Content Based Image Retrieval Using Signature Representation

Chathurani N.W.U.D., Geva S., Chandran V., Chappell T.
School of Electrical Engineering and Computer Science
Queensland University of Technology
Brisbane, Australia
(n.chathu@qut.edu.au, s.geva@qut.edu.au, v.chandran@qut.edu.au,
t.chappell@connect.qut.edu.au

Abstract

Retrieving relevant images from a large, diversified collection using visual queries (image content) as search argument is a challenging and important open problem. It requires an efficient and effective content-based image retrieval (CBIR) system. Image representation has a profound effect on the performance of CBIR. This paper presents a CBIR system based on a novel image representation using a new approach to the generation of image signatures (CBIR-ISIG). Image signatures are generated by applying random indexing (RI) to a Bag-of-visual Words (BoW) representation of the images. RI is an efficient and scalable approach to dimensionality reduction, based on random projection which avoids the computational cost of matrix factorization. Most importantly, it can be performed incrementally as new data arrives, as is crucial for online systems. The retrieval quality of the proposed approach is evaluated using a benchmark dataset for image classification (a subset of the Corel dataset). The proposed approach shows promising results with comparable retrieval quality to state of the art approaches while retaining the benefits of the highly efficient representational scheme.

Keywords: Content based image retrieval, Image signature, Random Indexing, Invariance, Bag-of-Words.

1 Introduction

The development of the Internet and increased availability of image capturing devices have enabled collections of digital images to grow at a fast pace in recent years and to become more diverse. This created an ever growing need for efficient and effective image browsing, searching, and retrieval tools. Retrieving relevant images accurately to satisfy an information need from a large, diversified collection using visual queries is a challenging and important problem to address. Despite many years of research in this area, an effective general solution still eludes researchers. The most familiar CBIR system in wide use today is offered by Google Images, where a user can upload an image as a query, and the system responds with similar images, based on content. The performance of this system can only be described as less than satisfactory for all but a very small fraction of images – when it works well, it is almost invariably for very famous images, but even then it often fails. Figure 1 shows an example of a failed search for a query image (Eiffel tower) in Google Images search. All the while Google Images are quite effective for text based search using accompanied image text annotations, it is not effective for (visual) content-based image retrieval.

Selecting an appropriate image representation is the most important factor in implementation of an effective CBIR system. Among the many image representations that have been studied, the BoW approach is one of the most promising (Yuan, et al. 2011 and Mansoori, et al. 2013). The BoW approach is flexible with respect to geometry, deformations and viewpoint and it provides vector representation for sets. Finally it gives a compact summary of image content. This research adapts the BoW approach and presents an efficient content based image retrieval system using image signatures (CBIR-ISIG) and, introduces a new approach to image signature definition. Image signatures are generated by applying random indexing on a BoW representation of the image, using Topsig (Geva and De Vries 2011). This improves the computational efficiency of conventional BoW CBIR approaches. The system performance is evaluated by the use of a well categorized subset of the standard Corel dataset and is shown to produce superior results when compared with other independently developed and tested approaches.

There are numerous examples in the literature where signatures are used to detect near-duplicates in image processing and text processing at high speed. Topsig is a tool that is used to generate and search signatures, with response time at the millisecond scale to search millions of documents (Chappell, et al. 2013). Signature representation is a concise representation of real vectors, which help to reduce the storage space (eg: in this research around 5840 bytes of a real valued image feature vectors are represented by 1024 bytes). It can be defined as a representation of documents, images and other searchable abstract objects as binary strings of fixed length (Chappell, et al. 2013). The motivation behind this research is to find a representation that is efficient to search while retaining retrieval quality by feature fusion and signatures.

In image retrieval there are essentially two approaches. The first is the text based (TBIR) approach. Here the user describes the image content with a textual specification of the image content, such as tags. However the representation of images using text was found to be too difficult: requiring excessive human effort to support, time consuming and expensive, incapable of describing

Copyright © 2014, Australian Computer Society, Inc. This paper appeared at Australasian Data Mining Conference (AusDM 2014), Brisbane, 27-28 November 2014. Conferences in Research and Practice in Information Technology, Vol. 158. Richi Nayak, Xue Li, Lin Liu, Kok-Leong Ong, Yanchang Zhao, Paul Kennedy Eds. Reproduction for academic, not-for-profit purposes permitted provided this text is included.

Figure 1: Retrieval results for an example query image (Eiffel tower) in Google image search

rich image features, and very much dependent on human perception. To overcome the limitations of TBIR, content based image retrieval (CBIR) was introduced, allowing queries to be specified visually. In CBIR image features are extracted and indexed automatically to support storage and retrieval of images with visual queries.

Over the last 30 years significant attention has been paid to CBIR. Extensive research has been conducted (Liua, et al. 2007) to develop sophisticated algorithms to extract low-level image features such as colour, shape, texture, edges, point of interest, and spatial relationships. These algorithms then measure the similarity between pairs of images based on image feature vectors. Much work had been dedicated to exploring and provision of solutions to the problems of image rotation, translation and scale invariance (RTSI). However, these algorithms cannot adequately model image semantics and have many limitations when it comes to dealing with broad content image databases, especially with regards to response time and retrieval accuracy. CBIR has been assessed comprehensively in (Liua, et al. 2007).

In this paper authors offer a different approach to the construction of image signatures. This starts from a set of standard image features, use vector quantisation to generate a BoW representation. Sub-image feature sets as well as full image feature set are further used. Then random indexing is applying to fuse the various feature sets and then the representation used to store and search images. This approach is described in more detail later in this paper.

The rest of the paper is organized as follows. Section 2 outlines the previous work done on CBIR. Section 3 describes an overview of the system and the proposed method. Section 4 provides details of the experiment and the results obtained. Conclusions drawn from the research findings are included in section 5.

2 Previous Work

CBIR systems measure visual similarity between a query image and multiple database images and then retrieve the top ranked images using various similarity measures. These similarity measures are computed on extracted features (Colour, Texture, and Shape) from images. In (Liua, et al. 2007) a comprehensive assessment of prior work with CBIR was performed.

Most of the systems (Saad, et al. 2011) have used global feature representation whereas some other systems (Hiremath and Pujari 2007, Mansoori, et al. 2013, Takala et al. 2005) have used local feature representation. In global representation, features are extracted from the whole image while in local representation features are extracted from either, segmented regions, or from a regular grid or points of interest. Local representation is applied to a wide range of CBIR systems and applications to achieve robustness. However a precise image segmentation method that can be applied to general image collections has not yet been found. In the absence of an accurate segmentation approach, a sliding window approach over location and scale has shown to be quite effective (Hiremath and Pujari 2007).

Feature extraction is an important step largely responsible for the performance of CBIR. Colour is one of the most cognitively significant features in an image and it is the simplest and most extensively used feature in image retrieval (Hiremath and Pujari 2007, Mansoori, et al. 2013, Saad, et al. 2011) (notwithstanding human ability to work well with Greyscale images). The texture feature has been used in CBIR systems in different ways (Hiremath and Pujari 2007, Takala et al. 2005, Yuan, et al. 2011), and these feature extraction methods can be classified as statistical, spectral or structural. Texture describes the content of many real world images and provides important characteristics for surface and object identification. Similar to the aforementioned features, shape is also an important feature in CBIR (Hiremath and Pujari 2007, Saad, et al. 2011), especially when dealing with objects that have clear shapes. Shape features are classified as both boundary-based and region-based methods. Different systems use different features and combinations of these features. In this approach BoW approach uses all these extracted image features to generate visual vocabularies.

During the past decade, the BoW approach has achieved popularity in the fields of classification and retrieval in CBIR (Yuan, et al. 2011, Mansoori, et al. 2013) because of its simplicity and good relative performance. It is also suitable for large databases as it scales efficiently to large collections. This approach was introduced by Sivic and Zisserman (Sivic and Zisserman 2003) to the computer vision community and it was inspired by the BoW model in text document retrieval. The visual vocabulary or visual codebook is formed by clustering image features that are extracted from images in the database. Firstly, similar features are gathered together where each cluster centre stands for a visual word. After that, feature vectors are mapped to those visual words and each image is represented as a histogram of visual words. Spatial information has been introduced to the BoW approach (Lazebnik et al.2006) to improve the results. However, like the other representations, the BoW approach still requires dealing with high-dimensionality data, which presents scalability challenges.

High dimensional indexing is one of the prevailing challenges in CBIR. Quite a few systems have addressed the problem of high dimensionality features. The well-known techniques used in CBIR include wavelet transform (Elharar et al. 2007), discrete cosine transform (DCT) (Elharar et al. 2007), latent semantic analysis (LSA) (Gorman et al. 2006), principal component analysis (PCA) (Banda et al. 2013), singular value decomposition (SVD) (Banda et al. 2013) and locality sensitive hashing (LSH) (Gorman et al. 2006). They are designed to reduce the dimensionality of feature vectors while maintaining the information in the descriptors as much as possible. Features encoded in a lower dimensional space must contain enough information to usefully distinguish between classes of images and perform well. Random indexing (RI) has been used and has shown great promise as a dimensionality-reduction technique in text retrieval (De Vries et al. 2009, Gorman et al. 2006). Compared to other methodologies RI has low computational cost, lower complexity, competitive accuracy, and most importantly it is an incremental approach (Magnus 2005, De Vries et al. 2009). Reducing the dimensionality of features has a considerable effect on the way that feature vectors are stored and retrieved. The reduced dimensional space provides a compressed representation of the original feature space.

The BoW representation as used in the literature is generally a histogram of visual words (Mansoori, et al. 2013). The BoW approach that is presented in this study is different however, in that it literally considers images as text documents and represents them as sets of visual-words (represented as symbolic tokens). The process developed in this research moves image retrieval further into the text retrieval domain. In the literature different techniques are adapted to reduce the dimensionality of images and this paper introduces RI as a dimensionality reduction technique used in text retrieval, to CBIR.

The objective of this research is to develop a novel CBIR system in order to achieve faster and adequately precise image classification and retrieval. An extended BoW feature representation method is developed, therefore utilizing subdivisions of each image into equally sized non overlapping tiles to generate a codebook. In order to achieve this objective, RI is applied to the BoW representation of images, introducing a new approach to the definition of image signatures. This BoW representation is also novel in the way that it translates an image into a bag of visual words document. Empirical evaluation is performed on a standard Corel dataset to validate the performance of this method against other independently evaluated methods.

3 Overview of CBIR-ISIG System

Feature extraction plays a major role in CBIR. The extracted features are used to index the images in CBIR. Comparisons of defined techniques indicate that a single feature for image retrieval is not an adequate solution. General images may have some colour images, images with texture, images of objects, and so on. Therefore it is concluded that multiple feature representation for image retrieval is necessary and this study proposes a novel approach that uses a combination of low-level features including colour, texture, shape, and GIST. The main reason for the selection of these descriptors is to address retrieval in a heterogeneous database of images. The following section defines features with the feature descriptors that are critical for accurate retrieval. These particular features are selected because of their reported performance and variation of feature description as described in the literature.

3.1 Colour Feature

Colour plays an important role in image retrieval and has been widely considered in feature extraction in the literature (Liua et al. 2007). There are a number of colour descriptors and the three most popular colour descriptors; the colour histogram, colour moments and colour coherence vector are selected for this study. Colour histograms are efficient and insensitive to small changes in camera view point. Colour histogram was adapted from (Qiu 2002) as it achieved better retrieval results using the YCbCr colour space, providing a closer match with human perception. The process of generating the histogram is as described in (Qiu 2002), and is most comparable to the colour set approach. Colour moments overcome the quantization effects in histograms and gives colour distribution. First order original moment, second order central moment, and third order central moment are calculated. Colour coherence vector (Pass 1996) includes spatial information; it classifies each pixel in a colour bucket as coherent or incoherent. Eight colour components are used in this approach.

3.2 Texture Feature

Notwithstanding the fact that texture is not well defined, it is very helpful to describe real world images. The well-known Gabor wavelet, Wavelet transforms and Edge histogram descriptor are selected as texture descriptors in this study. The Gabor features are widely adapted and have performed well in CBIR and they are also used in this proposed system. In this paper five scales and eight orientations are used. The rotation and scale invariance property is achieved by simple circular shift operation proposed in (Rahmanan, et al. 2011). Mean and standard deviation of each filter are used as a feature vector. The Wavelets transform provides a good multi resolution tool for texture description and allows the representation of a texture having various spatial resolution. It effectively describes both global and local information. Daubechies wavelets are chosen because they are better for general-purpose images search (Manthalkar, et al. 2003). Here decomposition is done up to three levels and each time the low frequency sub band is decomposed. The mean and standard deviation of each sub band are computed. 2003) is used to achieve the rotation and scale invariance

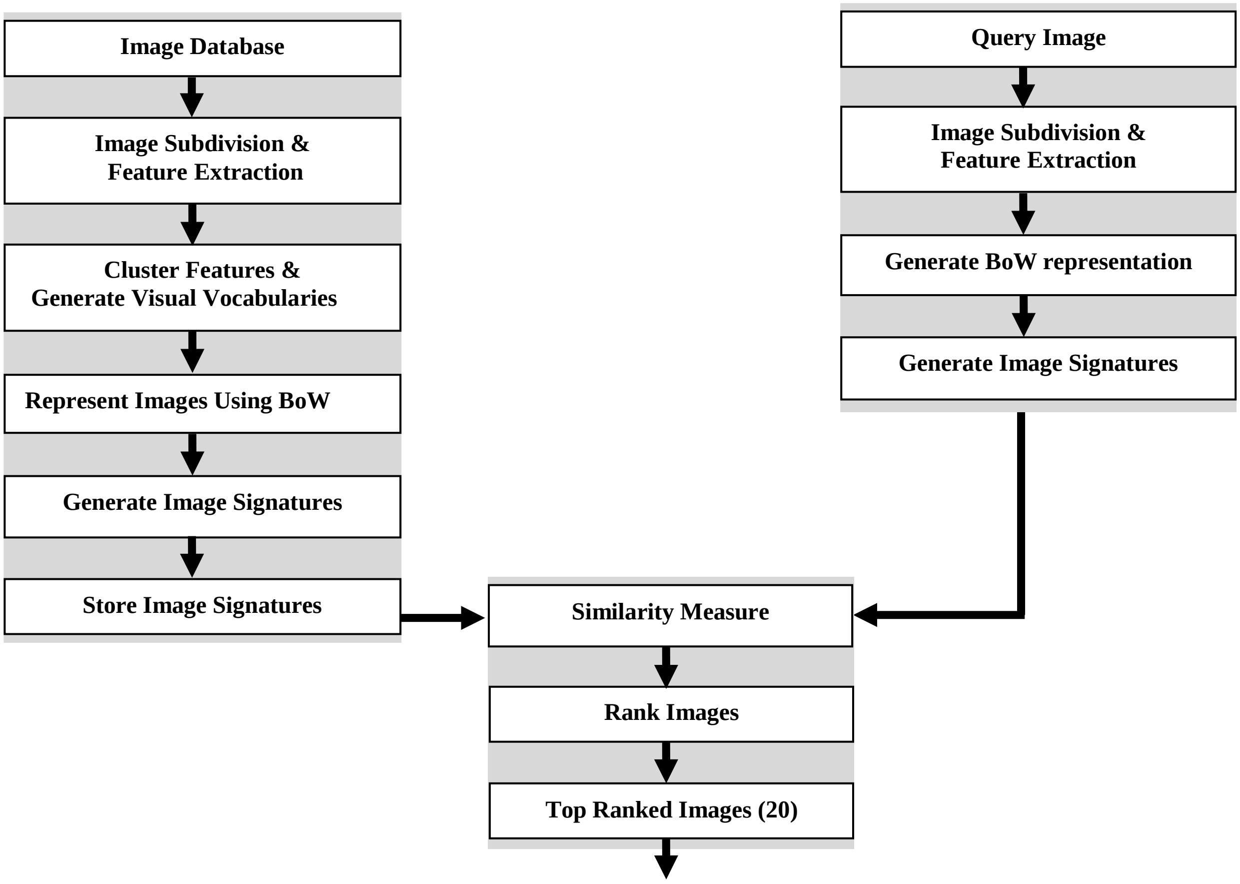

Figure 2: An Overview of the CBIR-ISIG System

The simple operation, proposed in (Manthalkar, et al. property. The Edge histogram descriptor has shown reliable performance and it effectively describes heterogeneous textures (Agarwal, et al. 2013). It captures the spatial distribution of edges and helps to extract different textures using five filters.

3.3 Shape Feature

The most common Generic Fourier descriptor and the invariant moments are used for shape description. The Generic Fourier descriptor is a region based method and suitable for general image retrieval. It is translational, rotational, scale invariant and robust to noise and occlusion too (Minaqiang et al. 2008). Four radial frequencies and 15 angular frequencies are used here. Although this is computationally expensive, it generates excellent retrieval performance. The invariant moment is an invariant feature and widely used for shape retrieval task. It gives compact representation on pixel distribution of a shape image. Moments are limited to seven by the calculation that use of higher order moments result in being sensitive to noise thus cause hindrance to accuracy.

3.4 GIST

The GIST descriptor describes the spatial envelope of the image and has shown good retrieval performance (Torralba, et al. 2008) in the literature. So the GIST feature used in this approach as well. This system uses code developed for the SUN database (Jianxiong, et al. 2010) where greyscale with eight orientations and three scales are used.

All the above features are adapted in CBIR-ISIG with the aim of improving retrieval performance. MATLAB is used for feature extraction. Each image is partitioned into 9 equal sized sub-images, by dividing each image into 3 by 3 grid. A family of image features, like shape, texture, and colour, characterizes every sub-image. In this research it was intended to provide some degree of spatial invariance through these sub-images. Even though only one signature is generated at the end, it is derived from a combination of sub-image signatures.

Different researchers try to achieve rotation translation scale invariance (RTSI) in different ways. Some low level features already have these properties and adapted existing techniques, but these are still low-level features. Rather than relying exclusively on explicit RTSI features, this proposed approach also relies on the BOW invariance property – sub-images are not indexed by position (in the same manner that the positions of words in text are not used in BoW text indexing.) Underlying features are the basic features like Colour histogram, Wavelet transform, GIST.

It is necessary to adapt the BoW approach used in document retrieval in order to apply it in an analogous way to images. After extracting low level features it is necessary to select an appropriate multidimensional indexing algorithm to index them. Clustering is a promising technique among indexing techniques. It is first necessary to cluster the image features in order to obtain discrete representation of feature sets. K-means is one of the simplest and best-known unsupervised clustering algorithms that can be easily implemented for feature vocabulary generation.

After extracting features, independent visual vocabularies are generated. In order to achieve this, all the features are clustered, separately into groups where similar feature vectors are placed together. Here the size of the vocabulary is the number of clusters generated and each cluster centre is considered as a visual word in the vocabulary. Then each sub-image and the full image are represented by visual words from these vocabularies through codebook lookup of each raw image feature. The images are represented symbolically, just like text, by using the codebook label of each cluster as a visual word to encode the feature. Other approaches which use BoW to represent images use histogram representation, by counting how many times each word appears in an image. However in our approach an image is regarded as analogous to a document and sub-image is regarded as analogous to a paragraph in a document.

If a BoW representation of a full image is denoted as Z it can be defined as,

$Z = X_a$, $a \in \{1,....,M\}$,Where X_a is a sub-image representation and M is the number of sub-images in an image;

$X = \{f_i\_c_j\}$, $i \in \{1,....,N\}$ and $j \in \{1,......,K\}$,

Where f_i is a local feature, N is the number of features used in the system and each feature is given a number, c_j is a cluster number (to which cluster that word belongs) and K is the vocabulary size. Each feature is given a number to denote it within the representation.

Image signatures are generated for sub-images as well as the full image. This is done to significantly reduce the dimensionality of the representation. The descriptors' dimensionality is important and heavily influences the complexity of the similarity measure in retrieval, and the memory requirements for storing the descriptors. There are several approaches for dimensionality reduction including RI (Magnus 2005), which is an efficient, scalable and incremental approach, based on random projection to avoid the computational cost for matrix factorization (Geva and De Vries 2011). One prime advantage of RI is that it can work directly with symbolic features, for instance, words in documents. RI is used effectively in text retrieval applications to reduce the dimensionality of documents without significant degradation in retrieval quality. RI can produce binary object signatures. The representation of objects as bit vectors lends itself to efficient processing, with low level bitwise operations supported on all conventional processor architectures. Most importantly, RI can be performed incrementally aligning with the new data arrival, as is crucial for online systems. Therefore in our approach RI is used for dimensionality reduction and to create image signatures. This allows the feature vector space to be reduced in dimensionality without expensive factorization such as, for instance latent semantic analysis (LSA) techniques. Seeding a pseudorandom number generator with the feature hash, and then generating a feature signature is used to create pseudo-random sparse ternary feature vector having values from {-1,0,+1}. A common choice with RI is to assign the proportion of vector elements with each value {-1,0,+1} to be respectively 1/6, 2/3 and 1/6.

All feature vectors in an image, of the entire image and of each sub-image, are then summed to produce a single image index vector. The image index vector is then squashed into a binary signature by assigning 1 bits to positive values and 0 bits to negative values. Similar images that share similar features will have similar signatures. Image signatures can then be compared for similarity by taking the bitwise (Hamming) distance between them. This technique can be used as a highly efficient replacement for a cosine similarity calculation in the original feature vector space. This approach uses a signature search-engine for searching. The motivation for using signatures to represent images comes from the fact that computation time quickly becomes a bottleneck when dealing with large databases and signature search engines can retrieve results from web-scale collections in milliseconds (Chappell, et al. 2013). Topsig (Geva and De Vries 2011) which is available in open source, is used to generate and search signatures in our CBIR system. This paper is concerned with identifying the ability of our approach to represent and then find images, rather than the signature matching and searching mechanism itself, so in fact it allows us to use any signature search engine regardless of specifications. Our concern is with how well the signatures would represent the images. The proposed signature based approach was evaluated using a gold standard benchmark i.e. subset of Corel dataset which is described in section 4.

An overview of the proposed CBIR system is depicted in Figure 2.

4 Evaluation

The Wang dataset of 1000 images was used for both evaluation of the system and comparison with the other systems. The Wang dataset of 1000 images is a subset of manually selected images from Corel image database and it was previously used in CBIR as a standard dataset for evaluation purposes; hence, it is convenient to re-use here since it provides a baseline for comparison with other independently developed and tested approaches. It consists of 10 classes with 100 images in each and they are African people, Beaches, Buildings, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountains, and Food. These images are JPEG with the resolution of 384x256 or 256x384.

During the evaluation features are extracted from the query images and represented using nearest cluster centres as in section 3. Then an image signature is generated as described in section 3. Finally the query signature is compared with that of the image database and top k images are retrieved from the database. Hamming distance is used for the similarity measure.

The most common evaluation measure in information retrieval is precision and it is used to evaluate the CBIR-ISIG system. Precision is the fraction of retrieved images that are relevant to the query and it is defined as;

(a) (b)

Figure 3: Search Results of CBIR-ISIG system for two queries (Query image is the top left most one)

Class	2005 [1]	2007 [2]	2011 [3]	2011 [4]	2013 [5]	CBIR-ISIG
Africans	0.23	0.48	0.57	**0.90**	0.70	0.75
Beach	0.23	0.34	0.58	0.38	0.28	**0.64**
Building	0.23	0.36	0.43	**0.72**	0.56	0.50
Bus	0.23	0.61	0.93	0.49	0.84	**0.85**
Dinosaur	0.23	0.95	0.98	1.00	0.81	**1.00**
Elephant	0.23	0.48	0.58	0.39	0.58	**0.70**
Flower	0.23	0.61	0.83	0.56	0.55	**0.95**
Horse	0.23	0.74	0.68	0.87	0.87	**0.94**
Mountain	0.23	0.42	0.46	0.45	0.48	**0.58**
Food	0.23	0.50	0.53	**0.87**	0.66	0.69
Average	0.23	0.55	0.66	0.66	0.63	**0.76**

Table 1: Average Precision (AP) of each class along with whole dataset with performance in the literature (AP for the top 20 images)

Class	2000 [6]	2002 [7]	2008 [8]	2009 [9]	2012 [10]	CBIR-ISIG
Africans	0.48	0.47	0.48	0.45	0.49	**0.50**
Beach	0.33	0.33	0.34	0.35	0.40	**0.45**
Building	0.33	0.33	0.33	0.35	**0.39**	0.33
Bus	0.36	0.60	0.52	0.60	0.58	**0.62**
Dinosaur	0.98	0.95	0.95	0.95	0.96	**0.98**
Elephant	0.40	0.25	0.40	**0.60**	0.50	0.44
Flower	0.40	0.63	0.60	0.65	0.75	**0.75**
Horse	0.72	0.63	0.70	0.70	**0.80**	0.68
Mountain	0.34	0.25	0.36	0.40	**0.40**	0.36
Food	0.34	0.49	0.46	0.40	**0.51**	0.41
Average	0.47	0.49	0.51	0.55	0.55	0.55

Table 2: Average Precision (AP) of each class along with whole dataset with performance in the literature (AP for the top 100 images)

$$\text{Precision} = \frac{|\{\text{Relevant images}\} \cap \{\text{Retrieved images}\}|}{|\{\text{Retrieved images}\}|}$$

Average Precision $P(c)$ ($1 \le c \le 10$) is taken for each class as follows.

$$P(c) = \frac{1}{N}\sum_{i=1}^{N} p(i)$$

Where $p(i)$ is the average precision of i^{th} query image and N is the number of images used for evaluation.

When N=20, each class has achieved more than 50% accuracy in the dataset and most of the classes have achieved the highest average precision among the compared systems. When N=100 performance of this system is relatively reduced, but still performs well by comparison with other systems. The systems which are compared are represented by year of publication and reference. Maximum average precision values of each class are highlighted Table 1 shows the results of CBIR-ISIG compared with the other systems (average precision for the top 20) and it shows the CBIR-ISIG system generates better results using this approach. They are not statistically significantly better (at 95% significance level), but averages are higher. In addition Table 2 shows the results of CBIR-ISIG compared with other systems (average precision for the top 100). Here total average precision reaches a highest in the CBIR-ISIG system for the top 20 and average performance for the top 100. Bold values show the highest among the compared systems.

References for the compared systems as given bellow for Table 1 and Table 2.

[1]- (Takala, et al. 2005), [2]- (Hiremath and Pujari 2007), [3]- (Yuan, et al. 2011), [4]- (Saad. et al. 2011), [5]- (Mansoori, et al. 2013) [6]- (Li, et al. 2000), [7]- (Chen 2002), [8]- (Hiremath and Pujari 2007), [9]- (Banerjee, et al. 2009), [10]- (Chowdhury 2012).

Figure 3 shows the top 20 images for a given query. The query image is at the top left. (a) sample from flowers class (19 out of 20) and (b) horses class (20 out of 20).

Our approach inherits the scalability of the particular signature search engine which was used here . This signature search engine (Geva and De Vries 2011) was reported to be capable of searching millions of signatures in milliseconds and so our approach can operate at that speed when the same signature size is used. However as this approach inherits the properties of signatures it will improve retrieval speed and reduce the memory size required to store features. But there is a trade-off between efficiency and accuracy. Signature size can be selected according to our need. Table 3 depicts the changes in AP with signature size. In this research 8192 bits (1024 bytes) signature size is chosen as it is

Class	Signature Size (in bits)										
	64	**128**	**256**	**512**	**1024**	**2048**	**4096**	**8192**	**16384**	**32768**	**65536**
Africans	0.40	0.55	0.64	0.70	0.70	0.73	0.73	0.75	0.75	0.74	0.74
Beach	0.31	0.47	0.51	0.59	0.61	0.63	0.65	0.64	0.64	0.64	0.64
Building	0.35	0.38	0.40	0.40	0.46	0.48	0.47	0.50	0.48	0.48	0.48
Bus	0.55	0.76	0.77	0.80	0.81	0.84	0.84	0.85	0.84	0.84	0.84
Dinosaur	0.88	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Elephant	0.31	0.59	0.57	0.62	0.65	0.69	0.69	0.70	0.70	0.70	0.70
Flower	0.71	0.90	0.95	0.94	0.94	0.94	0.94	0.95	0.95	0.95	0.95
Horse	0.65	0.83	0.88	0.91	0.91	0.92	0.93	0.94	0.93	0.93	0.93
Mountain	0.29	0.46	0.40	0.46	0.58	0.53	0.56	0.58	0.58	0.58	0.58
Food	0.32	0.55	0.54	0.64	0.65	0.69	0.68	0.69	0.69	0.69	0.68
Average	0.477	0.658	0.666	0.706	0.731	0.745	0.749	**0.76**	0.756	0.755	0.754
	0.48	0.66	0.67	0.71	0.73	0.74	0.75	**0.76**	0.76	0.76	0.75

Table 3: Average Precision (AP) of each class along with whole dataset with different signature size (AP for the top 20 images)

sufficient to demonstrate improvement. It only accommodates around 20% of the real valued feature vector in size, but even if use 128 bits (16 bytes) the results are respectable. Signature size can be chosen according to the requirement. There is a significant computational cost in generating the raw image features Feature extraction involves relatively heavy image processing, which is time consuming. This cost is not unique to our system and is incurred by any system that uses the same feature set. The cost of image signature generation comes from clustering features in order to generate a BoW representation, and the cost of generating signatures from the BoW representation. This entire process is done only once during indexing and can be trivially parallelised since there is no dependence between images. At search time the process is performed only on the search argument (a single image). and it is fast enough to support immediate response to the user.

The speed at which a signature can be generated is limited by the complexity of feature extraction, essentially conventional image processing and not by random indexing which consumes negligible amount of time, by comparison. In our computational configuration the average time needed to extract the above features from a 64 by 64 image is 0.1347 seconds, using Windows Core i5, 1.8GHz, and using MATLAB for image processing.

5 Conclusion

This paper presented a novel approach to represent images using image signatures which are derived by applying RI to a BoW representation of images. These image signatures improve the retrieval speed and reduce the requirement of memory for storage. The CBIR-ISIG system shows more than 50% average precision for the top 20 images in each class and achieved 75% average precision which is best among all the compared system against a standard subset of the Corel dataset. The results however are still limited and the approach will be tested with larger datasets in future work. The most obvious applications of this approach are large scale CBIR for heterogeneous collections and near-duplicate detection.

6 Reference

Yuan, X. Yu, J. Qin, Z. and Wan, T. (2011): A SIFT-LBP Image Retrieval Model Based on Bag-of-Features. *Proc. IEEE ICIP 18th International Conference on Image Processing*, Brussels, Belgium, 1061-1064, Brussels, Belgium.

Mansoori, N.S. Nejati, M. Razzaghi, P. and Samavi, S. (2013): Bag of visual words approach for image retrieval using colour information. *Proc. ICEE 21st Iranian Conference on Electrical Engineering*, Mashhad, 1-6, IEEE.

De Vries C. M., V. Lance D. V. and Shlomo G. (2009): Random Indexing K-tree. *Proc. ADCS of the 14th Australian Document Computing Symposium*, Sydney, Australia, 1-8, University of Sydney.

Liua, Y. Zhanga, D. Lua, G. and Mab, W.Y. (2007): A Survey of Content-based Image Retrieval with High-Level Semantics. *Journal of Pattern Recognition*, **40**(1):262-282.

Saad, M. H. Saleh, H. I. Konbor, H. and Ashour, M. (2011), Image retrieval based on integration between YCbCr colour histogram and shape feature, *Proc. ICENCO 7th International Computer Engineering Conference*, Giza, 97-102, IEEE.

Hiremath, P. S. and Pujari, J. (2007): Content Based Image Retrieval based on Colour, Texture and Shape features. *Proc. 15th International Conference on Advance Computing and Communications*, Guwahati, Assam, 780-784, IEEE.

Takala, V. Ahoen, T. and Pietikainen, M. (2005): Block-Based Methods for Image Retrieval Using Local Binary Patterns. *Proc. of the 14th Scandinavian Conference on Image Analysis*, Joensuu, Finland, **3540**: 882-891, Springer Berlin Heidelberg.

Sivic, J. and Zisserman, A. (2003): Video Google: a text retrieval approach to object matching in videos. *Proc.*

9th IEEE International Conference on Computer Vision, Nice, France, **2**:1470-1477, IEEE.

Chen, Y. and Wang, J. (2002): A region-based fuzzy feature matching approach to content-based image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**(9):1252-1267.

Elharar, E. Stern, A. Hadar, O. and Javidi, B. (2007): A hybrid compression method for integral images using discrete wavelet transform and discrete cosine transform. *Journal of display technology*, **3**(3):321--325.

Gorman, J. Curran, J. R. (2006): Scaling Distributional Similarity to Large Corpora. *Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics*, Sydney, Australia, 361-368, Association for Computational Linguistics, Stroudsburg, PA, USA.

Banda, J.M. Angryk, R.A. and Martens, P.C.H. (2013): On Dimensionality Reduction for Indexing and Retrieval of Large-Scale Solar Image Data, Journal of Solar Physics, **283**:113-141.

Magnus, S. (2005): An introduction to random indexing. *Proc. 7th International Conference on Terminology and Knowledge Engineering- Methods and Applications of Semantic Indexing Workshop*, TKE, 1-6, IEEE.

Geva, S. and De Vries. C. M. (2011): TOPSIG : Topology Preserving Document Signatures, *Proc. CIKM '11 20th ACM international conference on Information and knowledge management*, New York, NY, USA, 333-338, ACM Press.

Li, J. Wang and J. Z. Wiederhold, G. (2000): IRM: Integrated Region Matching for Image Retrieval. *Proc. ACM 8th international conference on Multimedia*, New York, NY, USA, 147-156, ACM Press.

Qiu, G. (2002): Indexing chromatic and achromatic patterns for content-based. *Journal of pattern recognition*, **35**:1675—1686.

Pass, G. Zabih, R. and Miller, J. Comparing Images Using Colour Coherence Vectors. (1996), *Proc. ACM fourth international conference on Multimedia*, New York, NY, USA, 65-73, ACM Press.

Rahmana, M.H. Pickering, M.R. Frater, M.R. (2011) "Scale and Rotation Invariant Gabor Features for Texture Retrieval," *Proc. DICTA International Conference on Digital Image Computing Techniques and Applications*, Noosa, QLD, 602-607, IEEE.

Manthalkar, R. Biswas, P. K. Chatterji, B.N. Rotation (2003) Scale invariant Texture Features Using Discrete Wavelet Packet Transform. *Journal of Pattern Recognition*, **24**(14):2455-2462.

Agarwal, S. Verma, AK. and Singh, P. (2013): Content Based Image Retrieval using Discrete Wavelet Transform and Edge Histogram Descriptor. *Proc. ISCON International Conference on Information Systems and Computer Networks*, Mashhad, 19-23, IEEE.

Minaqiang, Y. Kidiyo, K. and Joseph, R. (2008): A survey of shape feature extraction techniques. *Journal of Pattern Recognition*, 43-90.

Torralba, A. Fergus, R. and Freeman, W.T. (2008): 80 Million Tiny Images: A Large Data Set for Nonparametric Object and Scene Recognition. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, **30**(11):1958-1970.

Jianxiong, X. Hays, J. Ehinger. and K.A, Oliva. A, Torralba. (2010): A, SUN database: Large-scale scene recognition from abbey to zoo. *Proc. IEEE CVPR Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, 3485-3492, IEEE.

Chappell, T. Geva, S. Nguyen, A. and Zuccon, G. (2013): Efficient Top-k Retrieval with Signatures. *Proc. ADCS 18th Australasian Document Computing Symposium*, Brisbane, Australia, 10-17, ACM Press.

Hiremath, P. S. and Pujari, J. (2008): Content Based Image Retrieval Using Colour Boosted Salient Points and Shape Features of an Image. *Proc. International Journal of Image Processing*, **2**(1):10-17.

Banerjee, M. Kundu, M. and Maji, P. (2009): Content-based image retrieval using visually significant point features. *Journal of Fuzzy Sets and Systems*, 160(23):3323--3341.

Chowdhury, M. Das, S. and Kundu, M. (2012): Interactive content based image retrieval using Ripplet transform and fuzzy relevance feedback. *Proc. 1st Indo-Japan Conference*, Kolkata, India, 243-251, Springer.

Lazebnik, S., Schmid, C and Ponce, J. (2006): Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, **2**:2169-2178