Multimed Tools Appl (2014) 70:1651–1665
DOI 10.1007/s11042-012-1188-8

Secure and robust image hashing via compressive sensing

Rui Sun · Wenjun Zeng

Published online: 1 August 2012

Abstract Image hash functions find extensive applications in content authentication, database search. This paper develops a novel algorithm for generating a secure and robust image hash based on compressive sensing and Fourier-Mellin transform. Firstly, we incorporate Fourier-Mellin transform into our method to improve its performance under rotation, scale, transition attacks. Secondly, we exploit the property of dimension reduction inherent in compressive sensing for hash design. The statistic structure and sparse of the wavelet coefficients assure efficient compression in situation of including maximum the image features. The hashing method is computationally secure without additional randomization process. Such a combined approach is capable of tackling all types of attacks and thus can yield a better overall performance in multimedia identification. To demonstrate the superior performance of the proposed schemes, receiver operating characteristics analysis over a large image database is performed. Experimental results show that the proposed image hashing is robust to a wide range of distortions and attacks. When compared with the current state-of-the-art methods, the proposed method yields better identification performances under geometric attacks such as rotation attacks and brightness changes.

Keywords Compressive sensing · Fourier-Mellin transform · Image hashing · Image identification

1 Introduction

Digital media has profoundly changed our daily life during the past decades. However, the massive proliferation and extensive use of media data arising from its easy-to-copy nature

R. Sun (✉)
School of Computer and Information, Hefei University of Technology, Hefei 230009, People's Republic of China
e-mail: sunziyun@yahoo.com.cn

W. Zeng
Dept. of Computer Science, University of Missouri-Columbia, Columbia, MO 65211, USA
e-mail: zengw@missouri.edu

also pose new challenges to effectively manage such abundance of data (e.g., fast media searching, indexing) and protection of intellectual property of multimedia data. Among the various techniques proposed to address these challenges, image hashing has been proven to be an efficient tool because of its robustness and security.

An image hash is a compact and exclusive feature descriptor for a specific image. There are two important design criteria for image hash functions, namely, robustness and security [20, 27]. By robustness, we mean that when the same key is used, perceptually similar images should produce similar hashes. Here, the similarity of hashes is measured in terms of some distance metric, such as the Euclidean or Hamming distance. We consider two images to be similar if one image can be obtained from the other through a set of content-preserving manipulations. This set of manipulations includes moderate levels of additive noise, JPEG compression, filtering operations, geometric distortions, and watermark embedding. The security of image hash functions is introduced by incorporating a secret key in generating the hash. Without the knowledge of the key, the hash values should not be easily forged or estimated. Additionally, some design criteria for generic data hash also apply to image hash functions, namely, the one-way and collision-free properties. Although some generic data hash functions, such as MD5, satisfy these criteria [18], they are highly dependent on every bit (or pixel) of the input data rather than on the content. Hence, most of the them are not suitable for the emerging multimedia applications and the need for building robust and secure image hash is paramount.

A number of media-specific hash functions have been proposed for multimedia authentication. In addition to content authentication, multimedia hashes are used in content-based retrieval from databases [15] and image and video watermarking [6, 19]. It is worth mentioning that different applications may impose different requirements in a hashing design. For the purpose of image authentication, it is required that minor unmalicious modifications which do not alter the content of the data should preserve the authenticity of the data [29]. The robustness of image hash assures its capability to authenticate the content by ignoring the effect of minor unmalicious modifications on the original data. The desirable hash method can achieve not only tampering detection but also tampering localization. It increases the hash length for including the mount of information about original image. For the management of large image databases [14], image hashing allows efficient media indexing, identification, and retrieval by avoiding exhaustively searching through all the entries, thus reducing computational complexity of similarity measurements. The desirable hash method is computationally effective. The hash length is short for storage with the original data in the form of a lookup table. In this paper, we are particularly interested in image identification and indexing and explore how to design image hashing in this direction.

The procedure of deriving an image hash has two steps. The first step extracts a feature vector from the image, whereas the second stage compresses this feature vector to a final hash value. In the feature extraction step, the 2-D image is mapped to a 1-D feature vector. This feature vector must capture the perceptual qualities of the image. That is, two images that appear identical to the human visual system should have feature vectors that are close in some distance metric. Likewise, two images that are clearly distinct in appearance must have feature vectors that differ by a large distance. At the same time, using such features alone makes the system susceptible to forgery attacks, which may be carried out by an attacker that creates a new image with different visual content but with the same feature values. Thus, security mechanisms [25] must be combined into the feature extraction stage, e.g., by introducing some pseudorandom key in the hashing system.

2 Literature review

Various approaches have been proposed in literatures for constructing image hashes, although there is no universal hashing approach that is robust against all types of attacks. Swaminathan's hashing scheme [25] incorporates pseudo randomization into Fourier-Mellin transform to achieve better robustness to geometric operations. However, it suffers from some classical signal processing operations such as noising. It was also proposed in [21] to generate the hash by detecting invariant feature points, though the expensive searching and removal of feature points by malicious attacks such as cropping and blurring limit its performance in practice. Kozat proposed using low-rank matrix approximations obtained via the well-known singular value decomposition (SVD) for image hashing [12]. While the SVD-based hashing scheme exhibits good geometric attack robustness, it does so at the expense of significantly increasing misclassification. Monga introduced nonnegative matrix factorization (NMF) into their new hashing algorithm [22]. The major benefit of NMF hashing is the structure of the basis resulting from its nonnegative constraints, which lead to a parts-based representation. In contrast to the global representation obtained by SVD, the non-negativity constraints result in a basis of interesting local features [13]. Based on the results in [22], the NMF hashing possesses excellent robustness under a large class of perceptually insignificant attacks, while it significantly reduces misclassification for perceptually distinct images. It was shown to provide the best performance among NMF based hashing schemes investigated in [22], simply as NMF hashing in this paper. Other content-preserving features based on image statistics [9], wavelet transform [1, 7], DCT transform [10], Radon transform[24, 30], Fast Johnson-Lindenstrauss Transform [16, 17] have also contributed to the development of image hashing and enlightened some novel directions.

In this paper, we propose a hashing technique based on compressive sensing principles and Fourier-Mellin transform, which is robust legitimate content-preserving manipulations such as moderate affine transform, filtering, cropping and secure against malicious forgeries. According to the sampling theory and the Nyquist- Shannon sample theorem, exact reconstruction of a continuous-time signal from its samples is possible if the signal is band-limited and the sampling rate is more than twice the signal bandwidth. In recent years, a new theory Compressive Sensing (CS) also referred as Compressive Sensing or Compressive Sampling, has been proposed as a more efficient sampling scheme. The theoretical framework of CS was developed by Candes et al. [3] and Donoho [5]. The CS principle claims that a sparse signal can be recovered from a small number of random linear measurements. The CS theory provides a great reduction of sampling rate, power consumption and computational complexity to acquire and represent a sparse signal. In [26], an image authentication scheme based on CS and distributed source coding (DSC) was proposed, where the image hash is derived from the DSC-encoded quantized random projection coefficients of an image. To perform authentication, a DSC decoder decodes the received hash bits with the test image serving as the side information, where the authenticity depends on the success/fail of the DSC decoding. But the method has very long hash length and is computational complex so that it is limited in application. Kang presents a compressive sensing-based image hashing [11]. The method introduces visual information fidelity for hash comparison. Based on hash comparison, the distortion and visual quality of query image can be estimated. But the comparison process consumes so much time as to impact image retrieval efficiency. Our scheme is low complexity in hash extraction and comparison. It has short hash length that is suitable in image identification and indexing.

In the experiments, we study the performance of our algorithm under the attacks of rotation, scaling, shifting, luminance adjustment, filtering, additive Gaussian white noise. The results show that our algorithm achieves a good balance between robustness and discrimination. The experimental results of NMF hashing [22] and CS hashing [11] are compared with ours on the same dataset, our algorithm outperforms under most of attacks.

The rest of this paper is organized as follows. We first introduce the background and theoretic details about Fourier-Mellin transform and Compressive Sensing in Section 3. We propose the geometric invariant hashing methods by combining the Fourier-Mellin transform and CS to achieve better geometric robustness in Section 4. The analytical and experimental results are exhibited in Section 5 to demonstrate the superior performance of the proposed schemes. The conclusion and suggestions for future work are given in Section 6.

3 Theoretical background

In this section, we provide a brief summary of two topics that play a central role in the proposed method. In Section 3.1 we discuss Fourier-Mellin transform, which has been shown to be invariant to two-dimensional (2-D) affine transformations. In Section 3.2 we illustrate the foundations of compressive sensing, that is employed in order to efficient dimension reduction from a limited number of random projections.

3.1 Fourier-Mellin transform

Various translation, rotation and scale invariant methods such as integral transforms, moment invariants and Neural Network approaches have been proposed . These techniques provide good invariance theories but suffer from the presence of noise, computation complexity or accuracy problem [28]. Fourier-Mellin transform (FMT) performs well under noise and can be applied efficiently by using Fast Fourier Transform. FMT is translation invariant and represents rotation and scaling as translations along the corresponding axes in parameter space.

Consider an image $f_2(x,y)$ that is a rotated, scaled and translated replica of $f_1(x,y)$;

$$f_2(x,y) = f_1[\sigma(x\cos\alpha + y\sin\alpha) - x_0, \sigma(-x\sin\alpha + y\cos\alpha) - y_0] \tag{1}$$

where α is the rotation angle, σ the uniform scale factor, and x_0 and y_0 are translational offsets. The Fourier Transform of $f_1(x,y)$ and $f_2(x,y)$ are related by

$$F_2(u,v) = e^{-j\Phi_s(u,v)}\sigma^{-2}\left[F_1\left[\sigma^{-1}(u\cos\alpha + v\sin\alpha), \sigma^{-1}(-u\sin\alpha + v\cos\alpha)\right]\right] \tag{2}$$

where $\Phi_s(u,v)$ is the spectra phase of the image $f_2(x,y)$. This phase depends on the translation, scaling and rotation, but the spectral magnitude

$$|F_2(u,v)| = \sigma^{-2}\left|\left[F_1\left[\sigma^{-1}(u\cos\alpha + v\sin\alpha), \sigma^{-1}(-u\sin\alpha + v\cos\alpha)\right]\right]\right| \tag{3}$$

is translation invariant.

Equation (3) shows that a rotation of the image rotates the spectral magnitude by the same angle, and that a scaling by σ scales the spectral magnitude by σ^{-1}: Rotation and scaling can be decoupled by defining the spectral magnitudes of f_1 and f_2 in the polar coordinates (θ, r) ;

$$f_{2p}(\theta,r) = |F_2(r\cos\theta, r\sin\theta)|, f_{1p}(\theta,r) = |F_1(r\cos\theta, r\sin\theta)| \tag{4}$$

The (2) can be written using polar coordinates as

$$f_{2p}(\theta, r) = \sigma^{-2} f_{1p}(\theta - \alpha, r/\sigma) \tag{5}$$

Hence an image rotation shifts the function $f_{1p}(\theta, r)$ along the angular axis. A scaling is reduced to a scaling of the radial coordinate and to a magnification of the intensity by a constant factor σ^2: Scaling can be further reduced to a translation by using a logarithmic scale for the radial coordinate, thus

$$f_{2pl}(\theta, \lambda) = f_{2p}(\theta, r) = \sigma^{-2} f_{1pl}(\theta - \alpha, r - \eta) \tag{6}$$

Where $\lambda = \log(r)$ and $\eta = \log(\sigma)$. In this polar-logarithmic representation, both rotation and scaling are reduced to translation. By Fourier transforming the polar-logarithm representations, Eqs. (5) and (6),

$$F_{2pl}(\varsigma, \xi) = \sigma^{-2} e^{-j2\pi(\varsigma\eta + \xi\lambda)} F_{1pl}(\varsigma, \xi) \tag{7}$$

thereby rotation and scaling now appear as phase shifts. The Fourier magnitude of the two LPM mappings is related by

$$\left|F_{2pl}(\varsigma, \xi)\right| = |\sigma|^{-2} \left|F_{1pl}(\varsigma, \xi)\right| \tag{8}$$

Equation (8) demonstrates that the amplitude of Fourier–Mellin spectrum is scaled by $|\sigma|^{-2}$ caused by scaling transform, and is invariant to rotation and translation. $|\sigma|^{-2}$ will cause no problem at all if we use image resizing in advance, so the Fourier– Mellin transform is truly invariant to RST.

3.2 Compressive sensing

Compressive sensing theory asserts that it is possible to perfectly recover a signal from a limited number of incoherent nonadaptive linear measurements, provided that the signal can be represented by a small number of nonzero coefficients in some basis expansion.

Let $\mathbf{x} \in \mathbf{R}^{\mathbf{n}}$ denote the signal of interest and $\mathbf{y} \in \mathbf{R}^{\mathbf{m}}$, $m < n$, a number of linear random projections (measurements) obtained as $\mathbf{y} = \mathbf{Ax}$. The measurement matrix must be chosen in such a way that it satisfies a *restricted isometry property* (RIP) of order k [4], which says that all subsets of k columns taken from $\mathbf{A}$ are in fact nearly orthogonal or, equivalently, that linear measurements taken with $\mathbf{A}$ approximately preserve the Euclidean length of k sparse signals. The entries of $\mathbf{A} \in \mathbf{R}^{m \times n}$ the measurement matrix can be random samples from a given statistical distribution, e.g., Gaussian or Bernoulli. At first, let us assume that $\mathbf{x}$ is k sparse, i.e., there are exactly $k << n$ nonzero components. The goal is to reconstruct $\mathbf{x}$ given the measurements $\mathbf{y}$ and the knowledge that $\mathbf{x}$ is sparse. This can be formulated as the following optimization problem:

$$\min \|\mathbf{x}\|_0 \text{s.t.}\, \mathbf{y} = \mathbf{Ax} \tag{9}$$

where the ℓ_0 norm (represented as $\| \|_0$) simply counts the number of nonzeros entries of $\mathbf{x}$. Unfortunately, an exact solution to this problem requires an exhaustive search over all the possible k-sparse solutions and is, therefore, computationally intractable. Nonetheless, the recent results of compressive sensing have shown that, if $\mathbf{x}$ is sufficiently sparse, an approximation of it can be recovered by solving the following minimization problem:

$$\min \|\mathbf{x}\|_1 \text{s.t.}\, \mathbf{y} = \mathbf{Ax} \tag{10}$$

which can be immediately cast as a linear program. The solution of (10) is the same as (9) provided that the number of measurements satisfies$m \geq C \cdot k\log_2(n/k)$, where C is some small positive constant.

These results also hold when the signal is not sparse, but it has a sparse representation in some orthonormal basis. Let $\Phi \in \mathbf{R}^{n\times n}$ denote an orthonormal matrix, whose columns are the basis vectors. Let us assume that we can write $\mathbf{x} = \Phi\theta$, where θ is k sparse. Given the measurements $\mathbf{y} = \mathbf{Ax}$, the signal can be reconstructed by solving the following problem:

$$\min\|\theta\|_1 \text{s.t.}\, \mathbf{y} = \mathbf{A}\Phi\theta \tag{11}$$

For the case of noisy measurements, the signal model can be expressed as $\mathbf{y} = \mathbf{Ax} + \mathbf{z}$, where the noise amplitude is assumed to be bounded, i.e., $\|\mathbf{z}\|_2 \leq \varepsilon$. This situation occurs when the measurements are quantized. An approximation of the signal can be obtained by solving the following problem:

$$\min\|\theta\|_1 \text{s.t.}\, \|\mathbf{y} - \mathbf{A}\Phi\theta\| \leq \varepsilon \tag{12}$$

In this work, the wavelet transform is adopted to make the original signal become sparse. Recent research has demonstrated that if one exploits the structure in the transform coefficients characteristic of typical data or imagery, one often may significantly reduce the number of required CS measurements [2]. The structure associated with typical wavelet coefficients has been utilized in a statistical setting, building on recent research on Bayesian CS [8].

4 Proposed hashing algorithm

4.1 The performance of CS

Motivated by the hashing approaches based on SVD [12] and NMF [22], we believe that dimension reduction is a significantly important way to capture the essential features that are invariant under many image processing attacks. For CS, three benefits facilitate its application in hashing. First, CS is a random projection, enhancing the security of the hashing scheme. Second, CS's low distortion guarantees its robustness to most routine degradations and malicious attacks. The last one is its low computation cost when implemented in practice.

We will study the capability of CS to capture image features by comparison of SVD, NMF and CS. A sample case for the Lena image is illustrated in Fig. 1. Figure 1(a) shows the Original 128×128 Lena image. Approximations to the Lena image with similar compression ratio are shown in Fig. 1(b) (c) (d) by using SVD, NMF and CS, respectively. CS reconstruct method use the wavelet-based Bayesian CS [8]. It may be seen that perceptually Fig. 1(b) (c) (d) are of about the same quality.

We will test the stability and sensitivity of CS through L2 norm the difference among the lena image, the man image and the lena JPEG version (QF=10). The image use Haar wavelet transform firstly and then m=2,000 CS measurements. We divide the sampling signal in 50 blocks and compute the average of each block.

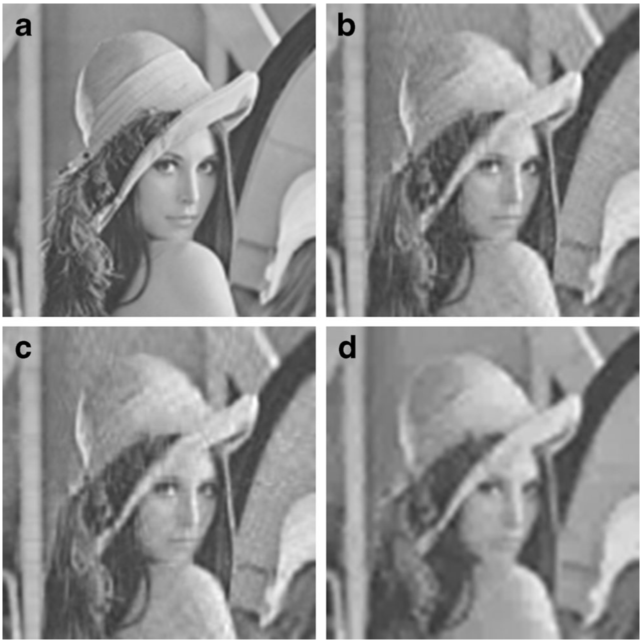
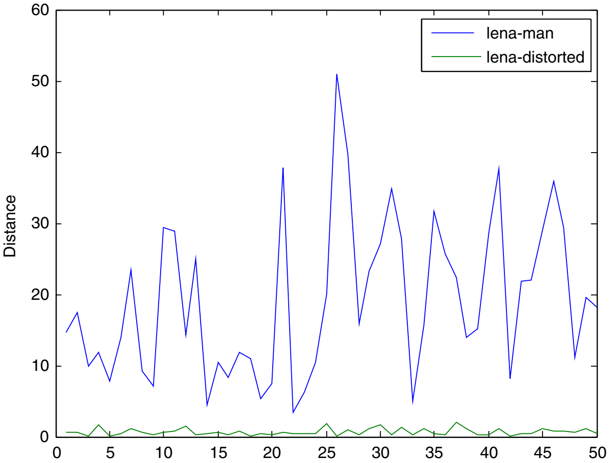


Fig. 1 Example of approximation of the Lena image via SVD, NMF and CS. The corresponding PSNR values: Fig. (**b**) is 30.2 dB, Fig. (**c**) is 29.5 dB and Fig. (**d**) is 30.5 dB. **a** Original Lena image. **b** Low-rank SVD approximation. **c** Low-rank NMF approximation. **d** CS approximation

Finally, we get the feature vector of length 50. Figure 2 shows the L2 norm of the component wise difference in features vector. The distinction between the different image versus the distorted version is easily made because of CS's inherent ability to capture local image features. It make our method has good classification ability.

Fig. 2 L2 norm of the difference between feature vector of the original Lena images, Man image and Lena distorted version after JPEG compression



4.2 FMCS hashing method

In this section, we proposed FMCS hashing method based on FM transform and CS principles. Figure 3 shows the framework of FMCS hashing method.

1) Image Preprocess: we let the original image X undergo a sequence of pre-processing, including image re-sizing, color space conversion, Since the luminance plane contains most of the geometric and visually significant information, for a color image we only consider the luminance component. Image resizing changes the image into a standard size $N \times N$ using bi-linear interpolation. This is done to ensure that the zgenerated image hash is scale invariant.
2) Appling FM Transform: the FMT could be divided into three steps, which result in the invariance to geometric attacks.

   a) Fourier Transform. It converts the translation of original image in spatial domain into the offset of angle in spectrum domain. The magnitude is translation invariant.
   b) Cartesian to Log-Polar Coordinates. It converts the scaling and rotation in Cartesian coordinates into the vertical and horizontal offsets in Log-Polar Coordinates.
   c) Mellin Transform. It is another Fourier transform in Log-Polar coordinates and converts the vertical and horizontal offsets into the offsets of angles in spectrum domain.

The final magnitude matrix $\mathbf{F} \in \mathbf{R}^{\mathbf{N} \times \mathbf{N}}$ is invariant to translation, rotation, and scaling.

3) Matrix Decimation: The magnitude matrix **F** is partitioned into blocks of size $B \times B$. The average of the component of each block is computed and stored in a vector $\mathbf{v} \in R^n$, where n denotes the number of blocks in the image, i.e., $n = N^2/B^2$.
4) Discrete Wavelet Transform: Appling wavelet transform to the vector **v** get wavelet coefficients feature vector $\mathbf{w} \in R^n$. The feature vector is sparse and satisfied to CS requirement. These papers demonstrate that one may achieve accurate CS inversions with substantially fewer projection measurements

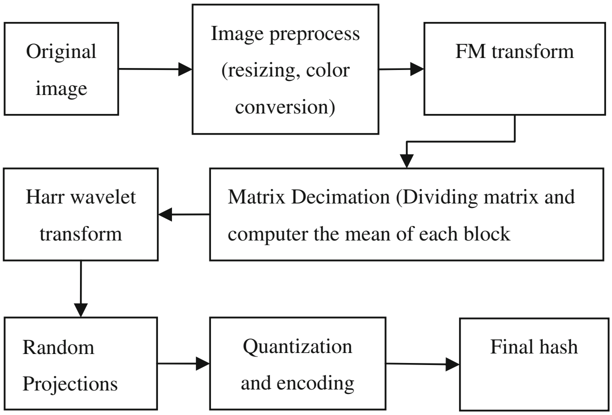

Original image
Image preprocess (resizing, color conversion)
FM transform
Matrix Decimation (Dividing matrix and computer the mean of each block
Harr wavelet transform
Random Projections
Quantization and encoding
Final hash

Fig. 3 The framework of FMCS hashing method

(smaller) if known properties of the structure of are exploited properly. The utility of exploiting prior knowledge about the structure of the wavelet coefficients is particularly valuable to represent feature vector with a small number of CS measurement.

5) Random Projections: A number of linear random projections $\mathbf{p} \in R^m, m < n$ is produced as $\mathbf{p} = \mathbf{Aw}$. The entries of the matrix $\mathbf{A} \in R^{m \times n}$ are sampled from a Gaussian distribution, generated using a random seed *S*, which will be sent as part of the hash to the user. The random seed *S* works as a sort of secret key to guarantee computational security against malicious attacks which may exploit the knowledge of the nullspace of the projection matrix **A** to break the system. The choice of the number of random projections depends on the expected sparsity and the structure of the vector **w**.
6) Post Processing: We quantize the resulting vector **p** and apply gray coding to obtain the binary hash sequence **h**. Furthermore, we can enhance the security using randomly permuted according to a permutation table generated using the key.

5 Analytical and experimental results

5.1 Performance evaluation

Let $S = \{s_i\}$ be the set of original images in the tested database and define a space $H(S) = \{H(s_i)\}$ as the set of corresponding hash vectors. We use Hamming distance as the performance metric to measure the robustness against content preserving manipulations and discriminating capability between two hash vectors, defined as

$$\mathrm{HD} = \sum_{i=1}^{n} |h_i(s_1) - h_i(s_2)| \tag{13}$$

where $H(s_i) = \{h_1(s_i), h_2(s_i), \cdots, h_n(s_i)\}$ means the corresponding hash vector with length n of the image s_i. Given a tested image s, we first calculate its hash $H(S)$ and then obtain its distances to each original image in the hash space $H(S)$. Intuitively, the query image s is identified as the $\widehat{i}$ th original image which yields the minimum corresponding distance, expressed as

$$\widehat{i} = \arg\min\left\{\|H(s) - H(s_i)\|_2\right\}, i = 1, \cdots, N \tag{14}$$

Except investigating robustness and identification accuracy, we also study the receiver operating characteristics (ROC) curve to visualize the performance of different hashing approaches, including NMF hashing, CS hashing and our method. The ROC curve depicts the relative tradeoffs between benefits and cost of the identification and is an effective way to compare the performances of different hashing approaches. To obtain ROC curves to analyze the hashing algorithms, we may define the probability of true identification P_T and probability of false alarm P_F as

$$P_T = \Pr\left(\|H(I) - H(I_{simi})\|_2 < T\right) \tag{15}$$

$$P_F = \Pr\left(\left\|H(I) - H\left(I_{diff}\right)\right\|_2 < T\right) \tag{16}$$

where T is the identification threshold. The images I and I_{diff} are two distinct original images and the image I_{simi} is manipulated versions of the image I. Ideally, we hope that the hashes of the original image I its manipulated version I_{simi} should be similar and thus be identified accurately, while the distinct images I and I_{diff} should have different hashes. In other words, given a certain threshold T, an efficient hashing should provide a higher P_T with a lower P_F simultaneously. Consequently, when we obtain all the distances between manipulated images and original images, we could generate a ROC curve by sweeping the threshold T from the minimum value to the maximum value, and further compare the performances of different hashing approaches.

5.2 Identification results

In order to evaluate the performance of the proposed new hashing algorithms, we test our method on a database of 100 000 images. In this database, there are 1,000 original color nature images, which are mainly selected from the ten sets of categories in the content-based image retrieval database of the University of Washington [23].we generate 99 similar versions by manipulating the original image according to a set of content preserving operations (CPOs) listed in Table 1. All the operations are implemented using Matlab.

We firstly test identification accuracy for the standard test images such as Baboon, Lena, and Peppers. Here we will measure the proposed hashing on the new database. Ideally, it is robust to all routine degradations and malicious attacks, no matter what content-preserving manipulation is done, the image with any distortion should still be correctly classified into the corresponding original image.

Following the algorithms designed in Section 4, we test our hashing with the parameters chosen as as summarized in Table 2. Since the NMF-NMF-SQ hashing has been shown to outperform the SVD-SVD and PR-SQ hashing algorithms having the best known robustness properties in the existing literature. The CS hashing exploits CS mechanism too. We Choose NMF-NMF-SQ hashing and CS hashing for

Table 1 Types and parameters of CPOs

Operations	Parameters	Number
Gaussian noise	Sigma:0–0.1	10
Salt & pepper noise	Sigma:0–0.1	10
Speckle noise	Sigma:0–0.1	10
Gaussian blurring	Size:3–21, sigma:5	10
Circular blurring	Radius:1–10	10
Motion blurring	Len:5–15	9
Rotation	Degree:5–45	9
Cropping	5 %–35 %	6
Scaling	25 %–200 %	5
JPEG comprssion	QF:5–50	10
Gamma correction	Gamma:0.75–1.25	10

Table 2 Parameter setting

Parameter	Value
Standard size	N=256
Block size	B=8
Level of wavelet tansform	2
Length of projection	m=50
Length of gray code	6
Length of hash vector	300

comparing the performance of our proposed hashing algorithm. For the NMF approach, the parameters are set as m=64, p=10, $r1$=2, $r2$=1, and M=40 according to [22]. It is worth mentioning that, to be consistent with the FCMS approach, we chose the same size of subimages and length of hash vector in NMF hashing. We first examine the identification accuracy of both hashing algorithms under different attacks, and the identification results are shown in Table 3. It is clearly noted that the proposed hashing consistently yields a higher identification accuracy than that of NMF hashing and CS hashing under different types of tested manipulations and attacks.

5.3 ROC analysis

We then present a statistical comparison of the proposed FCMS and NMF hashing algorithms by studying the corresponding ROC curves. We generate the overall ROC curves for all types of tested manipulations when applying different hashing schemes, and the resulting ROC curves are shown in Fig. 4. From Fig. 4, one major observation is that the proposed FCMS hashing outperforms NMF hashing and CS hashing in various CPOs.

Table 3 Identification accuracy for FMCS,NMF and CS hashing

Operations	NMF	CS	FMCS
Gaussian noise	58.24 %	49.16 %	65.62 %
Salt & pepper noise	78.53 %	76.73 %	94.23 %
Speckle noise	76.81 %	77.40 %	96.28 %
Gaussian blurring	97.51 %	88.33 %	99.55 %
Circular blurring	97.12 %	95.72 %	99.26 %
Motion blurring	98.23 %	95.57 %	99.82 %
Rotation	17.16 %	47.38 %	55.34 %
Cropping	15.28 %	37.29 %	87.34 %
Scaling	98.78 %	99.38 %	100 %
JPEG comprssion	98.70 %	96.82 %	100 %
Gamma correction	7.34 %	23.57 %	67.45 %

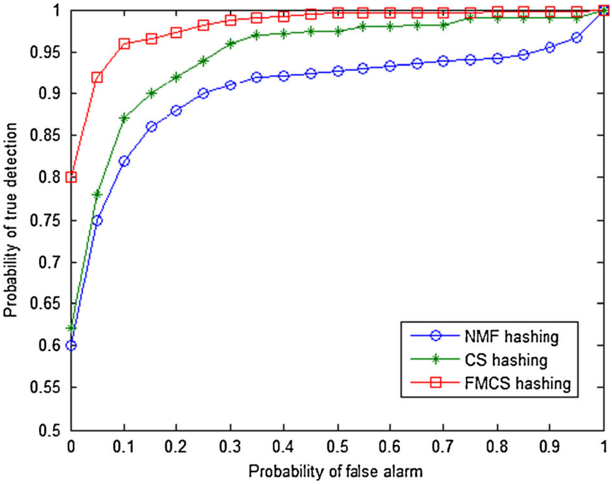


Fig. 4 The overall ROC curves of NMF hashing, CS hashing and FCMS hashing under all types of tested operations

5.4 Security analysis

Collision occurs if the Hamming distance between two hash values of visually distinct images is sufficiently small, say, less than a given threshold *T*. In order to find the collision probability, we generated hashes of 1,000 different color images from the image database of Washington University. Assume the Hamming distances follow one of the common distributions, i.e., Poisson, lognormal, and normal distributions. We apply chi-square test to determine which is the closest. Parameters of these distributions are obtained based on the maximal likelihood estimation, and the probability density functions (PDF) are computed at the values ranging from 0 to the hash length *L*. Figure 5 gives comparison between the actual distribution and the ideal normal distribution. We can identify the distribution of Hamming distances as the normal

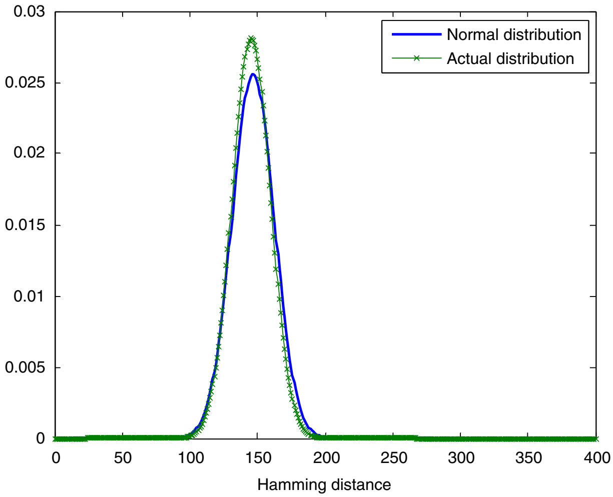


Fig. 5 Distribution of Hamming distances between different image hashes

distribution with its mean and standard deviation being $\mu = 146.8$ and $\sigma = 15.7$, respectively. Given a threshold *T*, the collision probability can be obtained as

$$P(HD \leq T) = \frac{1}{\sqrt{2\pi}\sigma} \int_0^T e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{2} erfc\left(-\frac{T-\mu}{\sqrt{2}\sigma}\right) \tag{27}$$

Then, a very low collision probability 3.52×10^{-14} is achieved when T=30.

5.5 CPU time cost

Compared with NMF hashing and CS hashing, which use prefixed regions of interest determined by a secret key for feature extraction and CS random projection and reconstruction, the major and additional computation cost of the proposed FMCS hashing lies in the FM and wavelet transform. Therefore, the computation cost of the proposed FMCS hashing is higher than NMF hashing and CS hashing. As an example, we test these approaches on 50 images using a desktop computer with CPU 3.0 G and 2 G RAM and report the average computational time in Table 4. After the hash is formed offline, the FMCS hashing has faster hash compassion speed than CS hashing. It is suitable to image identification and indexing.

6 Conclusion

In this paper, we develop new image hashing algorithms using compressive sensing principle. We have incorporated Fourier-Mellin transform to our hashing against rotation, scaling, and transition attacks and exploited the property of dimension reduction inherent in compressive sensing for hash design. The advantage of CS, relative to conventional compress approaches, is that the number of (projection) Measurements may be significantly smaller than the number of measurements in traditional sampling methods. The statistic structure and sparse of the wavelet coefficients assure efficient compression in situation of including maximum the image features. Based on our experimental results, it is noted that the FMCS-based hashing is robust to a large class of routine distortions and geometric attacks. Compared with the NMF hashing and the CS hashing, the proposed FMCS hashing can achieve comparable, sometimes better, performances than that of NMF, while requiring less computational cost. The random projection and low distortion properties of FMCS make it more suitable for hashing in practice than the NMF approach.

Furthermore, we plan to explore the CS-based hashing in image authentication application. Most of hash-based image authentication methods don't localize the tampering area. We will exploit inversion reconstruction of CS procedure to obtain the estimate of the image tampering. Another concern that is of great importance in practice but is rarely discussed in the context of image hashing is automation. Automatic estimation/choice of design parameters removes the subjectivity from the design procedure and can yield better performances. We will study some optimization algorithms for automatic estimation of parameters of the FMCS hashing using could improve the identification performance.

Table 4 The average CPU times of NMF,CS and FMCS hashing

Computational cost	NMF	CS	FMCS
Time(s)	0.92	0.82	2.31

Acknowledgements This work was supported by the Natural Science Foundation of China (61001201). The authors appreciate the anonymous reviewers for their constructive comments.

References

1. Ababneh S, Ansari R, Khokhar A (2008) Scalable multimedia-content integrity verification with robust hashing. in Proceedings of IEEE International Conference on Electro/Information Technology, 263–266
2. Blumensath T, Davies EM (2009) Sampling theorems for signals from the union of linear subspaces. IEEE Trans Inf Theory 55(4)
3. Candès E, Romberg J, Tao T (2006) Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. IEEE Trans Inf Theory 52:489–509
4. Candés E, Wakin BM (2008) An introduction to compressive sampling: a sensing/sampling paradigm that goes against the common knowledge in data acquisition. IEEE Signal Process Mag 25(2):21–30
5. Donoho D (2006) Compressed sensing. IEEE Trans Inf Theory 52:1289–1306
6. Fridrich J, Goljan M (2000) Robust hash functions for digital watermarking. in Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC '00),178–183
7. Gerold L, Andreas U (2008) Key-dependent JPEG2000-based robust hashing for secure image authentication. EURASIP J Inf Secur 8(1):1–19
8. He L, Carin L (2009) Exploiting structure in wavelet-based Bayesian compressive sensing. IEEE Trans Signal Process 57(9):3488–3497
9. Kailasanathan C, Naini SR (2001) Image authentication surviving acceptable modifications using statistical measures and k-mean segmentation. Proc. IEEE-EURASIP Work. Nonlinear Sig. Image
10. Kailasanathan C, Naini S R, Ogunbona P (2003) Compression tolerant DCT based image hash. in: Proceedings of International Conference on Distributed Computing Systems, 562–567
11. Kang WL, Lu SC, Hsu YC (2009) Compressive sensing-based image hashing, in Proc. of 2009 IEEE Int. Conf on Image Processing, Cairo, Egypt, November 2009: 1285–1288
12. Kozat S, Venkatesan R, Mihcak KM (2004) Robust perceptual image hashing via matrix invariants. Proc IEEE Intl Conf Image Process 5:3443–3446
13. Lee D, Seung H (2001) Algorithms for non-negative matrix factorization. Adv Neural Inform Process Syst 13:556–562
14. Lew M, Sebe N, Djeraba C, Jain R (2006) Content based multimedia information retrieval: state of the art and challenges. ACM Trans Multimed Comput Commun Appl 2(1):1–19
15. Lin S, Ozsu TM, Oria V, and Ng R (2001) An extendible hash for multi-precision similarity querying of image databases. in Proc. 27th Very Large Data Bases (VLDB) Conference
16. Lv XD, Wang Jane Z (2008) Fast Johnson-Lindenstrauss Transform for Robust and Secure Image Hashing. Proc. of IEEE MMSP, 725–729
17. Lv XD, Wang Jane Z (2009) An Extended Image Hashing Concept: Content-Based Fingerprinting Using FJLT. EURASIP Journal on Information Security
18. Menezes A, Oorschot V, Vanstone S (1998) Handbook of applied cryptography Boca Raton. CRC, FL
19. Mihcak KM, Venkatesan R (2001) Video watermarking using image hashing. Microsoft Research Tech. Report
20. Monga V (2005) Perceptually based methods for robust image hashing, Dissertation, University of Texas
21. Monga V, Evans LB (2006) Perceptual image hashing via feature points: performance evaluation and tradeoffs. IEEE Trans Image Process 15(11):3452–3465
22. Monga V, Mihcak KM (2007) Robust and secure image hashing via non-negative matrix factorizations. IEEE Trans Inform Forensic Secur 2(3):376–390
23. Object and Concept Recognition for Content-Based Image Retrieval, University of Washington. http://www.cs.washington.edu/research/imagedatabase/
24. Seo SJ, Haitsma J, Kalker T, Yoo DC (2004) A robust image fingerprinting system using the radon transform. Signal Process Image Comm 19(4):325–339
25. Swaminathan A, Mao Y, Wu M (2006) Robust and secure image hashing. IEEE Trans Inform Forensic Secur 1(2):215–230
26. Tagliasacchi M, Valenzise G, Tubaro S (2009) Hash-based identification of sparse image tampering. IEEE Trans Image Process 18(11):2491–2504
27. Venkatesan R, Koon MS, Jakubowski HM, Moulin P (2000) Robust image hashing. in Proceedings of the International Conference on Image Processing (ICIP '00), vol. 3: 664–666
28. Wood J (1996) Invariant pattern recognition: a review. Pattern Recogn 29(1):1–17

29. Wu WC (2002) On the design of content-based multimedia authentication systems. IEEE Trans Multimed 4(3):385–393
30. Wu D, Zhou XB, Niu XM (2009) A novel image hash algorithm resistant to print-scan. Signal process 89:2415–2424

Rui Sun is currently an associate professor of Hefei University of Technology, China. He received BS degree in Central South University of China in 1998, MS degree in Harbin Engineering University of China in 2000, Ph. D degree in Huazhong university of Science and Technology of China in 2003. He worked as senior software engineer in TCL mobile communication company, China from 2003 to 2005. He was a visiting scholar in Computer Science department, University of Missouri-Columbia, USA from 2010 to 2011. He is senior member of Chinese institution of electronics. His research interests include multimedia security, computer vision, and multimedia mobile networking.

Wenjun Zeng received the B.E., M.S., and Ph.D. degrees, all in electrical engineering, from Tsinghua University, Beijing, China, the University of Notre Dame, Notre Dame, IN, and Princeton University, Princeton, NJ, in 1990, 1993,and 1997, respectively. He is currently an Professor with the Department of Computer Science, University of Missouri, Columbia. Prior to joining the University of Missouri in 2003, he was with Packet Video Corporation, San Diego, CA, Sharp Laboratories of America, Camas, WA, Bell Laboratories, Murray Hill, NJ, and Matsushita Information Technology Laboratory, Panasonic Technologies, Princeton. He has consulted with Microsoft Research Asia, Beijing, Huawei Technologies, Shenzhen, China, and a couple of startup companies. From 1998 to 2002, he was an active contributor to the MPEG4 Intellectual Property Management and Protection Standard and the JPEG 2000 Image Coding Standard, where four of his proposals were adopted. He was elected an IEEE fellow his outstanding contributions to multimedia communication and security in 2011. His current research interests include multimedia communications and networking, distributed source and video coding, and content and network security.