IET Computer Vision

*Research Article*

IET Journals
The Institution of Engineering and Technology

# Content-based image retrieval system via sparse representation

ISSN 1751-9632
Received on 20th September 2014
Revised on 20th July 2015
Accepted on 27th July 2015
doi: 10.1049/iet-cvi.2015.0165
www.ietdl.org

*Sajad Mohamadzadeh, Hassan Farsi* ✉

*Department of Electronics and Communications Engineering, University of Birjand, Birjand, Iran*
✉ *E-mail: hfarsi@birjand.ac.ir*

**Abstract:** The aim of image retrieval systems is to automatically assess, retrieve and represent relative images-based user demand. However, the accuracy and speed of image retrieval are still an interesting topic of many researches. In this study, a new method based on sparse representation and iterative discrete wavelet transform has been proposed. To evaluate the applicability of the proposed feature-based sparse representation for image retrieval technique, the precision at percent recall and average normalised modified retrieval rank are used as quantitative metrics to compare different methods. The experimental results show that the proposed method provides better performance in comparison with other methods.

## 1 Introduction

Image retrieval is fundamentally important to numerous multimedia information processing systems and applications. Image retrieval is a process which searches a query image among image datasets and then finds and represents the images related to the user's demand. Image retrieval systems are studied based on text and content which are the two main research fields [1]. Research on image retrieval had been started from 1970. In initial image retrieval systems, all images were manually annotated using text descriptors and then a dataset management system was used. However, these systems have many shortcomings. For example, the concept of an image is higher than a series of words for description of images, manual indexing is hard, and this process is costly and time consuming [2]. In the beginning of 1990s, due to rising capacious images used in Internet and also the shortcomings mentioned above, text-based image retrieval was known as an inefficient method and demand for content-based image retrieval (CBIR) increased. The CBIR techniques use vision features for interpretation of image concept instead of annotation of manual texture [2].

Feature extraction is one of the most important procedures used for interpreting and indexing images in CBIR systems [3]. This process produces the feature vector for all images in a dataset and represents the image concept in image classification methods. The size of feature vector has to be smaller than the image size. This reduces the search time, simplifies the search process and retrieves the same image as fast as possible. There are many proposed methods and approaches for classifying, indexing, searching and retrieving of visual information based on analysis of low-level image features like colour, texture, shape and so on [4]. The combination of these features has been shown more efficient performance in image retrieval systems [5].

### 1.1 Colour and texture features

Colour feature is one of the most common and determinant features used in image retrieval systems and is stable against direction variations, size of image and background complexity [6]. In this paper, we have used HSI and CIE-$L^*a^*b^*$ colour spaces, because these colour spaces contain uniform perception which means the distance between colours has to be proportional with human perception distinctions [7]. Texture feature refers to visible patterns containing compatible features where these features are not obtained by only one colour or light intensity. The texture contains natural and virtual features for some surfaces such as clouds, bricks and fabric and includes some important information about surface structure arguments and their relation to the environment or the circumstance. Due to the importance of texture feature, it has been widely used in different fields such as pattern recognition, machine vision and image retrieval [8]. Several methods have been proposed for description of image texture. Jain and Tuceryan have divided the texture analysis methods into four groups: statistical, geometrical or structural, model based and signal processing methods [9]. In this paper, we have used signal processing method because it results in higher performance in comparison with the other above mentioned methods [10].

### 1.2 Compressed sensing (CS) and sparse representation

During the recent years, CS has attracted considerable attention in many areas of applied mathematics, computer science, and electrical engineering by suggesting that it may be possible to surpass the traditional limits of sampling theory. CS represents many signals using only a few non-zero coefficients in a suitable basis or dictionary. Non-linear optimisation can then enable recovery of such signals from very few measurements [11].

Transform coding is one of the most popular techniques for signal compression, and typically relies on finding a basis or frame that provides compressible or sparse representations for signals in a class of interest [11, 12]. Compressible representation is well-approximated of a signal with only $k$ non-zero coefficients and sparse representation can represent a signal of length n with $k \ll n$ non-zero coefficients. Compressible and sparse signals can be represented with high fidelity by preserving only the values and locations of the largest coefficients of the signal. This process is called sparse approximation, and it forms the foundation of transform coding schemes that exploit signal compressibility and sparsity, including the JPEG, JPEG2000, MPEG, and MP3 standards [11]. Compressed sensing enables a potentially vast reduction in the sampling and computation costs for sensing signals that have a sparse or compressible representation. Sparse representation of signal can extensively reduce the number of measurements that need to be stored, while the Nyquist-Shannon sampling theorem states that a certain minimum number of samples is required in order to perfectly capture an arbitrary band-limited signal. The CS field has central challenges such as the design of a small set of linear, non-adaptive measurements

schemes and their extensions to practical data models and acquisition systems [13]. Candes, Romberg, Tao, Donoho and Elad [11–14] have grown out the CS field such that a signal having a sparse representation can be recovered exactly from a small set of linear, no adaptive measurements. These results suggest that it may be possible to sense sparse signals by taking far fewer measurements, hence the name compressed sensing. The CS differs from a classical sampling in three important respects. First, a sampling theory typically considers in finite length, continuous-time signals. In contrast, the CS is a mathematical theory focused on measuring finite-dimensional vectors in $R^n$. Second, rather than sampling the signal at the specific points in time, the CS systems typically acquire measurements in the form of inner products between the signal and more general test functions. This is in fact in the spirit of modern sampling methods which similarly acquire signals by more general linear measurements. Third, the two frameworks differ in the manner in which they deal with signal recovery. In the CS, signal recovery is typically achieved using highly non-linear methods. However, in the Nyquist-Shannon framework, the signal recovery is achieved using linear process that requires little computation and has a simple interpretation [11, 14].

Sparse representation has been shown to have applications in many areas such as geophysics, speech recognition [15], image compression, image and signal processing, image enhancement, sensor networks and computer vision [12]. As we have shown in Section 4, the main focus of this paper is to improve retrieval accuracy by using convex optimisation techniques.

### 1.3 Motivation

In spite of proposing various methods in image retrieval during the recent years, there is no efficient and perfect algorithm to find the related images. The aim of this paper is to propose a new algorithm for searching in several image datasets (independent from image datasets) with highly accurate retrieval, low complexity, and high speed retrieval. Our studies have been focused on colour and texture features. The proposed method combines new colour and texture features to increase retrieval performance. We use two colour spaces as colour feature and discrete wavelet transform (DWT) as signal processing texture feature because it is a powerful technique extensively used in the community of computer vision and pattern recognition to detect and describe local features in images [10]. In this paper, we propose a novel image retrieval system to find all related images and to compute similarity between query and dataset images via sparse representation technique. Finally, the proposed method is evaluated by five datasets and compared with several methods.

This paper has been organised as follows: In Section 2, the related research from the existing literature will be presented. In Section 3, the proposed image retrieval system via sparse representation is described step by step. Feature vector construction and sparse representation have been explained. In Section 4, evaluation measures, datasets and indexing results are detailed. Finally, the conclusion is drawn in Section 5.

## 2 Related work

During the recent years, different systems have been proposed that provide different and several algorithms for content based image retrieval. The main purpose of these systems is to provide the improvement of results based on the user's demand. In [16], Liz and Tan have proposed a system for designing an information retrieval system with ability of dealing with imaged document stored in digital libraries. In [17], it has been proposed an interesting solution for multi-attribute image retrieval by using a structural learning paradigm to model interdependency across query attributes. Several researchers have improved the retrieval accuracy of this approach from different perspectives, e.g. by building high-order features [18]. Most of these methods induce considerable computations, e.g. the spatial verification or the query expansion is much slower than the retrieval. Also, the identification of reliable composite features is costly. Instead, our method addresses both retrieval accuracy and efficiency. In [19], scale invariant feature transform feature-based sparse representation has been used in image retrieval and assessed the similarity between two images. In [20], Yang *et al.* have proposed a mobile image retrieval approach by exploiting the advantage of mobile end in which people usually take multiple photos of an object in different focuses and viewpoints. They have used visual similarity in mobile end and learned salient visual words by exploring saliency from relevant images, and finally determine the contribution order of salient visual words to carry out scalable retrieval. In their method, the soft spatial verification is proposed to re-rank the results. In [21], a clustering method using dictionary learning has been used to retrieve the medical images. In addition, the sparse representation based on learning dictionaries simultaneously via K-singular value decomposition has been proposed. A query image has been matched with the existing dictionaries to identify the dictionary with the sparsest representation using an orthogonal matching pursuit (OMP) algorithm. Then images in the cluster associated with this dictionary have been compared using a similarity measure to retrieve images similar to the query image. Liu *et al.* have proposed an image classification using constrained sparse coding in [22]. They have considered the visual similarities between local features jointly with the corresponding label information of local features. This has been achieved by combining the label constraints with the encoding of local features. In this way, they can ensure that similar local features with the same label are encoded with similar parameters. Local features with different labels have been encoded with dissimilar parameters to increase the discriminative power of encoded parameters. Meanwhile, they have combined the spatial and contextual information of local features [22]. In [23], we have used the HMMD colour space and Hadamard DWT (HDWT) to retrieve the related images. We have used discrete cosine transform (DCT) and principal component analysis to represent and retrieve images [24]. Given the aforementioned researches, the purpose of the following sections is to explore a procedure in which the accuracy of image retrieval may be retained. This, also, ensures that the processing time is reduced and the retrieval performance is improved. Meanwhile, the proposed method has been applied for the several datasets and all images of datasets, and the performance criteria have been obtained.

## 3 Image retrieval based on sparse representation

In this section, the image retrieval framework has been described in the following.

### 3.1 Role of feature extraction

Feature extraction method plays a very important role in image processing systems. In computer vision and image processing literature, numerous feature extraction methods have been investigated for finding projections that better separate the classes in lower dimensional spaces often referred to feature spaces. Feature vector of each image should represent the content of the image accurately [25]. Meanwhile, the size of feature vector has to be smaller than the image size. This reduces the search time and the storage memory, simplifies the search process and retrieves the same image as fast as possible. In the proposed method, the CIE-$L^*a^*b^*$ and HSI colour spaces have been used to extract iterative DWT (IDWT) features. The DWT has been used because it consists of recursively dividing the image into low- and high-frequency components. The DWT uses a multistage filtering to increase a frequency resolution. Firstly, the filtering is performed horizontally and then vertically on the image. Therefore, four components are extracted from the input image in each filtering stage [10]. Low-low, low-high, high-low and high-high frequency components are sub-bands of DWT. The


*IET Comput. Vis.*, 2016, Vol. 10, Iss. 1, pp. 95–102

low-low frequency components provide a coarse scale approximation of the image, while the other frequency components fill in details and resolve edges. Most coefficients of DWT are very small when we compute a wavelet transform of a typical natural image. Hence, we can obtain an accurate approximation of the image by setting the small coefficients to zero, or thresholding the coefficients, to obtain a sparse representation [26].

*3.1.1 Feature extraction based on IDWT:* IDWT feature has been used to extract colour and texture features of HSI and CIE-$L^*a^*b^*$ planes to construct the feature vector. The steps of IDWT method are represented as follows:

*Step 1:* Apply DWT on each HSI and CIE-$L^*a^*b^*$ planes separately with size of $N \times N$ to generate approximation (low-low), horizontal (low-high), vertical (high-low) and diagonal (high-high) components. Approximation components have been used for the next step because wavelet transform analyses the signal at various frequency bands and gives higher frequency resolution and lower time resolution at the lower frequencies [10]. The flowchart of DWT is represented in Fig. 1 and formulated by the following equation

$$[A_{N/2\times N/2}, H_{N/2\times N/2}, V_{N/2\times N/2}, D_{N/2\times N/2}] = \text{DWT (image plane}_{N\times N}) \quad (1)$$

where $A$, $H$, $V$ and $D$ are approximation, horizontal, vertical and diagonal components, respectively.
*Step 2:* Apply DWT only on the approximation component. The size of this component is $(N/2) \times (N/2)$ which shown in (1). The DWT method is repeated $L$ times on approximation component of each DWT output as shown in Fig. 2.

The IDWT feature vector of each datasets formulated by the following equation:

$$\text{IDWT feature vector} = \left[A_{N/2^L\times N/2^L}\right] \quad (2)$$

where $L$ is the number of iteration.

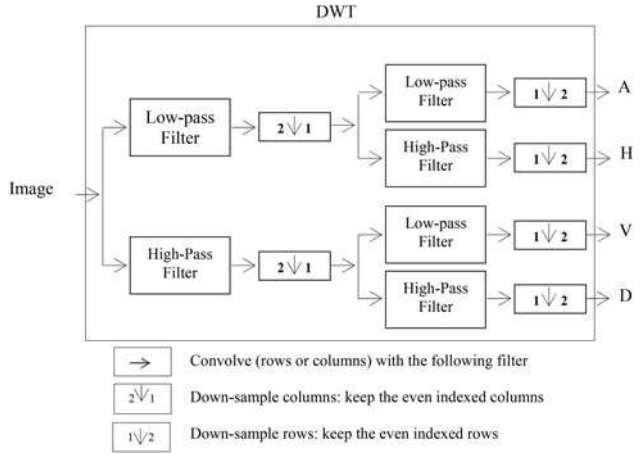


**Fig. 1** *Decomposition of the image by DWT [10]*

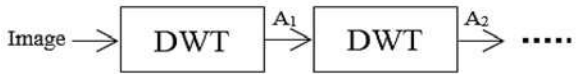


**Fig. 2** *Iterative DWT feature*

### 3.2 Sparse representation

Compressive sensing has been one of the hot topics in the signal and image processing in the past ten years. We, first, review some basics of the sparse representations. Then, we explain our proposed method for image retrieval application using sparse representation. In this paper, all data are assumed in the real domain. The concatenation of two vectors is written by

$$[\boldsymbol{x}_1; \boldsymbol{x}_2] \doteq \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix}; \quad [\boldsymbol{x}_1, \boldsymbol{x}_2] \doteq [\boldsymbol{x}_1 \quad \boldsymbol{x}_2]$$

and $l_p$-norm is annotated by $\|\cdot\|_p$.

Given a signal vector $\boldsymbol{b} \in R^m$, signal (or atomic) decomposition is the linear combination of $n$ basic atoms $a_i \in R^m$, $(1 \le i \le n)$ which constructs the signal vector $\boldsymbol{b}[n]$ as

$$\boldsymbol{b} = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = \boldsymbol{A}x \quad (3)$$

where $\boldsymbol{A} = [a_1, a_2, \ldots, a_n]$, $x = (x_1, x_2, \ldots, x_n)$.

The dictionary $\boldsymbol{A}$ comprises $n$ signals $(a_1, a_2, \ldots, a_n)$ called atoms. In discrete Fourier transform or classical signal decomposition, the number of atoms $(n)$ is equal to the length of signals $(m)$, an unique solution exists for this problem. However, when these two parameters are not equal $(n > m)$ means that the decomposition is not unique. The sparse decomposition aims to seek for a solution in which as few as possible atoms contribute in decomposition. This is equivalent to seeking the sparsest solution of the undetermined system of linear equation $\boldsymbol{b} = \boldsymbol{A}x$. The sparsest solution is sought for this equation by solving following optimisation problem

$$(P_0): \min_{x_0} \|x\|_0 \quad \text{subject to } \boldsymbol{b} = \boldsymbol{A}x \quad (4)$$

The $l_0$-norm counts the number of non-zero entries in a vector. It has been shown that (3) has got a solution for $\|x\|_0 < \text{spark}(A)/2$, where the spark of a given matrix $\boldsymbol{A}$ is the smallest number of columns from $\boldsymbol{A}$ that are linearly-dependent and $x$ is the unique sparsest solution [12]. However, the problem of finding the sparsest solution of an underdetermined system of linear equations is NP-hard and difficult even to approximate [27]. In the recent years, the several development methods have been proposed to solve (3) which some important methods are mentioned as follows.

However, if $x_0$ is sufficiently sparse (i.e. most of its entries in the canonical coordinates are zero), and the sensing matrix $\boldsymbol{A}$ is incoherent with the basis under which $x_0$ is sparse (i.e. the identity matrix for the canonical coordinates), then the solution of the $P_0$ minimisation problem (4) is equal to the solution to the following $P_1$ minimisation problem

$$(P_1): \min_{x} \|x\|_1 \quad \text{subject to } \boldsymbol{b} = \boldsymbol{A}x \quad (5)$$

This problem can be solved in polynomial time by standard linear programming methods [12]. Even more efficient methods are available when the solution is known to be very sparse. For example, homotopy algorithms recover solutions with $t$ non-zero in $O(t^3 + n)$ time, linear in the size of the training set [28].

So far, we have assumed that (3) holds exactly and have real data. However, $\boldsymbol{b}$ often contains noise in practice since real data are noisy. It may not be possible to express sparse representation of samples. In such cases, the model (3) can be modified by writing

$$\boldsymbol{b} = \boldsymbol{A}x + z \quad (6)$$

where $z \in R^m$ is a noise term with bounded energy $(\|z\|_2 \le \varepsilon)$. The sparse solution $x_0$ can still be approximately recovered by solving the following stable $l_1$-minimisation problem. The equality constraint can be relaxed, resulting in the constrained basis pursuit

de-noising (BPDN) problem [29]

$$(P_{1,2}): \min_x \|x\|_1 \quad \text{subject to } \|\boldsymbol{b} - \boldsymbol{A}x\|_2 \leq \varepsilon \tag{7}$$

where $\varepsilon > 0$ is a pre-determined noise level. This convex optimisation problem can be efficiently solved via second-order cone programming [29]. A variant of this problem is also well known as the unconstrained BPDN problem with a scalar weight $\lambda$

$$(QP_\lambda): \min_x \frac{1}{2}\|\boldsymbol{b} - \boldsymbol{A}x\|_2^2 + \lambda\|x\|_1 \tag{8}$$

Theoretical analysis of the BPDN problem [30] has shown that, although exact recovery of the ground-truth signal $x_0$ is not possible with noise in many cases (e.g. when the observation is corrupted by random Gaussian noise), it can be well approximated by the solution of $(P_{1,2})$ or $(QP_\lambda)$.

While problem $P_1$ associated with the compressive sensing can be formulated as a linear program (LP) and readily solved by classical methods in convex optimisation such as interior-point methods, the computational complexity of these methods is often too large for large-scale high-dimensional image datasets [12, 31]. In the past decade, many new efficient algorithms of real applications in various fields have been proposed. In this paper, several acceleration techniques have been used to solve (3) such as smoothed L0 (SL0), dual augmented Lagrangian method (DALM), primal augmented Lagrangian method (PALM), homotopy method, fast iterative soft-thresholding algorithm (FISTA) and approximate message passing (AMP) [12, 31, 32].

### 3.3 Sparse image retrieval algorithm

The main purpose of an image retrieval system is to retrieve relevant images from datasets. We focus on investigating sparse representation of IDWT features and finding its use in image retrieval applications. Therefore, the following algorithm via sparse representation and weighted elements has been applied. The flowchart of the proposed image retrieval algorithm has been represented in Fig. 3.

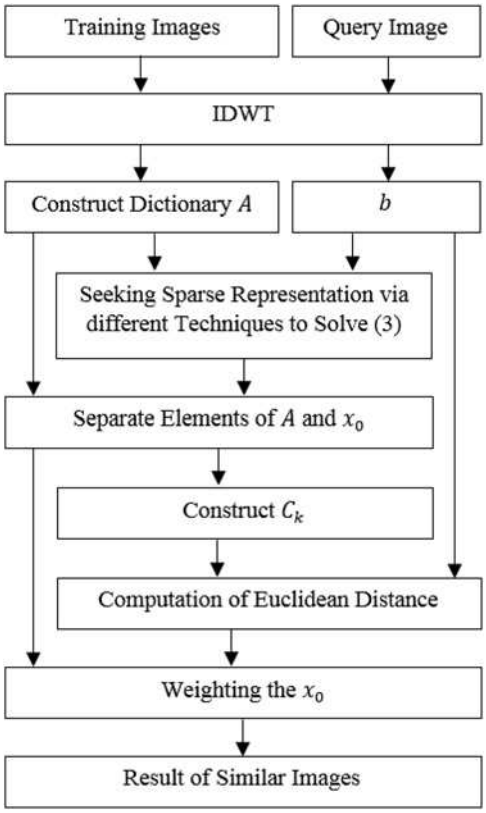


**Fig. 3** *Flowchart of the proposed algorithm*

(i) Construct dictionary $\boldsymbol{A}$ by using sufficient training samples of the $i$th image, $A_i = [v_{i,1}; v_{i,2}; \ldots; v_{i,m}] \in R^{m\times 1}$, where $v_{i,j}$ represents the $j$th feature of the $i$th extracted image by applying IDWT method on image. Therefore

$$\boldsymbol{A} = [A_1, A_2, \ldots, A_n] \in R^{m\times n} \tag{9}$$

(ii) Extract the feature vector of the query image by applying IDWT method, $\boldsymbol{b} \in R^m$.

(iii) Seek sparse representation, $x_0 \in R^n$, by solving (3) and using SL0, DALM, PALM, homotopy, FISTA and AMP techniques. Therefore, some elements of $x_0$ are zero except those associated with the $k$th class.

(iv) Separate elements of $\boldsymbol{A}$ and $x_0$ into $k$ clusters

$$x_0 = \left[\underbrace{\alpha_1; \alpha_2; \ldots}_{1}; \underbrace{\alpha_j \ldots}_{2}; \ldots; \underbrace{\ldots; \alpha_n}_{k}\right] = [x_{0,1}; x_{0,2}; \ldots; x_{0,k}] \tag{10}$$

$$\boldsymbol{A} = \left[\underbrace{A_1, A_2, \ldots}_{1}, \underbrace{A_j \ldots}_{2}, \ldots, \underbrace{\ldots, A_n}_{k}\right] = [D_1, D_2, \ldots, D_k] \tag{11}$$

(v) Define $C_k = D_k x_{0,k} \in R^m$ by using distinguished elements in the previous step.

(vi) Compute the Euclidean distance (ED) between the feature vector of query image ($b$)and $C_k$ by

$$\mathrm{ED}_k = \sqrt{\sum_{i=1}^{m} (b_i - C_i)^2} \tag{12}$$

(vii) Compute the weighting of the elements of $x_0$ by considering ED and (10)

$$x_{\text{weighted}} = \left[\frac{|x_{0,1}|}{\mathrm{ED}_1}; \frac{|x_{0,2}|}{\mathrm{ED}_2}; \cdots; \frac{|x_{0,k}|}{\mathrm{ED}_k}\right] \in R^n \tag{13}$$

(viii) viii.Finally, retrieve the best relevant images by using sorted element of $x_{\text{weighted}}$.

## 4 Experimental results

### 4.1 Evaluation measures

To evaluate the performance of the proposed retrieval system, we use three evaluation metrics. Montagna and Finlayson in [6] proposed a method using the combination of the precision and recall criteria as performance measures for CBIR systems. The precision and recall criteria are given by (14) and (15), respectively,

$$\text{Precision} = \frac{\text{Number\_of\_Relevant\_Images\_Retrieved}}{\text{Total\_Number\_of\_Images\_Retrieved}} \tag{14}$$

$$\text{Recall} = \frac{\text{Number\_of\_Relevant\_Images\_Retrieved}}{\text{Total\_Number\_of\_Relevant\_Images\_in\_Database}} \quad (15)$$

According to [6], the following measures have been adopted:

(a) $P(0.5)$, precision at 50% recall (i.e. precision after retrieving 1/2 of the relevant documents).
(b) $P(1)$, precision at 100% recall (i.e. precision after retrieving all of the relevant documents).

These two values have been defined because precision and recall are considered in relation to each other and they are not meaningful if taken separately.

(c) The average normalised modified retrieval rank (ANMRR) is an objective measure which summarises the system performance into a scalar value. It is defined from MPEG-7 research group [33, 34]. First, NG($q$), $K(q)$ and $R(k)$ are denoted as follows:

- $NG(q)$: The number of the ground truth images for a query $q$.
- $K(q) = \min\left[4.\left|\text{NG}(q)\right|, 2.\max\left\{\left|\text{NG}(q)\right|, \forall q\right\}\right]$.
- $R(k)$= rank of an image, $k$, in retrieval results.

Rank($k$) is obtained by

$$\text{Rank}(k) = \begin{cases} R(k) & \text{if } R(k) \le K(q) \\ 1.25K & \text{otherwise} \end{cases}. \quad (16)$$

Using (16), average rank, AVR($q$), for query, $q$, is given by

$$\text{AVR}(q) = \langle \text{Rank}(k) \rangle. \quad (17)$$

However, with ground truth sets of different size, the AVR($q$) value depends on NG($q$). To minimise the influence of variations in NG($q$), modified retrieval rank, MRR($q$), is obtained by

$$\text{MRR}(q) = \text{AVR}(q) - 0.5\left[1 + \left|\text{NG}(q)\right|\right]. \quad (18)$$

The upper bound of MRR($q$) depends on NG($q$). To normalise this value, normalised MRR, NMRR($q$), is obtained by

$$\text{NMRR}(q) = \frac{\text{AVR}(q) - 0.5\left[1 + \left|\text{NG}(q)\right|\right]}{1.25K(q) - 0.5\left[1 + \left|\text{NG}(q)\right|\right]}. \quad (19)$$

This measure is zero for perfect performance and approaches to one as performance worsens. The ANMRR of the dataset is finally given by averaging the NMRR($q$) overall the $q$'s

$$\text{ANMRR} = \langle \text{NMRR}(q) \rangle. \quad (20)$$

### 4.2 Datasets

To evaluate the proposed image retrieval system, five common different datasets have been collected and considered. (i) The Flower dataset [35] including 1360 variable size images in 17 classes. The images have large scale, pose and light variations and there are also classes with the large variations of images within the class and close similarity to the other classes. (ii) The Corel dataset [36] including 1000 variable size images classified in 10 classes of human being, horse, elephant, flower, bus, manmade thing and natural scenery. (iii) The Amsterdam Library of Object Image (ALOI) dataset [37] including 72,000 variable size images in 1000 classes. (iv) The VisTex dataset [33, 38] is composed of 1200 RGB colour images with $128 \times 128$ pixels, divided into 75 classes. Each class contains 16 images coming from their mother image, which is one of 75 images with $512 \times 512$ pixels chosen from the VisTex collection from the MIT Media Lab (v) MPEG-7 dataset [39] including 1400 variable size images in 70 classes. The MPEG-7 is a difficult dataset for colour-based image retrieval. Instead of objects, its images are composed of photos and sequences of video frames. Also, we use only the photos part of each image. All classes in each dataset are well balanced. We have tried to collect diverse datasets to evaluate the proposed method. As described in the previous sections, the features of all images in every datasets are extracted by using the IDWT feature, and then the best relevant images are shown by using the proposed image retrieval method via sparse representation.

### 4.3 Indexing results

In this section, we represent the results of the proposed method and compare it with several other methods. $P(0.5)$, $P(1)$ and ANMRR metrics are obtained to compare and evaluate the proposed method using different acceleration algorithms to solve (3) with level-5 of HMMD-HDWT [23], multiresolution colour and texture features (MCTF) [33], wavelet-based colour histogram (WBCH) [40], generalised Gaussian density and Kullback–Leibler distance (GGD & KLD) [41], wavelet-based features for colour texture classification (WFCTC) [42], colour layout descriptor (CLD) [43], dominant colour descriptor (DCD) [44], scalable colour descriptor (SCD) [45], Padua point (PP) and histogram intersection (HI) [6]. We apply the proposed method on all images of the datasets. Type, plane, the size of feature vector and the results of each metric are shown in Tables 1–5, for flower, Corel, ALOI, VisTex and MPEG-7 datasets, respectively. The best scores for each metric are highlighted.

In Table 1, The experimental results of the proposed method via SL0, DALM, PALM, homotopy, FISTA and AMP algorithms for HSI and CIE-$L^*a^*b^*$ planes, HMMD-HDWT, MCTF, WBCH, GGD & KLD, WFCTC, CLD, DCD, SCD, PI and HI methods for Flower dataset have been shown. In this experiment, the best performance rate is $P(1) = 92.62\%$ and ANMRR = 0.0823 for the proposed method via DALM algorithm for the $S$ plane (DALM-S).

**Table 1** Size of feature vector, $P(0.5)$, $P(1)$ and ANMRR of the retrieval methods for Flower dataset

| Type | Sparse algorithm | Plane | Size of feature vector | $P$ (0.5), % | $P(1)$, % | ANMRR |
|---|---|---|---|---|---|---|
| the proposed method | SL0 | $L^*$ | 64 | 35.86 | 38.08 | 0.5465 |
| | | $a^*$ | 64 | 28.95 | 29.75 | 0.6454 |
| | | $b^*$ | 64 | 34.29 | 35.31 | 0.5830 |
| | | $H$ | 64 | 35.96 | 34.76 | 0.6045 |
| | | $S$ | 64 | 34.35 | 36.19 | 0.5717 |
| | | $I$ | 64 | 35.18 | 36.81 | 0.5645 |
| | DALM | $L^*$ | 64 | 32.58 | 29.75 | 0.6656 |
| | | $a^*$ | 64 | 27.30 | 25.41 | 0.7126 |
| | | $b^*$ | 64 | 26.84 | 23.24 | 0.7237 |
| | | $H$ | 64 | 84.31 | 83.32 | 0.1586 |
| | | $S$ | 64 | 92.87 | 92.62 | 0.0823 |
| | | $I$ | 64 | 37.51 | 34.97 | 0.6181 |
| | PALM | $L^*$ | 64 | 16.15 | 12.79 | 0.8268 |
| | | $H$ | 64 | 16.67 | 12.60 | 0.8310 |
| | | $S$ | 64 | 16.51 | 12.93 | 0.8263 |
| | | $I$ | 64 | 16.58 | 12.99 | 0.8541 |
| | Homotopy | $L^*$ | 64 | 11.09 | 9.66 | 0.8431 |
| | | $H$ | 64 | 73.09 | 72.96 | 0.4626 |
| | | $S$ | 64 | 89.87 | 89.61 | 0.1025 |
| | | $I$ | 64 | 9.52 | 8.12 | 0.8858 |
| | FISTA | $L^*$ | 64 | 7.70 | 7.52 | 0.8892 |
| | | $H$ | 64 | 8.15 | 7.62 | 0.8945 |
| | | $S$ | 64 | 6.90 | 6.91 | 0.8974 |
| | | $I$ | 64 | 7.74 | 7.16 | 0.8980 |
| | AMP | $L^*$ | 64 | 5.88 | 5.88 | 0.9119 |
| | | $H$ | 64 | 5.88 | 5.88 | 0.9119 |
| | | $S$ | 64 | 5.88 | 5.88 | 0.9119 |
| | | $I$ | 64 | 5.88 | 5.88 | 0.9119 |
| HMMD-HDWT | | – | $3 \times 64$ | 20.95 | 17.07 | 0.7669 |
| MCTF | | | 92 | 79.19 | 78.98 | 0.3626 |
| WBCH | | | 512 | 65.33 | 63.23 | 0.5278 |
| GGD & KLD | | | 18 | 57.86 | 55.67 | 0.5695 |
| WFCTC | | | 384 | 35.93 | 30.34 | 0.6698 |
| CLD | | – | 12 | 17.99 | 15.05 | 0.7948 |
| DCD | | – | 32 | 18.16 | 14.30 | 0.8060 |
| SCD | | – | 11*121 | 16.78 | 13.26 | 0.8211 |
| PP30 | | – | 496 | 16.55 | 13.06 | 0.8342 |
| HI | | – | 1024 | 17.87 | 14.04 | 0.8231 |

**Table 2** Size of feature vector, $P(0.5)$, $P(1)$ and ANMRR of the retrieval methods for Corel dataset

| Type | Sparse algorithm | Plane | Size of feature vector | $P$ (0.5), % | $P(1)$, % | ANMRR |
|---|---|---|---|---|---|---|
| the proposed method | SL0 | $L^*$ | 64 | 57.78 | 58.60 | 0.3422 |
| | | $a^*$ | 64 | 39.07 | 39.97 | 0.5441 |
| | | $b^*$ | 64 | 49.60 | 51.31 | 0.4162 |
| | | $H$ | 64 | 50.48 | 51.28 | 0.4242 |
| | | $S$ | 64 | 50.79 | 52.58 | 0.4083 |
| | | $I$ | 64 | 58.44 | 58.72 | 0.3424 |
| | DALM | $L^*$ | 64 | 61.09 | 56.42 | 0.3921 |
| | | $a^*$ | 64 | 39.85 | 38.25 | 0.5723 |
| | | $b^*$ | 64 | 43.61 | 39.41 | 0.5485 |
| | | $H$ | 64 | 89.36 | 88.89 | 0.0883 |
| | | $S$ | 64 | 89.98 | 89.46 | 0.0824 |
| | | $I$ | 64 | 65.35 | 60.85 | 0.3520 |
| | PALM | $L^*$ | 64 | 21.06 | 17.23 | 0.7624 |
| | | $H$ | 64 | 24.67 | 19.46 | 0.7374 |
| | | $S$ | 64 | 30.18 | 23.08 | 0.6978 |
| | | $I$ | 64 | 33.36 | 25.85 | 0.6664 |
| | Homotopy | $L^*$ | 64 | 11.99 | 11.14 | 0.8384 |
| | | $H$ | 64 | 68.60 | 69.32 | 0.2038 |
| | | $S$ | 64 | 69.64 | 69.32 | 0.2038 |
| | | $I$ | 64 | 12.74 | 11.52 | 0.8345 |
| | FISTA | $L^*$ | 64 | 10.60 | 10.24 | 0.8500 |
| | | $H$ | 64 | 12.36 | 10.85 | 0.8479 |
| | | $S$ | 64 | 12.98 | 11.22 | 0.8470 |
| | | $I$ | 64 | 10.07 | 10.09 | 0.8490 |
| | AMP | $L^*$ | 64 | 10.00 | 10.00 | 0.8501 |
| | | $H$ | 64 | 10.00 | 10.00 | 0.8501 |
| | | $S$ | 64 | 10.00 | 10.00 | 0.8501 |
| | | $I$ | 64 | 10.00 | 10.00 | 0.8501 |
| HMMD-HDWT | | – | $3 \times 64$ | 41.45 | 35.44 | 0.5615 |
| MCTF | | | 92 | 75.43 | 71.34 | 0.1837 |
| WBCH | | | 512 | 80.35 | 76.25 | 0.1209 |
| GGD & KLD | | | 18 | 65.79 | 62.35 | 0.3314 |
| WFCTC | | | 384 | 47.21 | 40.23 | 0.5489 |
| CLD | | – | 12 | 57.79 | 41.82 | 0.5391 |
| DCD | | – | 32 | 48.64 | 36.37 | 0.5826 |
| SCD | | – | 11*121 | 43.63 | 33.65 | 0.6106 |
| PP30 | | – | 496 | 38.56 | 28.65 | 0.6531 |
| HI | | – | 1024 | 40.21 | 29.92 | 0.6324 |

**Table 3** Size of feature vector, $P(0.5)$, $P(1)$ and ANMRR of the retrieval methods for ALOI dataset

| Type | Sparse algorithm | Plane | Size of feature vector | $P(0.5)$, % | $P(1)$, % | ANMRR |
|---|---|---|---|---|---|---|
| | SL0 | $L^*$ | 64 | 55.11 | 51.01 | 0.4080 |
| | | $a^*$ | 64 | 39.99 | 38.58 | 0.5598 |
| | | $b^*$ | 64 | 48.84 | 45.51 | 0.4748 |
| | | $H$ | 64 | 36.93 | 34.62 | 0.6008 |
| | | $S$ | 64 | 58.76 | 52.03 | 0.4055 |
| | | $I$ | 64 | 53.20 | 49.26 | 0.4300 |
| | DALM | $L^*$ | 64 | 54.61 | 49.19 | 0.4348 |
| | | $a^*$ | 64 | 38.04 | 33.25 | 0.6070 |
| | | $b^*$ | 64 | 46.91 | 39.05 | 0.5346 |
| | | $H$ | 64 | 89.41 | 88.45 | 0.0992 |
| | | $S$ | 64 | 89.99 | 89.89 | 0.0905 |
| | | $I$ | 64 | 53.68 | 48.77 | 0.4451 |
| | PALM | $L^*$ | 64 | 20.64 | 19.05 | 0.7976 |
| | | $H$ | 64 | 24.31 | 23.43 | 0.7539 |
| | | $S$ | 64 | 25.06 | 24.71 | 0.7468 |
| | | $I$ | 64 | 20.33 | 18.24 | 0.8013 |
| | Homotopy | $L^*$ | 64 | 7.21 | 4.63 | 0.9410 |
| | | $H$ | 64 | 23.97 | 22.85 | 0.7640 |
| | | $S$ | 64 | 23.39 | 19.99 | 0.7879 |
| | | $I$ | 64 | 7.52 | 4.78 | 0.9390 |
| | FISTA | $L^*$ | 64 | 6.03 | 4.92 | 0.9384 |
| | | $H$ | 64 | 3.51 | 3.33 | 0.9506 |
| | | $S$ | 64 | 3.20 | 2.90 | 0.9613 |
| | | $I$ | 64 | 5.81 | 4.87 | 0.9388 |
| | AMP | $L^*$ | 64 | 1.56 | 1.56 | 0.9766 |
| | | $H$ | 64 | 1.56 | 1.56 | 0.9766 |
| | | $S$ | 64 | 1.56 | 1.56 | 0.9766 |
| | | $I$ | 64 | 1.56 | 1.56 | 0.9766 |
| HMMD-HDWT | | – | $3 \times 64$ | 82.29 | 66.68 | 0.2551 |
| MCTF | | | 92 | 81.60 | 63.33 | 0.3213 |
| WBCH | | | 512 | 61.59 | 50.11 | 0.4319 |
| GGD & KLD | | | 18 | 64.73 | 55.29 | 0.4027 |
| WFCTC | | | 384 | 43.03 | 30.45 | 0.6935 |
| CLD | | – | 12 | 19.49 | 17.56 | 0.8526 |
| DCD | | – | 32 | 50.21 | 41.87 | 0.6354 |
| SCD | | – | 11*121 | 53.21 | 43.55 | 0.5215 |
| PP30 | | – | 496 | 37.79 | 7.71 | 0.9023 |
| HI | | – | 1024 | 36.26 | 6.77 | 0.9186 |

The $S$ plane provides better performance than the $L^*$, $a^*$, $b^*$, $H$ and $I$ planes. In the following comparisons, we consider the $L^*$, $H$, $S$ and $I$ planes to compare the SL0, DALM, PALM, homotopy, FISTA and AMP algorithms in the proposed image retrieval has been described in Section 3.3, because the $a^*$ and $b^*$ have the lowest performance than other planes. In comparing performance, after the proposed method via DALM-S, the proposed method via homotopy algorithm for the $S$ plane with $P(1) = 89.61\%$ and ANMRR = 0.1025 and the proposed method via DALM algorithm for the $H$ plane with $P(1) = 83.32\%$ and ANMRR = 0.1576 provide better performance than other methods.

The results of the same experiment for Corel dataset are summarised in Table 2. Once again, it is observed that the proposed method via DALM-S provides the best performance rate that is $P(1) = 89.46\%$ and ANMRR = 0.0824. Corel dataset provides the same results as mentioned for Flower dataset. After the proposed method via DALM-S, the proposed methods via DALM algorithm for the $H$ plane with $P(1) = 88.89\%$ and ANMRR = 0.0883 provides better performance than other methods. Although, the size of feature vector for CLD, DCD, GGD & KLD methods and the proposed method is 12, 32, 18 and 64, respectively, however, the $P(1)$ of CLD, DCD and GGD & KLD are 47.64, 53.09 and 27.11% lower than DALM-S, respectively.

In Table 3, the results of the same experiment for ALOI dataset are shown. In this experiment, the best performance rate is $P(1) = 89.89\%$ and ANMRR = 0.0905 for the proposed method via DALM-S. Once again, the result of ALOI dataset provides the same results as mentioned for Flower and Corel datasets. In comparing performance, after the proposed method via DALM-S, the proposed method via DALM algorithm for the $H$ plane with $P(1) = 88.45\%$ and ANMRR = 0.0992 and HMMD-HDWT with $P(1) = 66.68\%$ and ANMRR = 0.2551 provide better performance than other methods.

The results of the same experiment for VisTex dataset are represented in Table 4. According to these results, the best performance rate is $P(1) = 85.07\%$ and ANMRR = 0.0840 for MCTF, however, the $P(1) = 83.59\%$ and ANMRR = 0.1001 of the proposed method via DALM-S are very close to MCTF method (only 1.8% less than MCTF). However, our proposed method significantly reduces the size of feature vector (0.69 times less than MCTF method), and this plays an important role in image retrieval and storage of feature datasets. Therefore, one may consider the proposed method as the best method. Once again, VisTex dataset provides the same results as mentioned for other datasets. In comparing performance, after the proposed method via DALM-S and MCTF method, GGD & KLD with $P(1) = 78.12\%$ and ANMRR = 0.2243 and the proposed method via homotopy algorithm for the $S$ plane with $P(1) = 74.51\%$ and ANMRR = 0.2533 provide better performance than other methods.

In Table 5, we show the results of the proposed method via SL0, DALM, PALM, homotopy, FISTA and AMP algorithms, HDWT, MCTF, WBCH, GGD & KLD, PP and HI for MPEG-7 dataset. Note that WFCTC, CLD, DCD, SCD methods need colour images with three dimensions (3D) including 256 values, whereas the images of MPEG-7 dataset include 1D with two values. Therefore, in Table 5, the results of the methods which use 1D images are shown. In this table, for the proposed method via DALM algorithm, $P(1) = 58.31\%$ and ANMRR = 0.4250 are better than other methods. $P(0.5)$ for HI method has better than other methods, however, the size of feature vector of HI method is more than the proposed method (>16 times).

Also, as observed in Tables 1–5, the proposed method via DALM algorithm for the $S$ plane provides even better performance than other sparse algorithms to solve (3), level-5 of HMMD-HDWT, MCTF, WBCH, GGD & KLD, WFCTC, CLD, DCD, SCD, PP-30 and HI. Therefore, the obtained results show that with respect to


*IET Comput. Vis.*, 2016, Vol. 10, Iss. 1, pp. 95–102

**Table 4** Size of feature vector, $P(0.5)$, $P(1)$ and ANMRR of the retrieval methods for VisTex dataset

| Type | Sparse algorithm | Plane | Size of feature vector | $P(0.5)$, % | $P(1)$, % | ANMRR |
|---|---|---|---|---|---|---|
| | SL0 | $L^*$ | 64 | 38.24 | 29.90 | 0.6632 |
| | | $a^*$ | 64 | 24.86 | 20.74 | 0.7706 |
| | | $b^*$ | 64 | 28.42 | 23.14 | 0.7413 |
| | | $H$ | 64 | 28.97 | 24.56 | 0.7340 |
| | | $S$ | 64 | 35.31 | 27.65 | 0.6919 |
| | | $I$ | 64 | 1.79 | 1.59 | 0.9767 |
| | DALM | $L^*$ | 64 | 42.75 | 39.50 | 0.5906 |
| | | $a^*$ | 64 | 22.77 | 20.79 | 0.7813 |
| | | $b^*$ | 64 | 24.77 | 21.45 | 0.7712 |
| | | $H$ | 64 | 51.99 | 50.83 | 0.4866 |
| | | $S$ | 64 | 84.55 | 83.59 | 0.1001 |
| | | $I$ | 64 | 43.18 | 39.98 | 0.5865 |
| | PALM | $L^*$ | 64 | 17.03 | 10.79 | 0.8639 |
| | | $H$ | 64 | 9.80 | 6.28 | 0.9181 |
| | | $S$ | 64 | 14.48 | 8.68 | 0.8891 |
| | | $I$ | 64 | 17.00 | 10.56 | 0.8657 |
| | Homotopy | $L^*$ | 64 | 14.60 | 8.87 | 0.8923 |
| | | $H$ | 64 | 52.32 | 52.29 | 0.4766 |
| | | $S$ | 64 | 75.18 | 74.51 | 0.2533 |
| | | $I$ | 64 | 14.31 | 8.51 | 0.8956 |
| | FISTA | $L^*$ | 64 | 1.90 | 1.65 | 0.9759 |
| | | $H$ | 64 | 2.38 | 1.79 | 0.9771 |
| | | $S$ | 64 | 1.95 | 1.61 | 0.9768 |
| | | $I$ | 64 | 1.79 | 1.59 | 0.9768 |
| | AMP | $L^*$ | 64 | 1.33 | 1.33 | 0.9801 |
| | | $H$ | 64 | 1.33 | 1.33 | 0.9801 |
| | | $S$ | 64 | 1.33 | 1.33 | 0.9801 |
| | | $I$ | 64 | 1.33 | 1.33 | 0.9801 |
| HMMD-HDWT | | – | $3\times64$ | 35.13 | 25.07 | 0.6767 |
| MCTF | | | 92 | 92.35 | 85.07 | 0.0840 |
| WBCH | | | 512 | 76.78 | 72.68 | 0.2654 |
| GGD & KLD | | | 18 | 80.22 | 78.12 | 0.2243 |
| WFCTC | | | 384 | 44.28 | 40.36 | 0.5739 |
| CLD | | – | 12 | 42.11 | 30.28 | 0.6273 |
| DCD | | – | 32 | 68.33 | 51.71 | 0.3914 |
| SCD | | – | 11*121 | 85.01 | 67.23 | 0.2586 |
| PP30 | | – | 496 | 34.43 | 24.22 | 0.7296 |
| HI | | – | 1024 | 29.25 | 21.76 | 0.7512 |

**Table 5** Size of feature vector, $P(0.5)$, $P(1)$ and ANMRR of the retrieval methods for MPEG-7 dataset

| Type | Sparse algorithm | Size of feature vector | $P$ (0.5), % | $P(1)$, % | ANMRR |
|---|---|---|---|---|---|
| the proposed method | SL0 | 64 | 52.97 | 43.83 | 0.5247 |
| | DALM | 64 | 59.41 | 58.31 | 0.4250 |
| | PALM | 64 | 40.87 | 30.59 | 0.6520 |
| | Homotopy | 64 | 16.15 | 9.50 | 0.8898 |
| | FISTA | 64 | 21.58 | 17.06 | 0.8074 |
| | AMP | 64 | 1.43 | 1.43 | 0.9787 |
| HDWT | | $3\times64$ | 70.41 | 53.90 | 0.4285 |
| MCTF | | 92 | 66.23 | 46.68 | 0.5012 |
| WBCH | | 512 | 39.69 | 30.21 | 0.6601 |
| GGD & KLD | | 18 | 40.23 | 35.40 | 0.7138 |
| PP30 | | 496 | 77.77 | 37.21 | 0.6821 |
| HI | | 1024 | 88.34 | 57.56 | 0.4432 |

performance rate and size of feature vectors, the proposed method scores extremely well. Thus, the proposed algorithm can be considered more powerful than the existing methods. Moreover, the proposed system not only reduces the size of the feature vector and storage space, but also improves the performance of the image retrieval system.

## 5 Conclusions

In this paper, we proposed a new scheme of image retrieval via sparse representation. The aim of the proposed algorithm is to provide a CBIR technique by using IDWT feature and sparse representation. HSI and CIE-$L^*a^*b^*$ colour spaces have been considered. The $P(0.5)$, $P(1)$ and ANMRR metrics of the proposed scheme and existing methods have been computed and compared. Flower, Corel, ALOI, VisTex and MPEG-7 datasets have been used to obtain the metrics. Experimental results for these datasets show that the proposed method via DALM algorithm for the $S$ plane yields higher retrieval accuracy than the other conventional methods with no greater feature vector size. In addition, the proposed method provided higher performance gain in both $P(1)$ and ANMRR over the other methods for five datasets. Moreover, the proposed system, both reduces the size of feature vector and storage space, and improves the performance of image retrieval. In the future, the proposed method can be used in the video retrieval systems, the image classification and so on.

## 6 References


1 Feng, J., Mingjing, L., Zhang, H.J., *et al.*: 'A unified framework for image retrieval using keyword and visual features', *IEEE Trans. Image Process.*, 2005, **14**, (7), pp. 979–989
2 Liu, Y., Zhang, D., Lu, G., *et al.*: 'A survey of content-based image retrieval with high-level semantics', *Pattern Recognit.*, 2007, **40**, (1), pp. 262–282
3 Starostenko, O., Chávez-Aragón, A., Burlak, G., *et al.*: 'A novel star field approach for shape indexing in CBIR system', *J. Eng. Lett.*, 2007, **1**, (2), pp. 10–21
4 Marakakis, A., Siolas, G., Galatsanos, N., *et al.*: 'Relevance feedback approach for image retrieval combining support vector machines and adapted Gaussian mixture models', *IET Image Process.*, 2011, **5**, (6), pp. 531–574
5 Li, F., Dai, Q., Xu, W., *et al.*: 'Multi-label neighborhood propagation for region-based image retrieval', *IEEE Trans. Multimedia*, 2008, **10**, (8), pp. 1592–1604
6 Montagna, R., Finlayson, G.D.: 'Padua point interpolation and $L^p$-Norm minimization in color-based image indexing and retrieval', *IET Image Process.*, 2012, **6**, (2), pp. 139–147
7 Liapis, S., Tziritas, G.: 'Color and texture image retrieval using chromaticity histograms and wavelet frames', *IEEE Trans. Multimedia*, 2004, **6**, (5), pp. 676–686
8 Lin, C.-H., Liu, C.-W., Chen, H.-Y.: 'Image retrieval and classification using adaptive local binary patterns based on texture features', *IET Image Process.*, 2012, **6**, (7), pp. 822–830
9 Nixon, M., Aguado, A.: 'Feature extraction & image processing' (Elsevier Ltd., 2002, 2nd edn., 2008)
10 Quellec, G., Lamard, M., Cazuguel, G., *et al.*: 'Fast wavelet-based image characterization for highly adaptive image retrieval', *IEEE Trans. Image Process.*, 2012, **21**, (4), pp. 1613–1623
11 Davenport, M.A., Duarte, M.F., Eldar, Y.C., *et al.*: 'Introduction to compressed sensing' (Department of Statistics at Stanford University, 2011), pp. 1–68
12 Elad, M.: 'Sparse and redundant representations' (Springer, New York, 2010)
13 Candes, E., Romberg, J., Tao, T.: 'Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information', *IEEE Trans. Inf. Theory*, 2006, **52**, (2), pp. 489–509
14 Donoho, D.: 'Compressed sensing', *IEEE Trans. Inf. Theory*, 2006, **52**, (4), pp. 1289–1306
15 Gemmeke, J.F., Hamme, H.V., Cranen, B., *et al.*: 'Compressive sensing for missing data imputation in noise robust speech recognition', *J. Sel. Top. Signal Process.*, 2010, **4**, (2), pp. 272–287
16 Lu, Y., Tan, C.L.: 'Information retrieval in document image databases', *IEEE Trans. Knowl. Data Eng.*, 2004, **16**, (11), pp. 1398–1410
17 Siddiquie, B., Feris, R., Davis, L.: 'Image ranking and retrieval based on multi-attribute queries'. CVPR, 2011, pp. 801–808
18 Zhang, S., Huang, Q., Hua, G., *et al.*: 'Building contextual visual vocabulary for large-scale image applications'. ACM Multimedia, 2010, pp. 25–29
19 Kang, L.W., Hsu, C.Y., Chen, H.W., *et al.*: 'Feature-based sparse representation for image similarity assessment', *IEEE Trans. Multimedia*, 2011, **13**, (5), pp. 1019–1030
20 Yang, X., Qian, X., Mei, T.: 'Learning salient visual word for scalable mobile image retrieval', *Pattern Recognit.*, 2015, **48**, (11), pp. 3093–3101
21 Srinivas, M., RamuNaidu, R., Sastry, C.S., *et al.*: 'Content based medical image retrieval using dictionary learning', *Neurocomputing*, 2015, **168**, (30), pp. 880–895
22 Liu, R., Chen, Y., Zhu, X., *et al.*: 'Image classification using label constrained sparse coding', *Multimedia Tools Appl.*, 2015, **74**, pp. 1–15
23 Farsi, H., Mohamadzadeh, S.: 'Colour and texture feature-based image retrieval by using Hadamard matrix in discrete wavelet transform', *IET Image Process.*, 2013, **7**, (3), pp. 212–218
24 Farsi, H., Mohamadzadeh, S.: 'Combining Hadamard matrix, discrete wavelet transform and DCT features based on PCA and KNN for image retrieval', *Majlesi J. Electr. Eng.*, 2013, **7**, (1), pp. 9–15
25 Mohamadzadeh, S., Farsi, H.: 'Image retrieval using color-texture features extracted from Gabor-Walsh wavelet pyramid', *J. Inf. Syst. Telecommun.*, 2014, **2**, (1), pp. 31–40
26 Mallat, S.A.: 'Wavelet tour of signal processing' (Academic press, San Diego, CA, 1999)
27 Amaldi, E., Kann, V.: 'On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems', *Theor. Comput. Sci.*, 1998, **209**, pp. 237–260
28 Donoho, D., Tsaig, Y.: 'Fast solution of l1-norm minimization problems when the solution may be sparse', preprint, 2006, pp. 4789–4812, http://www.stanford.edu/~tsaig/research.html

29 Chen, S., Donoho, D., Saunders, M.: 'Atomic decomposition by basis pursuit', *SIAM Rev.*, 2001, **43**, (1), pp. 129–159
30 Candès, E., Romberg, J., Tao, T.: 'Stable signal recovery from incomplete and inaccurate measurements', *Commun. Pure Appl. Math.*, 2006, **59**, (8), pp. 1207–1223
31 Yang, A.Y., Zhou, Z., Ganesh, A., *et al.*: 'Fast $l_1$-minimization algorithms for robust face recognition', *IEEE Trans. Image Processing*, 2013, **22**, (8), pp. 3234–3246
32 Mohimani, H., Babaie-Zadeh, M., Jutten, C.: 'A fast approach for over complete sparse decomposition based on smoothed norm', *IEEE Trans. Signal Process.*, 2009, **57**, pp. 289–301
33 Chun, Y.D., Kim, N.C., Jang, I.H.: 'Content-based image retrieval using multi-resolution color and texture features', *IEEE Trans. Multimedia*, 2008, **10**, (6), pp. 1073–1084
34 Veganzones, M.A., Graña, M.: 'A spectral/spatial CBIR system for hyper spectral images', *IEEE J. Sel. Top. Earth Obs. Remote Sens.*, 2012, **5**, pp. 488–500
35 Flower dataset, Available at: http://www.robots.ox.ac.uk/~vgg/data/flowers/17/
36 Coral dataset, last referred on June 2009, Available at: http://www.wang.ist.psu.edu/docs/related/
37 Geusebroek, J.M., Burghouts, G.J., Smeulders, A.W.M.: 'The Amsterdam library of object images', *Int. J. Comput. Vis.*, 2005, **61**, pp. 103–112
38 VisTex Texture Dataset. Available at: http://www.vismod.media.mit.edu/vismod/imagery/VisionTexture/
39 International organization for standardization, MPEG-7 overview 2004. Available at http://www.mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm, accessed 15 November 2011
40 Singha, M., Hemachandran, K.: 'Content based image retrieval using color and texture', *Signal Image Process., Int. J. (SIPIJ)*, 2012, **3**, (1), pp. 39–57
41 Minh, N.D., Vetterli, M.: 'Wavelet-based texture retrieval using generalized Gaussian density and Kullback–Leibler distance', *IEEE Trans. Image Process.*, 2002, **11**, (2), pp. 146–158
42 Hiremath, P.S., Shivashankar, S., Pujari, J.: 'Wavelet based features for color texture classification with application to CBIR', *IJCSNS Int. J. Comput. Sci. Netw. Secur.*, 2006, **6**, (9), pp. 124–133
43 Troncy, R., Huet, B., Schenk, S.: 'Feature extraction for multimedia analysis: 'Multimedia Semantics, Desktop Edition (XML): Metadata, Analysis and Interaction' (John Wiley & Sons Inc., New York, 2011, 1st edn.), pp. 36–54
44 Ibrahim, A., Zou'bi, A., Sahawneh, R., *et al.*: 'Fixed representative colours feature extraction algorithm for moving picture experts group-7 dominant color descriptor', *Int. J. Comput. Sci.*, 2009, **5**, (11), pp. 773–777
45 Jiang, Q., We, W., Zhang, H.: 'New researches about dominant color descriptor and graph edit distance', *Int. Conf. Intell. Human-Mach. Syst. Cybern. (IHMSC)*, 2011, **1**, pp. 50–52