Ministry of Higher Education Research

University of Technology

Computer Science Department

Software Branch

Image Similarity Search Engine

By:

SARAH MOHAMMED YOUSIF

RUSSULMOKDAM ABBAS

Supervised by:

Asst.Prof.Maisa'aAbid Ali Khoder.

2016

ACKNOWLEDGEMENT

I WOULD LIKE TO EXPRESS GRATITUDE TO MYSUPERVISOR ASSIST MS.MAESA'A ABD ALI FOR HER GUIDANCE , ENCOURAGEMENT , AND ADVICE THAT SHE GAVE ME THROUGHOUT THE WORK ON THIS .

T AM ESPECIALLYGRATEFUL TO MEMBERS FOR MY FAMILY (MY FATHER, MOTHER , BROTHERS , AND SISTER) FOR THEIR PATIENCE AND SACRIFICES.

SARAH & RUSSUL

ExaminationCommittee Certificate

We certify that we have read this project:

(*Image Similarity Search Engine*)

And as an examining committee, examined those students (sarah Mohamed Yousif, russul mokdam abbas) in its content and what is related to it and that in our opinion it meets the standard of a project for the degree of B.Sc. in computer science.

Signature:-
Name: -
Date: - 15 /6/2016

Signature:-
Name: -
Date: - 15 /6/2016

Supervisor's Certification

I hereby certify that project entitled:

(***Image Similarity Search Engine***)

Was prepared under my supervision at the department of computer science of the University of Technology, as a partial fulfillment of the requirements for the B.Sc. Degree in computer science.

15– 6 - 2016

Signature:-

Name: *Asst.Prof.Maisa'aAbid Ali Khoder.*

Date: 15 /6/2016

CONTENTS:-

Chapter one:-

Introduction:-

Chapter two:-

Image processing:-

Chapter three:-

Design and Implementation of System:-

Chapter four:-

Results and discussion:-

4.2- Results and discussion of Implemented System.

Chapter five:-

Conclusions and Recommendation:-

5.1 Conclusions.

5.2- Recommendation.

References.

Chapterone

Introduction

General Introduction

1.1 Introduction:-

Early 1920s - Bartlane cable picture transmission system - used to transmit newspaper images across the Atlantic. Images were coded, sent by telegraph, printed by a special telegraph printer took about three hours to send an image. First systems supported 5 gray levels. In 1964 NASA's Jet Propulsion Laboratory began working on computer algorithms to improve images of the moon. Images were transmitted by Ranger 7 probe. Corrections were desired for distortions inherent in on-board camera evolving technology and algorithms explosion of application areas.

1.2.Definitions:-

1.2.1 Computer Imaging:-

Computer imaging can be defined as the acquisition and processing of visual information by computer. The importance of computer imaging is derived from the fact that our primary sense is our visual sense, and the information that can be conveyed in image has been known throughout the centuries to be extraordinary.

One picture is worth a thousand words. This distinction allows us to separate the field of computer imaging into two primary categories: 1) computer vision and 2) image processing. In

computer vision applications the processed (output) images are for use by a computer, whereas in image processing applications the output images are for human consumption. The human visual system and the computer as a vision system have varying limitations and strengths, and computer imaging specialist needs to be aware of the functionality of these two very different systems.

These two categories are not totally distinct. The boundaries that separate the two are fuzzy, but this definition allows us to explore the difference between the two and to understand how they fit together see fig. (1.1).[2,5]

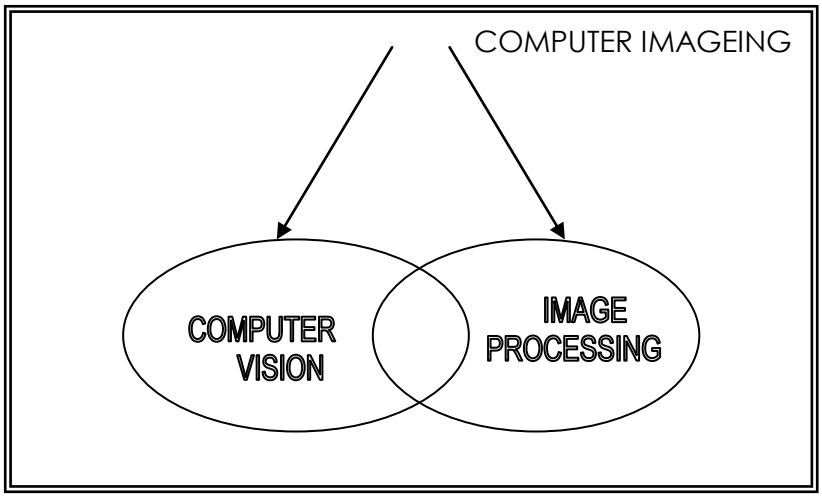


Fig. (1.1) Computer Imaging [1]

The field of image processing grew from electrical engineering as an extension of signal processing branch, whereas the computer science discipline was largely responsible for developments in computer vision. Recently, these two primary groups have come together to create the modern field of computer imaging. At many universities these two are still

separate and distinct, but thecommonalities and the perceived needs of each are rapidly bringing the two together.

1.2.2 Computer vision:

Computer vision is a computer imaging where the application does not involve a human being in the visual loop. The images are examined and acted upon by a computer. Although people are involved in the development of the system, the final application requires a computer to use the visual information directly. One of the major topics within the field of computer vision is image analysis [1,5]. Image analysis involves the examination of the image data to facilitate solving a vision problem. The image analysis process involves two other topics: feature extraction and pattern classification. Feature extraction is the process of acquiring higher-level image information, such as shape or color information, and pattern classification is the act of taking this higher-level information and identifying objects within the image. The field of computer vision may be best understood by considering different types of application. Many of these applications involve tasks that either are tedious for people to perform, require work in a hostile environment, require a high rate of processing, or require access and use of a large database of information. Computer vision systems are used in many various types of environments from manufacturing plants to hospital surgical suites to the surface of Mars[1,5] .

1. 3- Image Processing:

Image Processing is a computer imaging where the application involves a human being in the visual loop. The images are to be examined and acted by people.

For these types of application we require some understanding of how the human visual system operates. The major topics within the field of image processing include: image restoration, image enhancement, and image compression. The primary distinction between computer vision and image processing is that the output image is to be used a human being.

1.3.1- Image Restoration:

Image Restoration is process of taking an image with some known, or estimated, degradation, and restoring it to its original appearance. Image restoration is often used in the field of photography or publishing where an image was some show degraded but needs to be improved before it can be printed. For this type of application, we need to know something about the degradation process in order to develop a model for the distortion. When we have a model for the degradation process, we can apply the inverse process to the image to restore it to original form .[3,5]

1.3.2- Image Enhancement:

Image enhancement involves taking an image and improving it visually, typically by taking advantage of the human visual system's response. One of the simplest and often most dramatic enhancement techniques is to simply stretch the contrast of an image. Enhancement and restoration are similar in aim, to make an image look better, they differ in how they approach the problem. Restoration methods attempt to model the distortion to be image and reverse this degradation whereas enhancement methods use knowledge of human visual system's response to improve an image visually.[5,5]

1.4-Aims of Project:

1. To study the theory of image processing and computer vision.
2. To study the algorithms of similarity of Images.
3. To find a method for fast search engine in web.

1.5-Outline of Project:-

Addition of introduction the project consist to three chaptersChapter2:- This chapter describes Image processing.

Chapter3:- This chapter describes Design andImplementation System of Similarity Image

Chapter4:- This chapter describes Conclusion and Recommendation.

Chapter two

Image processing

Image processing

2.1- Introduction:

It 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 from 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 images as two dimensional signals while applying already set signal processing methods to them. It is among rapidly growing technologies today, with its applications in various aspects of a business. Image Processing forms core research area within engineering and computer science disciplines too.

Image processing basically includes the following three steps:

1-Importing the 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.

2.2- System Model:

Image analysis process can be broken down into three primary stages: 1) Preprocessing, 2) Data Reduction, and 3) Feature Analysis see fig. (2.1). First stage Preprocessing is used to remove noise and eliminate irrelevant, visually unnecessary information. Noise is unwanted information that can result from theimage acquisition process. [2,5]

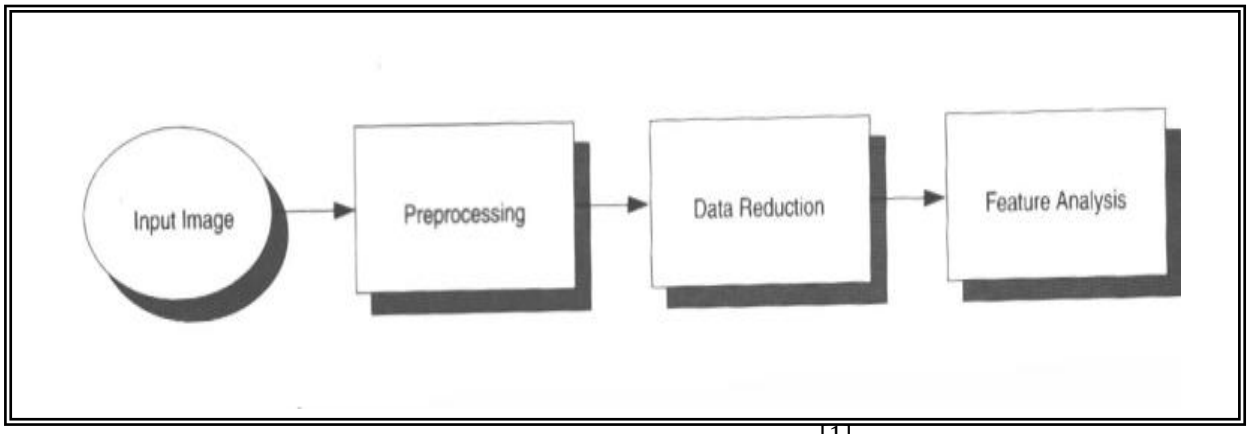


Fig. (2.1) Image Analysis [1]

Other Preprocessing steps might include gray-level or spatial quantization (reducing the number of bits per pixel or the image size) or finding regions of interest for further processing. The second stage, data reduction, involves either reducing the data in the spatial domain or transforming it into another domain called the frequency domain see fig. (2.2) and the extracting feature for

the analysis process. In the third stage feature analysis, the features extracted by the data reduction process are examined and evaluated for their use in the application [1,5].

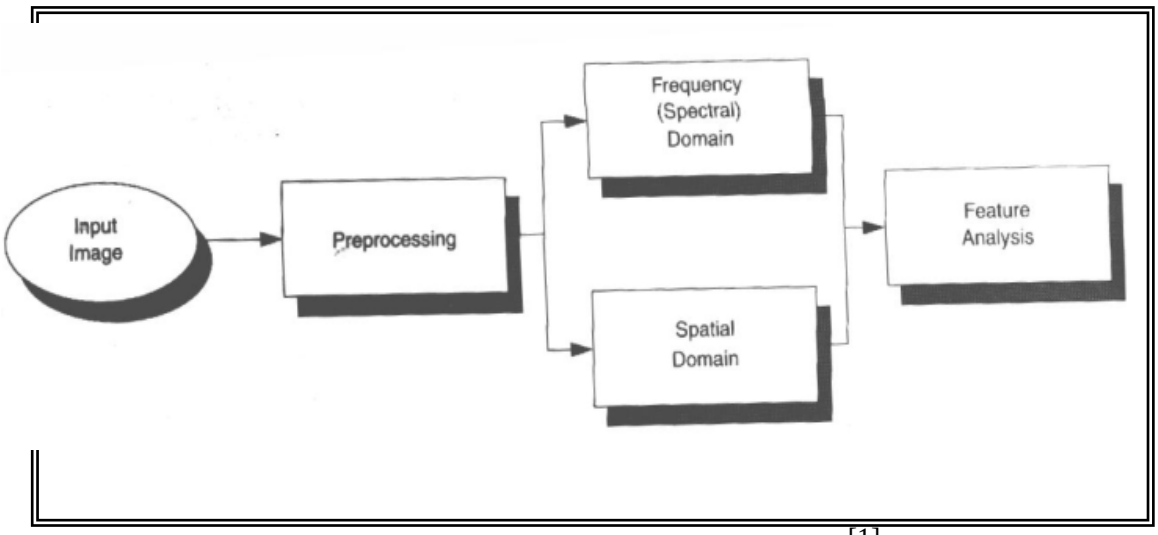


Fig. (2.2) Image Analysis Domain [1]

After preprocessing we can perform segmentation on the image in the spatial domain or convert it into the frequency domain via a mathematical transform. See fig. (2.3),which gives detailed diagram of this process. After either of these processes we may choose to filter the image. This filtering process further reduces the data and allows us to extract the features that we may require for analysis. After the analysis we have a feedback loop that provides for an application-specific review of the analysis results. This approach often leads to an iterative process that is not complete until satisfactory results are achieved. The application feedback loop is a key aspect of the entire process [1,5].

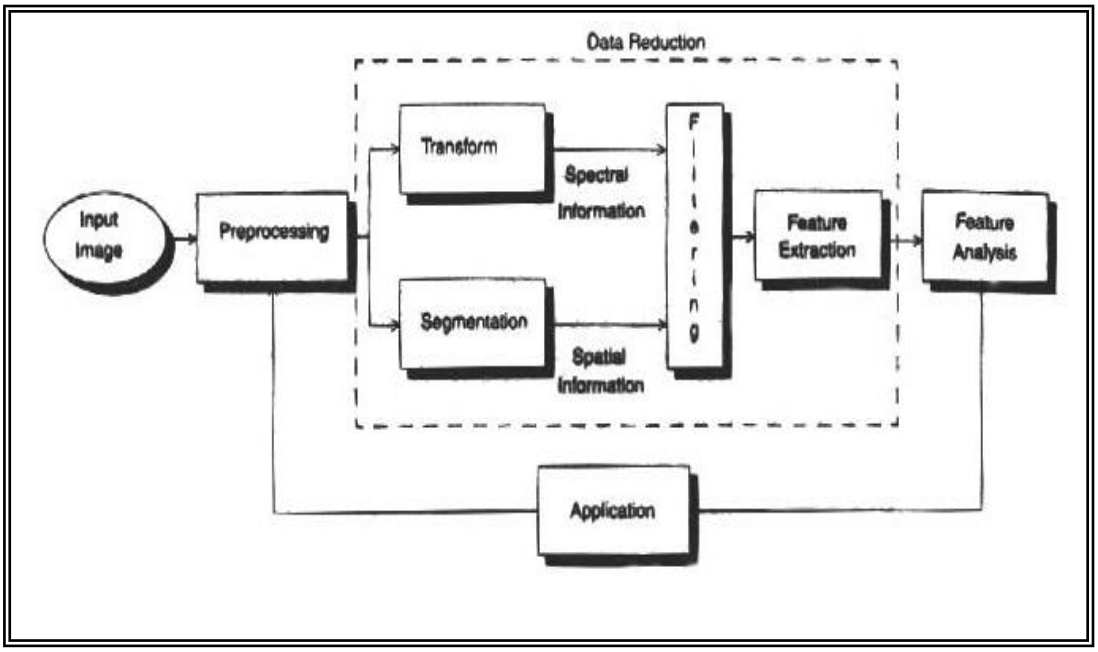


Fig.(2.3) Detailed Diagram of segmentation process.

2.3- Type of Images:

The digital image *I(r,c)* is analysis as a two-dimensional array of data, each pixel value corresponds to the brightness of the image at the point *(r,c).* A two-dimensional array like our image at our image model *I(r.c)* is referred to as a matrix, and one (row or column) is called a vector.

These are three types of images: 1) binary image, 2) gray-level image, 3) color image[3,5].

2.3.1- Binary Images:

Binary images are the simplest type of image analysis and can take on two values black and white or ‘0’ or ‘1’ see fig.

(2.4)[3,5]. A binary image is referred to as a (1-bit/pixel) image because it takes only one binary digital to represent each pixel[4,5].

FIG. (2.4) BINARY IMAGE [30]

2.3.2- Gray-Level Images:

Gray-level images are referred to as monochrome or one color images. They contain brightness information only, no color information see fig. (2.5) [2]. Typically gray-level image analysis contains 8-bit/pixel data which allow us to have (0-----225) different shades.

FIG. (2.5) GRAY-LEVEL IMAGE [2,5]

2.3.3- Color Images:

Color images can be modeled as three-band monochrome image data where each band of data corresponds Red, Green, Blue or *RGB* images see fig. (2.6) . When we use 8-bits monochrome standard as a model, the corresponding color image would have 24-bits/pixel 8-bits for each of the three-color bands (red, green, blue). In 8-bits colored image analysis the palette represents 256(0----225) colors variation .

a) 24-bit b) 8-bit

FIG.(2.6) COLOR IMAGES

2.4- Image File Format

These are several different types of graphics file format, these formats are (Bit map, Vector, Animation) and other [1,5].

As a general rule scanned images are bitmap files while drawing made in applications like Corel-Draw are saved as vector graphics. Vector are graphics that are completely

described using mathematical definition--Bitmap images are collection of bits that form individual dots (pixels) that all have their own color [4,5].

Bitmap file consists of header, bitmap data and other information, which may include a color palette and other data. The basic components of a simple bitmap file are the following [1,5].

Header
Bitmap Data

An image with palette appears immediately after the Header:

Header
Palette
Bitmap Data

Bitmap data can be saved in a wide variety of formats, among these are:

1-BMP file format.

2-GIF file format

3-JPEG file format

2.4.1- BMP File Format:

The basic design of BMP file format makes it a good general purpose format that can be used for color or black and white image storage. BMP is organized into four sections as illustrated in fig. (2.7) [3,5].

File Header
Information Header
Palette
Bitmap Data

FIG. (2.7) BMP FILE STRUCTURE

2.5-IMAGE SIMILARITY AND SEARCH:

When the information from images is captured in a feature set, there are two possibilities for endowing them with meaning: one derives an unilateral interpretation from the feature set or one compares the feature set with the elements in a given data set on the basis of a similarity function.

2.6 Sobel Filter:

The sobel operator is a derivate mask and is used for edge detection, the sobel operator is also used to detect two kinds of edges in an image:

1- Vertical direction.

2- Horizontal direction.

sobel operator the coefficients of masks are not fixed and they can be adjusted according to our requirement unless they do not violate any property of derivative masks, as shown in figure (2.8), and figure (2.9).

-1	**0**	**1**
-2	**0**	**2**
-1	**0**	**1**

Figure (2.8) Sobel filter vertical direction.

-1	**-2**	**-1**
0	**0**	**0**
1	**2**	**1**

Figure (2.9) Sobel filter horizontal direction.

- Compute sobel filter with image, the image performing theconvolution of Sobel kernels with the image

Chapter three

Design and Implementation of System

Design and Implementation of System

3.1- Introduction:

In this chapter, the design of similarity algorithm is explained. The first step of this algorithm contains loading the images from a database. Next, the detection for the image is done usingsoble filter. Finally, the similar image for the detected one is searched and loaded from the database.

3.2- Design System of Database:

Design the table1 to input image original as show in figure(3-1)

The table contains five pictures:-

The first picture name is (beautiful view),size (79),open picture(package).

The second picture name is (boat),size(81),openpicture(package).

The third picture name is (fire),size (82),open picture(package).

The fourth picture name is (forest),size (81),open picture(package).

The fifth picture name is (water), size (79),open picture(package).

Table1

name pic	size	open pic
beautiful view	79	Package
boat	81	Package
fire	82	Package
forest	81	Package
water	79	Package

Figure(3-1)The database of the program.

3.3- Design similarity Algorithm:

In this function we decides that every image that are be processed by edge detection must be identify by unique number added at the and of image's name like "fire.jpg" to be "fire1.jpg" after that that to retrieve the original image the function cut the unique number be multi- programming steps like "fire1.jpg" to "fire.jpg"

The Algorithm of similarity is explained below:-

Input: image, image detection

Output: original image similarity process

Step1: b = Split(f, "\")

Step2: a = Split(b(3), ".")

Step3:a(0) = Left(a(0), Len(a(0)) - 1)

Step4:aa = a(0) + "." + a(1)

Step5:pp = b(0) + "\" + b(1) + "\" + b(2) + "\" + aa

End sub

Using edge detection of image withSoble filter on the imageas shown in figures (4-1) and (4-2) inchapter four. The **Algorithm of edge detection is explained below:-**

Input: load image,soble.filter

Output: edge detection image

Process: x=load picture1.width

Y=load picture1.height

Step1:For i = 1 To x – 1

Step2:For j = 1 To y – 1

Step3:a1=picture1(i,j)

Step4: a2=picture1(i,j)

Step5:P = Sqr((a1 * a1) + (a2 * a2))

Step6: If (r < 0) Then r = 0

Step7:If (r > 255) Then r = 255

Step8:If (g < 0) Then g = 0

step9: If (g > 255) Then g = 255

step10: If (b < 0) Then b = 0

Step11:If (b > 255) Then b = 255

Step12 :Nexti

step13: Next j

End Sub

Chapter Four

Results and Discussion

Results and Discussion

4.1- Introduction:

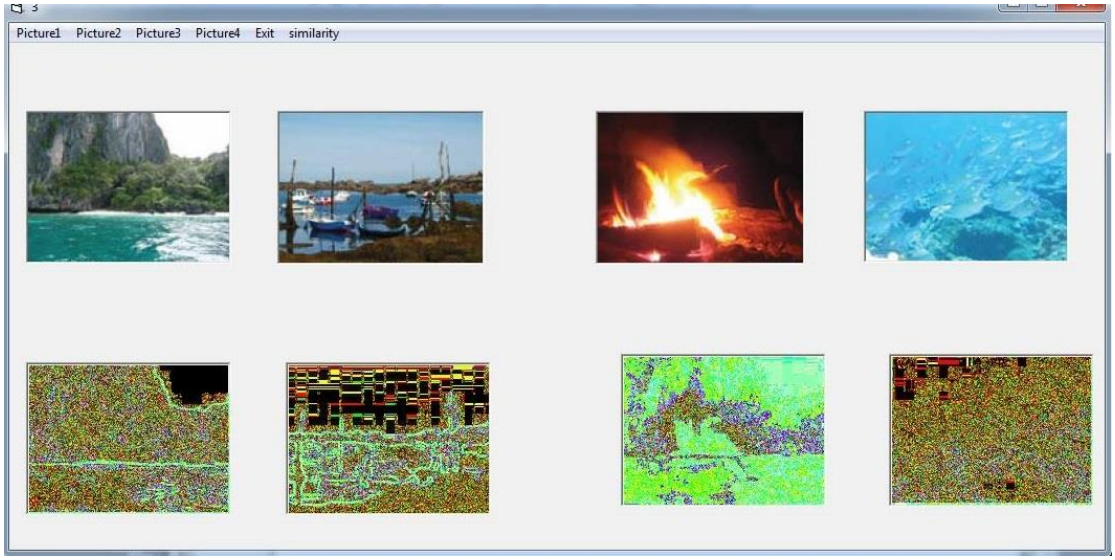
In this chapter, the results of the designed algorithm are displayed. Also, discussion for the results will be given.

4.2- Resultsand discussion of Implemented System:

This figure to display the images that it have downloaded from database ,shown it in below.

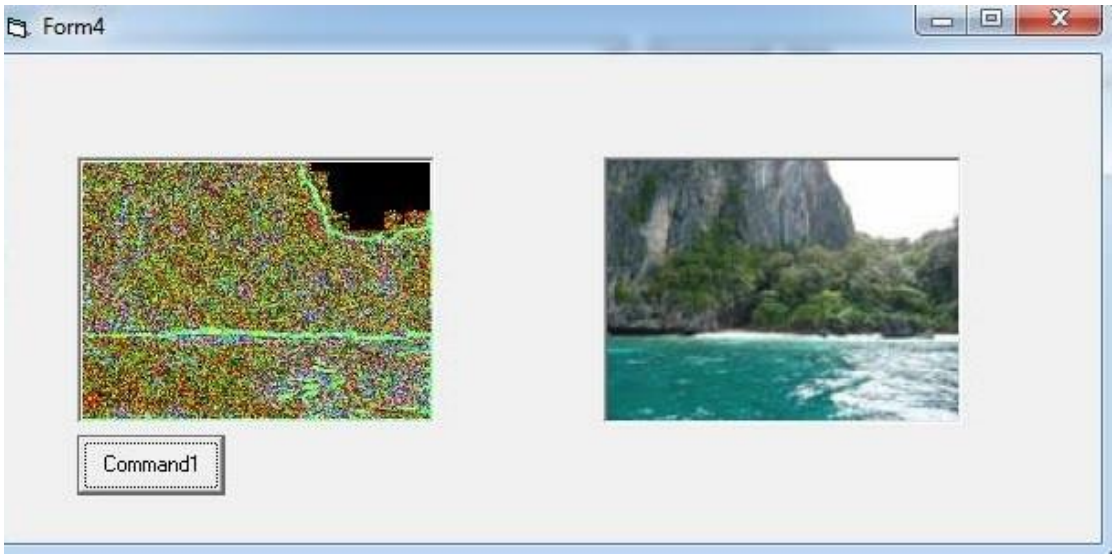
Figure(4-1)display the image.

- This figure is to make a detection soble filter of images and stored in the folder to be displayed, shown it in below.



Figure(4-2)To make detection of images

- This figure is holding similarity between the images that have been holding process them and store and is dispatched with the original images, show it in below .



Figure(4-3) similarity images

Chapter Five

Conclusions and Recommendation

Conclusions and Recommendation

5.1- Conclusions:

1-It can be conclusion of this project, that it is easy to get the original image using the soble filter.

2-And similarity by dividing the image into three parts, to get the original image.

3- It is a very fast and efficient way to search for the original image after it is stored in the original image data and image storage Sobel filter base exists in the calculator file.

5.2- Recommendation:

As suggestions for future work:

1- Developing this system to beimplementation tomodify this algorithm to search for 3D images.

2- Developingthis system thatimplementation to audio speech.

3- Developing this system thatimplementation tovideo images.

References

1- Scott E. Umbaugh, "Computer Vision and Image Processing", Prentice Hall, Inc 1998.

2-Maisa'aAbid Ali K., "Segmentation and Clustering For Image Color", University of Technology, 2004.

3- Th. Gevers , and A.W.M. Smeulders , " Image Search Engines An Overview", University of Amsterdam, 2003.

4- Stein Keijzers, Peter Maandag, Elena Marchiori, Ida Sprinkhuizen-Kuyper, "Image Similarity Search using a Negative Selection Algorithm", ECAL, 2013.

5- Nadav Ben-Haim, Boris Babenko, and Serge Belongie,"ImprovingWeb-based Image Search via Content Based Clustering".
Available at:

http://developer.yahoo.net/search/image/V1/imageSearch.html