IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

Visual Based Image Search using Perceptual Hash Codes for Online Shopping

Shradha kadam[1], Priyanka Gupta[2], Charul Veer[3], Akshay Loke[4]

B.E Student, Information Technology, Vidyalankar Institute of Technology, Mumbai, India [1, 2, 3]

Professor, Information Technology, Vidyalankar Institute of Technology, Mumbai, India [4]

Abstract: Visual based search aims at identifying a product, or retrieving similar products from a database based on a photo captured from a mobile phone camera or browsed from computer. Today, Keywords are used to find products on online shopping site. The fast and easy search has motivated the development of Visual based search which is the type of search that puts images at the centre of user's attention. Visual basedimage search works like traditional (key-word) search but without having to type any text or go through complicated menus to initiate the search process. We focus on a visual shopping system. To be more specific, we want to use visual information to find corresponding products in a precise and easy fashion. This is useful when people do not know how to describe the visual object in terms of text. In this paper we discuss the new hash technique called as perceptual hash which calculates the hash code for input image provided by user and then hamming distance is calculated with the images stored in the databases having their respective perceptual hash value and based on highest percentage of visual similarity, products are displayed to user.

Keywords: visual based image search, perceptual hash codes, hamming distance, visual similarity.

I. INTRODUCTION

Online Shopping system is enhancing day by day in terms of approximate search results based on the queries given by the customers. Traditional way of searching involves searching by keywords from the query given by user. Sometimes users are unable to form query as required by system which may not result in finding product which they are looking for.

Visual based search aims at filtering products through images that user wishes to buy and process query fast as possible and display similar types of images. In this work, we propose a user-friendly way which can provide user with superior performance and integrated user experience. User will be provided with multiple ways to select query image, directly taking a snapshot, choosing image from existing albums and downloading images from internet.
In this work we use perceptual hash codetechnique with hamming distance to refine query results based on the visual similarity.

II. PERCEPTUAL HASHING

Perceptual hash algorithm describes a class of comparable hash functions. Features in the image are used to generate a distinct (but not unique) fingerprint, and these fingerprints are comparable.
Perceptual hashes are a different as compared to cryptographic hash functions like MD5 and SHA1 [1]. Cryptographic hashes generate random hash value. The data used to generate the hash acts like a random seed, so the same data will generate the same result, but different data will create different results. Comparing to SHA1 hash, if these hashes are different, then the data is different and if the hashes are the same, then the data is likely the same. In this, there is a possibility of a hash collision, having the same hash values does not guarantee the same data. In contrast, perceptual hashes can be compared -- giving you a sense of similarity between the two data sets.

A. How to generate perceptual hash?

In this work we use one of the simplesthashes represent a basic average based on low frequencies [3].

Steps to create perceptual hash of an image:

1. Take any image.
2. **Reduce size**: The fastest way to remove high frequencies and detail is to shrink the image. In this case, shrink it to 8x8 so that there are 64 total pixels. This way, the hash will match any variation of the image, regardless of scale or aspect ratio.
3. **Reduce colour:** The tiny 8x8 picture is converted to a grayscale. This changes the hash from 64 pixels (64 red, 64 green, and 64 blue) to 64 total colours.
4. **Average the colours**: Compute mean value of the 64 colours.
5. **Compute the bits**: Each bit is simply set based on whether the colour value is above or below the mean.
6. **Construct the hash**: This will be 64 bit binary string or can be represented as hexadecimal value.

B. Hamming Distance

With the perceptual hash codes, image similaritycan be efficiently measured using logical XOR operations in Hamming space by Hamming distance, an integer value obtained by counting the number of bits at which the binary values are different.

C. Working of query adaptive image search

User uploads query image to the server, image is resized to the size of image stored in database. It is important to resize as it may affect the hash code that will be generated. Perceptual hash of resized input image is calculated. The 64 bit string is compared with each record, to bit string

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 3, March 2016

stored in database (hash value) .Hamming distance is calculated for each comparison and image is retrieved based on minimum hamming distance. A threshold of minimum hamming distance to filter the results is set and based on the fulfilment of the criteria productsare displayed. Fig. 1 depicts the working of query adaptive image search

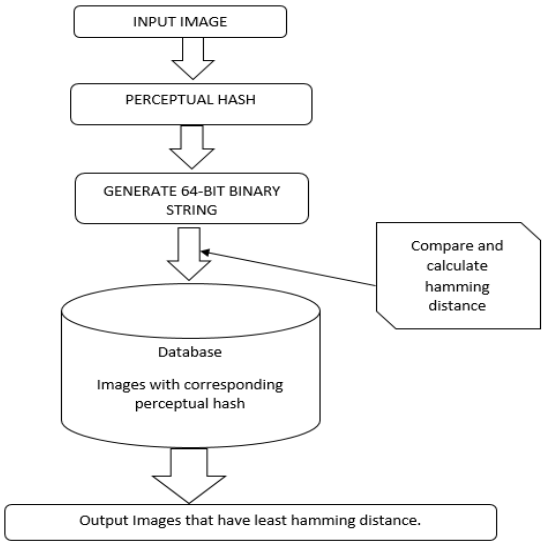

Fig 1 Working of query adaptive image search.

D. Pseudo-code:

1. Get input image from user.
2. Resize input image to size of image stored in database.
3. Hash1= perceptual hash of resized image.
4. while(all records from database are scanned)
5. {
6. Hash2=perceptual hash of each image in database.
7. Similarity=Compare Hash1 and Hash2 and calculate the number of bits they differ in.(hamming distance)
8. Represent hamming distance in terms of percentage.

Percentage= (Similarity/64 (string length of hash code))*100

9. If(Percentage>97)
10. Display Image.(this image is visually similar)
11. }

E. Experimental Results

Fig. 2 Input Image

To analyse the working of perceptual hash, we tested 1000 images consisting of different categories ex. Men and Women shirts, jeans, backpacks, Hand bags, wallets, watches, shoes. Images of these products are stored in the database with its corresponding perceptual hash value. Fig. 2 shows the input image

Fig. 3 displays visually similar images based on minimal hamming distance.

Fig. 3 Output of visually similar images

III.CONCLUSION AND FUTURE SCOPE

In this paper we have proposed a method to find visually similar images using perceptual hash technique. The experimental results shows effectiveness of the algorithm. Perceptual hash has varied implementation that can be useful for applications like content based image authentication and for image retrieval and matching in large scale image databases.

Future scope could be classification of objects in the database itself by assigning class weights to visually similar objects. Now, when user uploads image rather than calculating hamming distance for all records in database and then filtering results it would compare class weights and then get more approximate search result.

ACKNOWLEDGMENT

We thank our teachers for their valuable suggestions and their contribution that has helped us to accomplish our research work.

REFERENCES

[1] A. Menezes, V. Oorschot, and S. Vanstone, Handbook of Applied Cryptography, CRC Press, 1998.
[2] J. Wang, S. Kumar, and S.-F. Chang, "Semi-supervised hashing for scalable image retrieval," in IEEE Conf. on Computer Vision and PatternRecognition, 2010.
[3] www.bertolami.com