Faculty of Engineering

LAB SHEET

MULTIMEDIA TECHNOLOGY AND APPLICATIONS

ECE3086

Lab 1) MTA1 – Image Processing for Image Retrieval

Lab 1 Marking Scheme

Refer to rubric sheet. Total marks = 12 . Lab 1 contribute 5% to the course work mark.

Experiment 1: Investigation on the Semantic Gap Problem in the Content Based Image Retrieval System

Objectives of the Experiment

1) To implement basic image processing methods using Python
2) To perform image feature extraction and compare the feature effectiveness in image retrieval task

1.0 Brief Note

Please go through the following Python tutorial given here before coming to the lab. You are expected to solve most of the image processing problem before your lab session starts and show the result to the lab instructor during the lab session. Marks will be given based on the question & answer session (viva) with the instructor and lab report. Lab report is to be submitted within one week after the experiment date. You are required to use the lab report cover format and the content format given at Google classroom website. During the lab session you can get advice and feedback on the solution that you have implemented.

Test Images and Sample Code

All the test images and sample Python code can be downloaded from the Google classroom lab folder.

2.0 Introduction

In this experiment you will investigate on some technical issues related to the content-based image retrieval (CBIR) system. The CBIR system allows user to create searchable image database. The CBIR system will retrieve a set of images from the database that are visually similar to the given query. Before the images can be searched, it has to be indexed with a identifier. The simple approach is to use the keyword as an identifier but this requires manual human annotation in order to label all the images in the database with a suitable keyword. Such manual labelling is subjective and not consistent. A better approach is to automatically extract feature vector from the image and use the feature vector as a searchable index used by the image search engine.

Feature vector describes the attributes of the object. For example, human can be described by the feature vector that consists of the measurement obtained from its height and weight variable. The feature vector v = [height weight] is a vector value.

The feature vector encodes the measurements obtained from the various attributes of the image and thus serves as a compact image descriptor. Feature vector is used to categorize objects or images into groups. This process is called classification. The feature vector can be automatically extracted from images and compared for visual similarity. This enables images search by example image and is known as query by example (QBE) instead of the conventional query by keyword.

The CBIR system that uses the QBE approach involve two main process that are

a) The database preparation
b) Image searching

In the database preparation process, every image in the database will need to be indexed or represented by a compact image descriptor (feature vector). The annotation is done automatically whereby the CBIR system extracts and stores the feature vector from the database image. In the searching process, the CBIR search engine needs to extract the feature vector from the query image and compare this feature vector with the feature vectors of the images in the database.

Given a query image example, a content-based image retrieval (CBIR) system will retrieve a set of images similar to the query image. Image that belongs to the same category as the query image is considered similar and will be retrieved. In order to retrieve similar images, the retrieval algorithm will need to compare the dissimilarity between the query image and all the images in the database.

A dissimilarity distance metric is a function that takes in two feature vector of the same dimension and gives a numerical value that shows the degree of dissimilarity between the two vectors. One of the most common dissimilarity distance metric function is the Euclidian distance.

Given the two feature vectors $H_1 = [h_1(1), h_1(2), \dots, h_1(M)]$, $H_2 = [h_2(1), h_2(2), \dots, h_2(M)]$ H_1 represents image I_1 and H_2 represents image I_2

The Euclidian distance between the feature vector H_1 and H_2 is denoted as $D_E(H_1 , H_2)$

$$D_E(H_1, H_2) = \sqrt{\sum_{i=1}^{i=M} (h_1(i) - h_2(i))^2}$$

The larger Euclidian distance between the image descriptor H_1 and H_2 means the more dissimilar the image I_1 when compared to I_2.

Figure 1 shows the example of the query image and the set of similar images retrieved by the CBIR system. Figure 2 shows the internal block diagram of the CBIR system (query by example). The feature vector plays an important role in getting accurate retrieval result. In this experiment you will investigate the effectiveness of the colour histogram as the feature vector for image retrieval. A good feature vector will give small Euclidian distance value for the pair of images that belong to the same category. The Euclidian distance for the feature vector that belongs to different category will be large.

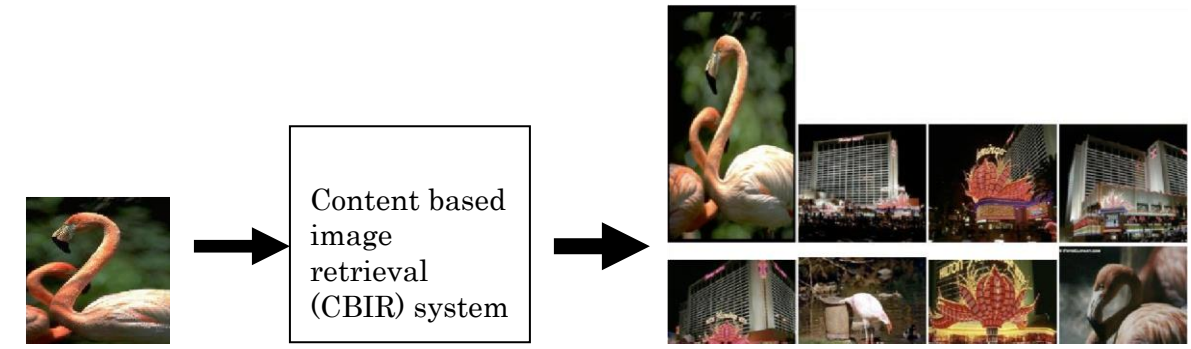

Figure 1: Content based image retrieval system (query by example)

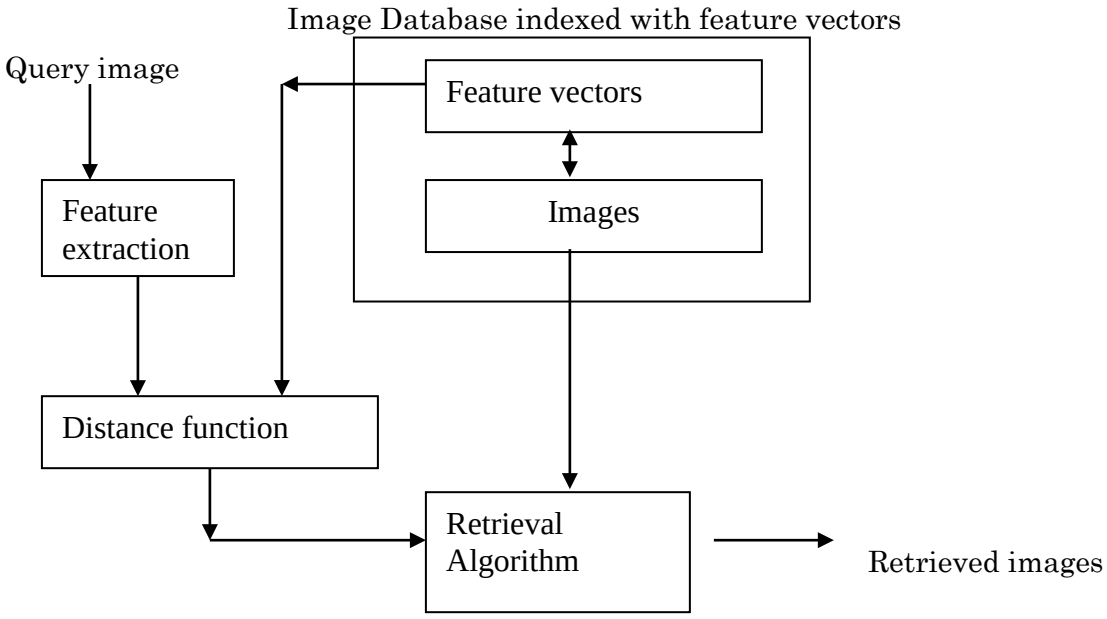

Figure 2: CBIR (query by example) Block Diagram

2. The Semantic Gap Problem in the CBIR System

Human search for images using high-level concept. The high-level concepts example includes scene types and expressed as semantic labels such as forest, beach, night sky, high rise building and ocean. It is common for human to use semantic labels to describe these concepts associated with the images. Such label allows the classification or categorization of the images into semantic groups. However, computers cannot understand semantic information in the image. The image categories are commonly approximately represented by low level features such as colour, shape and texture. The features are grouped as a vector and are referred to as the feature vector. The feature vector provides a compact numerical value of the image attributes. The major problem is that it is difficult to obtain the unique correspondence between the feature vector and the image label. Different semantic labels can have closely similar value for the feature vector. Semantic gap is the problem due to the difficulty in finding a good feature vector that can be used to categorize images into different semantic group. The semantic gap problem is commonly faced in the content-based image retrieval (CBIR) system. As a result, a human query to the system for a particular class of images, for example soccer players may result in images from unrelated categories (e.g forest) to be retrieved.

So how do images of different categories be differentiated? CBIR system uses distance function to give the dissimilarity distance between the feature vectors. One common dissimilarity distance is the Euclidian distance. The higher Euclidian distance value between feature vectors implies that image is less similar
(more dissimilar). In the query by example system, the retrieval algorithm typically computes the Euclidian distance between the query image and all the images in the database. Given a retrieval session to request K images from the database, the set of images with the K smallest distance are retrieved.

3. Problem Definition

One of the major problems in content-based image retrieval is the semantic gap. The research question that you will tackle is as follow; Can the low-level features using colour histogram be used to efficiently represent the image category (high level concept)? In this experiment you will evaluate the following image feature vectors

a) Image feature from pre-trained convolutional neural network (CNN).
b) Colour histogram that uses the red, green and blue (RGB) channels

Perform experiment using the provided Corel-1K dataset and investigate if the colour histogram and feature from pre-trained CNN meets the criteria of a good feature vector. Use the experiment results analysis to make your conclusion.

4. Dataset

The Corel-1K dataset that is used in this experiment consist of ten categories. Each of the images has been manually annotated to indicate its class type. It is pre-labelled so that you can evaluate the effectiveness of the feature vector and the CBIR system. For example, the images in the file 100.jpg until 199.jpg belong to the beach category. The beach category is given the label 2.

Image Database file: CBIR_database.pickle

Corel-1K CBIR dataset

Image Category	Category Label	Image file name in database
Africa	1	0000.jpg ,001.jpg, … .099.jpg
Beach	2	100.jpg, 111.jpg ……. 199.jpg
Building	3	200.jpg, 211.jpg ……..299.jpg
Bus	4	300.jpg, 311.jpg ……. 399.jpg
Dinosaur	5	400.jpg, 411.jpg ……. 499.jpg
Elephant	6	500.jpg, 511.jpg ……. 599.jpg
Flower	7	600.jpg, 611.jpg ……. 699.jpg
House	8	700.jpg, 711.jpg ……. 799.jpg
Mountain	9	800.jpg, 811.jpg ……. 899.jpg
Food	10	900.jpg, 911.jpg ……. 999.jpg

Table 1: Distribution of Images in Ten Categories

5. Methods

1. Load the given files from Google classroom into the selected folder
2. Use the Corel-1K dataset, which has been manually annotated. The dataset contains 1000 images with 10 classes. The prepared database has been saved in the file CBIR_database.pickle . Go to the working directory and load the database into Python workspace. Follow the given codes as your reference. The database is supposed to contain the feature vector for all images based on the pre-trained convolutional neural network (CNN) and the histogram of the RGB channels.

3. Use 256 bins for each colour channel to construct the colour histogram as the image descriptor. Use the extracted features from CNN in the database file. Compare the effectiveness of these two types of features.

4. Select 10 images from each category as the query images. For each query image, find the 10 most similar images with the use of Euclidian distance and the colour histogram feature. Repeat the experiment with CNN feature. Compute the average precision rate for all the retrieval attempts. Tabulate the result and discuss your analysis. Show the Python programming code to complete the task.

   Precision rate PR = CI/NR (NR=10)
   CI: number of relevant images retrieved for one query
   NR: number of images retrieved for one query

5. Tabulate the result in tables and discuss the results obtained. Show a few retrieval results in your report. Use the analysis of the result to justify your conclusion.

6.0 Report

Experiment is conducted with the purpose of evaluating a solution to a problem. It is also used to find information in a systematic manner. Your report should answer the following questions and the questions in the given Jupyter notebooks.

1) What is the purpose of the experiment?
2) What have you learned from the experiment?
3) Which type of feature is more suitable for image retrieval
4) What is the conclusion? Show evidence to support your conclusion.