in the green or red channel , there can be more noise in the blue channel.[6] at higher exposures , however , image sensor noise is dominated by shot noise , which is not Gaussian and independent of signal intensity.{form page WIKIPEDIA} [21].

2-7 Image retrieval

Retrieval is the process in which information in your memory can be recalled, information concerning events images and feelings are all stored in our memory. Just because you cannot remember something doesn't mean that it is not in your memory, just because you cannot remember something doesn't mean that it is not I your memory .it may be a problem with being able to locate it for retrieval. An image retrieval system is a computer system for browsing searching and retrieving images from a large database of digital images . Most traditional and common methods of image retrieval utilize some method of adding metadata such as captioning, keywords or description to the images so that retrieval can be performed iver the annotation words.[18] .

Chapter three

Practical work

- Introduction :
- Algorithms of the program :
- Explain the interface of program :
- Simplified explanation of user interfaces :

3-1 Introduction

In this chapter the practical work would be explained that contains the algorithms and the implementation of these algorithm through matlab programs .

In this algorithm, the program images are loaded and the statistical ladder is calculated for each image and stored in a database

in algorithm **3-2** configuration a database to gray

Algorithm **3-2**

- 1- start.
- 2- load the file that contain the images.
- **3-** divide the image into four parts.
- 4- calculate the histogram for each part.
- 5- storage the histogram.
- 6- go back to step 3.
- 7- the end

In this algorithm, the statistical ladder for each image is compared with the statistical scale to be searched in the database

in algorithm 3-3 explain searching and retrieval depends on converting the image to gray

Algorithm 3-3

- 1- start
- 2- image read.
- 3- divide the image to for parts.
- 4- Calculate histogram for each part.
- 5- choice destroyed part.
- 6- neglecting the distorted part and searching based on the undamaged part.
- 7- show image.
- 8- the end

In algorithm 3-4 configuration a database to color images the program images are loaded and the statistical ladder is calculated for each image and stored in a database

Algorithm 3-4

- 1- start.
- 2- load the file that contains the images.
- 3- analyze the image into chromatography(R,G,B).
- 4- calculate the histogram for each part of the image .
- 5- storage the histogram.
- 6- return to step 3.
- 7- the end

In this algorithm 3-5 searching and retrieval depends on color image the statistical ladder for each image is compared with the statistical scale to be searched in the database

Algorithm 3-5

- 1- start.
- 2- image read
- **3-** analysis image for three color
- 4- calculate histogram for each color
- 5- choice the damaged part
- 6- neglecting the distorted part and searching based on the undamaged part
- 7- the end

3-3 User interfaces

The main inter face of the program







load folder :open window to select the file contain image

brojecte3		
🔺 Select Folder to Open		×
🔶 🔶 🔺 📩 🔶 Ahmed Alshmare مشروع التخرج 🗧 Aa	مشروع التخرج Search 🗸 🗸	ο Aa 🛛 🔎
Organize 🔻 New folder		i≡ • (?)
This PC Name To Boots Destrop Documents Documents Downloads Music Pritures Videos Videos	Date modified T	Туре
Lead Disk (C) CD Disk (C) CD Disk (C) New Volume (f) New Volume (f New Volume (f Folden [Select Folder	Cancel

Select Image :open window to choice image

Open Image	▲ brojecte3	- 0
Organize New folder This C Name Determodified Type Downloads Noticems match your seach. Downloads Music Petures Load Folder Select Image Select Image Select Image Main program	Open Image X	Basec Image
Organize New folder Image Image Image <tr< td=""><td>A مشروع التخرج Aa Abmed Alshmare Alshmare Alshmare Abmed A</td><td>Dasec intage</td></tr<>	A مشروع التخرج Aa Abmed Alshmare Alshmare Alshmare Abmed A	Dasec intage
This PC Name Date modified Type Do Objects No items match your search. Downloads Downloads Music Pictures Local Folder New Volume (f New Volume (f	Organize 🕶 New folder 🔠 💌 🔟 👔	
 B O Objects Downleds Muicic Color (c) Color (c) Color (c) New Volume (f New Volume (f) File name I All Image Files ("reset":ntf:".nt,") Select Image Search Search Main program 	This PC Name Date modified Type	
Destrop Destrop Downlads Maix Pictures Udes Colored Disk (C) Open Cance Select Image Select Image Select Image Main program	3D Objects No items match your search.	
Bocuments Downloads Music Videos Videos Videos Co Drive (b) New Volume (f New Volume (f File name All Image Files ("uset" intf" int," int," Open Cancel Select Image Select Image Main program	Desktop	
• Downloads • Mais • Pictures • Cool Drive (D) • New Volume (f • New Volume	E Documents	
Muide In Values I values Values Values I	Downloads	
I vides	Music Dicturer	
Load Bik (C) CO Drive (b) New Volume (f New Volum	Videos	
Co Drive (D) New Volume (f New Volume (f Copen Cancel Select Image Select Image Main program	Local Disk (C)	
New Volume (f New Volume (f New Volume (f Cancel) Cancel) Cancel (Cancel) Cancel) Cancel (Cancel) Cancel) Cancel) Cancel (Cancel)	O D Drive (D:)	
New Volume (* Select Image Select Image Select Image Main program	New Volume (E	
File name All Image Files (".rset".ntt", ".nst. ~ File name All Image Files (".rset".ntt", ".nst. ~ Open Cancel Select Image Select Image Main program	New Volume (F	Load Folder
File name All mage files ("rast".nft", nft villes Open Cancel Select Image Search Main program		
Open Cancel Select Image Search Main program	File name: All Image Files (*.rset;*.ntf;*.nsf; >	
Search Main program	Open Cancel	Select Image
Search Main program		ould intrage
Search Main program		
Main program		
Main program		Search
Main program		
Main program		
		Main program

MENU — 🗆 🗙	Image hase be	een searched		Basec Image	
r the number of part wich have nois				Ū	
Have not any noise					
part one					
part two					
part three					
part fuor					
part one and part two					
part one and part three					
part one and part fuor					
part two and part three				Load Folder	
part two and part fuor					
part three and part fuor				Select Image	
rt one and part two and part three					
rt two and part three and part fuor					
rt three and part fuor and part one				Search	
art fuor and part one and part two				· · · · · · · · · · · · · · · · · · ·	
				Main program	

Search : open window to choice damage part

Main program: return back to main window



Chapter four

The conclusions and future vision

- Conclusion :
- Discussion :
- Feature work :

4-1 Discussion

The image retrieval can be done through analysis its content, However the image can be analyzed in a number of ways, including calculating the color intensity of the image and calculating the color distribution in the image and calculation of the statistical ladder of the image. In this research we relied on the calculation of the statistical ladder in the analysis of images and the composition of the databases on which the program is based in the search and matching and retrieval because the statistical method gives precise results in a short time and results of conformity to statistical peace is easy and simple and can be stored statistical ladder of images in a database and dealing with digital data as digital, Success rate this program is (%100), but in the case of the fact that the image you are looking for is present in the database or that the user has chosen the error of the parts distorted this program is unable to retrieve the pictures distorted until the proportion of distortion (%75) of the image to be retrieved on condition that the image you are looking for exit in the database and if were not exit, the program is retrieve closed to the image given in the search in terms of frequency colors in the picture This program was tested on a sample of (1000) images of sound images to create the database and a randomly distorted image was selected from a file containing distorted images in equal proportions (from 1% to 75%).the program returned the unmodified image depending on the undistorted part of the image given in the search.

4-2 Conclusion

Key interest of this project is to reduce the time and effort required to find the image of a given in the file contains the thousands of images and also solves the problem of suspecting that you get the images very competitive in the content chromatography (ie, images that have a difference among them a little is very to the degree no notice naked eye). In this project , we achieved good result , however the accuracy of retrieval is 100% for the simple image that were applied.

4-3 Feature work

Vision and future ideas for the project the vision of the future of this project is the improvement of algorithms to retrieve images depending on the body or shape in the picture taking into account the colors in the image to achieve the best result in the retrieval of the image in addition to organizing this program in the system to retrieve images quickly and costeffective.

Recourses

- 1- Gonzales, R. Woods, R. "Digital Images Processing". (Addison_Wrsley 2002).
- 2- Azriel, R. Avinash, "Digital Picture Processing". (Academic Press, 1982)
- 3- Rajeev, Ramanah. "Interpolation methods for the bayer coor array", (Masters thesis, North Carolina State University -2000).
- 4- Said A., Pearlman W.A. "New Fast and Effficient Image Codec Based on Set Partitioning in Hierarchical Trees", (Ieee Trans. On Circuits and Systems for Video Technology- June 1996)
- 5- Carson, C., Belongie, S., Greenspan, H., Malik, J., "Color- and Texture- Based Image Segmentation Using EM and Its application to Image Querying and Cassification ",(<u>http://elib.cs.berkeley.edu/papers/</u>-1997).
- 6- Smith, J.R. and Shih-Fu Chang. "Automated Image Retrieval Using Coor and Texture"(<u>http://www.ctr.columbia.edu/~jrsmith/html/pubs/-</u>1995).
- 7- Smith, J.R. and Shih-Fu Chang. "Automated Binary Texture Feature Sets for Image Retrival"(<u>http://www.ctr.columbia.edu/~jrsmith/html/pubs/-</u>1996).
- 8- Unser, M. Splines: Aperfect fit for signal and image processing. IEEE Signal Processing Magazine and IEEE Signal Processing Societys 2000 Magazine Award //16(6):22-38,November -1999.
- 9- Ardizzone, E., La CASCIA. M., Vito di Gesu, Valenti, C., "Content Based Indexing of Image and Video Databases by Global and Shape Features //http://www.cs.edu./associates/marco/publications.html -1996.

- 10- Sobel, I. "An isotropic image gradient operator. Machine Vision for Three-Dimensional Scenes" // pp.3376-379. Academic Press,1990.
- 11- Carson, C. and Ogle, V.E., "Storage and Retrieval of Feature Data for a Very Large Online Image Collection"//http://elib.cs.berkeley.edu/papers/ - 1996.
- 12- Mohammed Abdulshakoor Ameen "Content-Based Image Search Engine", The 17the national conference on Computer Science, King Abdul Aziz University,2004
- 13- Tou, J.T. and Gonzalez, R.C., Pattern Recognition Principles. Reading, MA: Addison- Weseley, 1974.
- 14- Bezdek, J. C., Pattern Recognition with Fuzzy Objective Function Algorithms . New York: Plenum, 1982.
- 15- James Ze Wang. Glo Wlederhold, Oscar Firschein and Sha Xin Wei, "Content-based image indexing and searching using Daubechies' wavelets", Department of Computer Science. Stanford University, Stanford. CA 94305, USA, US
- 16- Springer .Image Registration pp7-66| Cite as authors
- 17- AlleyDog.com . httos://www.alleydog.com/gloss
- 18- What is Image Processing : Tutorial with Introduction, Basics ,Types Articles ,http://www.enginerrsgarage.com
- 19- HTML ,dads ,http://xlinux.nistgov
- 20- The Physical Principals of Medical Imaging . Image Noise , Perry Sprawls , Ph.D.
- 21- Learn Image J sites.google.com Image Analysis .