# ROBUST IMAGE HASHING VIA NON-NEGATIVE MATRIX FACTORIZATIONS

*Vishal Monga*

Xerox Innovation Group
El Segundo, CA, 90245 USA
Email: *vishal.monga@xeroxlabs.com*

*M. Kıvanç Mıhçak*

Cryptography & Anti-Piracy Group,
Microsoft Research, Redmond, WA 98052 USA
Email: *kivancm@microsoft.com*

## ABSTRACT

In this paper, we propose the use of non-negative matrix factorization (NMF) for robust image hashing. In particular, we view images as matrices and the goal of hashing as a randomized dimensionality reduction that retains the essence of the original image matrix while preventing against intentional attacks of guessing and forgery. Our work is motivated by the fact that standard-rank reduction techniques such as the QR, and Singular Value Decomposition (SVD), produce low rank bases which do not respect the structure (i.e. non-negativity for images) of the original data. We observe that NMFs have two very desirable properties for secure image hashing applications: 1.) The additivity property resulting from the non-negativity constraints results in bases that capture *local characteristics* of the image, thereby significantly reducing misclassification, and 2.) the effect of geometric attacks on images in the spatial domain manifests (approximately) as independent identically distributed noise on NMF vectors, allowing design of detectors that are both computationally simple and at the same time optimal in the sense of minimizing error probabilities. ROC (receiver operating characteristics) analysis over a large image database reveals that the proposed algorithms significantly outperform existing approaches for robust image hashing.

## 1. INTRODUCTION

An image hash function maps an image to a short binary string based on the image's appearance to the human eye. In particular, a perceptual image hash function should have the property that two images that look the same to the human eye map to the same hash value, even if the images have different digital representations; e.g., being separated by a large distance in mean squared error. This differentiates a perceptual hash from traditional cryptographic hashes, such as SHA-1 and MD-5 [1]. SHA-1 and MD-5 hashes are extremely sensitive to the input data, i.e., a one bit change in the input changes the output dramatically.

An immediately obvious application for a *perceptual* image hash is identification/search of images in large databases. Several other applications have been identified recently in content authentication, watermarking [2], and anti-piracy search. Unlike traditional search, these scenarios are adversarial, and require the hash to be a randomized digest.

The underlying techniques for constructing image hashes can roughly be classified into methods based on image statistics [3], [4] [5], relations [6], [7], preservation of coarse image representation [8], [9], [10], and low-level image feature extraction [11], [12]. A common shortcoming of the schemes in [3] - [11] (excluding [10]) is poor robustness to geometric attacks, particularly lossy ones like cropping. While the singular value decomposition (SVD) based hashing scheme in [10] exhibits good geometric attack robustness, it does so at the expense of significantly increasing misclassification, i.e., different images mapping to similar hash values.

In this paper, we develop robust image hashing algorithms based on a recently-proposed dimensionality reduction technique by Lee *et al.* called *non-negative matrix factorization* (NMF) [13]. Our work is motivated partly by the SVD-based image hashing scheme proposed recently in [10]. NMF is distinguished from traditional matrix approximation methods by its use of non-negativity constraints. These constraints lead to a parts-based representation because they allow only additive, not subtractive, combinations. This is in contrast with SVD, which learns holistic and not parts-based representations. An immediate consequence of this property with respect to hashing, is far less misclassification (perceptually distinct images mapping to similar hash values) when NMF (instead of SVD) is employed for dimensionality reduction. In addition, we observe that geometric distortions on images result in approximately additive and independent, identically distributed noise on NMF vectors. We exploit this property to obtain pseudo-random linear statistics of NMF vectors, which significantly enhances hash robustness while allowing the hash to be of an acceptably small length for most practical applications.

The rest of this paper is organized as follows. Section 2 provides background on non-negative matrix factorizations. Section 3 describes our proposed hash algorithms. We present two variants: the NMF-NMF and NMF-NMF-SQ hashing schemes. Experimental results in the form of receiver operating characteristic (ROC) curves are presented in Section 4. Section 5 summarizes the contribution of the paper and provides directions for future work.

Work was carried out when V. Monga was with Microsoft Research.

## 2. NMF : BACKGROUND AND THEORY

Given a non-negative matrix $V$ of size $m \times n$, NMF algorithms seek to find non-negative matrix factors $W$ and $H$ such that

$$V \approx W \cdot H, \quad \text{where} \quad W \in \mathcal{R}^{m \times r} \text{ and } H \in \mathcal{R}^{r \times n}.$$

Equivalently, we have

$$v_j \approx W \cdot h_j, \quad v_j \in \mathcal{R}^m \quad h_j \in \mathcal{R}^r, \quad 1 \leq j \leq n,$$

where $\{v_j\}_{j=1}^n$ and $\{h_j\}_{j=1}^n$ denote columns of $V$ and $H$, respectively. For the class of non-sparse matrices, this factorization provides a reduction in storage whenever the number of vectors $r$, in the basis $W$ is chosen such that $r < \frac{mn}{m+n}$. In practice, $r$ is usually chosen such that $r \ll min(m, n)$.

### 2.1. NMF Cost Functions and Algorithms

To find an approximate factorization, $V \approx WH$, we first need to define cost functions that quantify the quality of the approximation. In the NMF literature, two popular cost functions have been studied. First is the classical Euclidean distance or Frobenius norm, given by

$$\Theta_E(W,H) \equiv \left( \sum_{j=1}^{n} \| v_j - Wh_j \|_2^2 \right)^{1/2} = \| V - WH \|_F \tag{1}$$

Another measure commonly used in practice is,

$$\Theta_D(V \| WH) \equiv \sum_{i=1}^{m} \sum_{j=1}^{n} \left( V_{ij} \ log \frac{V_{ij}}{[WH]_{ij}} - V_{ij} + [WH]_{ij} \right), \tag{2}$$

which is known as the generalized Kullback-Leibler (KL) divergence. It reduces to the standard KL divergence, or relative entropy, when $\sum_{ij} V_{ij} = \sum_{ij} [WH]_{ij} = 1$, so that the matrices can be regarded as normalized probability distributions.

The lack of convexity of the aforementioned costs in both factors $W$ and $H$ implies that it is unrealistic to expect a computationally efficient algorithm to find a global minimum. Using an approach analogous to Expectation-Maximization (EM) algorithm, Lee and Seung [13] developed algorithms commonly used to obtain NMFs.

### 2.2. Known NMF Results and Properties

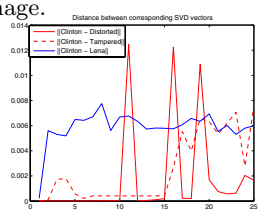
Using this representation, we see that the left factor $W$ contains a basis used for the linear approximation of $V$. The right factor $H$ is a coefficient matrix used to *add up* combinations of the basis vectors in $W$. The non-negativity constraint on $W$ allows us to visualize the basis columns in the same manner as columns in the original data matrix. This is the first benefit of NMF versus alternative factorizations like the SVD, where the basis vectors contain negative components that prevent similar visualizations.

The second, and very desirable, benefit of NMF is the structure of the resulting basis. For applications typically used, this basis will be $r$ conceptual or representative entities stored in the columns of $W$ that can sum up to (approximately) reconstruct the original data matrix. In particular,

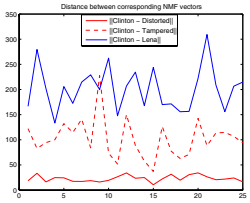
(a) Original *clinton* image

(b) Locally tampered version

Figure 1: Original and tampered versions of the *clinton* image.




(a) SVD




(b) NMF

Figure 2: L2 norm of the difference between corresponding SVD and NMF vectors of: *clinton* and *lena* images, *clinton* image and its attacked/distorted version with JPEG compression QF = 10, *clinton* image and its tampered version. Horizontal axis denotes the SVD/NMF vector index.

the non-negativity constraint means we obtain a basis containing interesting local features.

In the past, NMF has been applied with promising success to image classification [14], [15], text characterization [16], and even to de-construct music tones [17]. In this work, motivated by the ability of NMF to capture meaningful local components of the image matrix, we apply it to the robust image hashing problem. The next Section details the construction of our hash algorithms.

## 3. PROPOSED HASH ALGORITHMS

We first present some experimental observations that further motivate the use of NMF for image hashing, and also provide insight for the construction of our hash algorithms. Fig. 1 (a) shows the original 512 × 512 image of a former US President and the First Lady (*clinton* image). Fig. 1 (b) then shows a tampered version of the image in Fig. 1 (a) in which a malicious change is made to the First Lady's face. We obtain a rank 25 factorization of each of these images via both SVD and NMF. We also obtain rank 25 decompositions via SVD and NMF for the 512 × 512 *lena* image. This results in 25 left and right vectors (of length 512) for both SVD and NMF. Fig. 2(a) plots the $L2$ norm of the difference between corresponding left singular vectors of: 1.)*clinton* and *lena* images, 2.)*clinton* image and its compressed version (JPEG QF = 10), and 3.) *clinton* image and its tampered version in Fig. 1 (c). The same is repeated for NMF in Fig. 2 (b). The ability of NMF to capture robust local components is now readily apparent. In particular, it may be seen that SVD views JPEG compression (an allowable distortion), and local tampering (content change) as approximately the same. This distinction how-

ever, is clearly made in the case of NMF. Quantitatively, this is because the orthogonality constraints in SVD means that the singular vectors provide a holistic basis and hence ignore local changes.

### 3.1. NMF-NMF Hashing

A secret key is used as the seed of a cryptographically secure pseudo-random number generator that is employed for randomizing all steps below.

1. Given an image $I$, pseudo-randomly select $p$ subimages $A_i \in R^{k \times k}$, $1 \leq i \leq p$.
2. Obtain a rank $r_1$ NMF from each sub-image ($r_1 \ll k$)

$$A_i \approx W_i \cdot F_i^T,\ 1 \leq i \leq p. \tag{3}$$

   where $\{W_i\}$ and $\{F_i\}$ are all $k \times r_1$. This results in $2p$ NMF matrices of size $k \times r_1$ each.
3. Pseudo-randomly arrange these matrices to obtain a secondary image $J$ of size $k \times 2pr_1$.
4. Re-apply NMF to obtain a rank $r_2$ representation of $J$, $r_2 \ll min(k, 2pr_1)$

$$J \approx W \cdot H \tag{4}$$

   where $W$ is $k \times r_2$ and H is $r_2 \times 2pr_1$.
5. The concatenation of columns of $W$ and rows of $H$ gives the hash vector $h_K^{NMF-NMF}(I)$.

We employ NMF on pseudo-random (PR) image regions for security reasons. A two-stage cascade algorithm was constructed because we observe experimentally that it serves to significantly enhance hash robustness.

### 3.2. NMF-NMF-SQ Hashing

1. Obtain the NMF-NMF hash vector $h_K^{NMF-NMF}(I)$ as in the previous sub-section. Let $N$ be the length of hash vector.
2. Generate pseudo-random weight vectors $\{\mathbf{t}_i\}_{i=1}^M$ (with $M \ll N$) such that each $\mathbf{t}_i$ is of length $N$. The resulting hash vector of length $M$ is given by $\{< h_K^{NMF-NMF}(I), \mathbf{t}_1 >, ..., < h_K^{NMF-NMF}(I), \mathbf{t}_M >\}$, where $< \mathbf{a}, \mathbf{b} >$ denotes the inner product (which induces the Euclidean norm) of vectors $\mathbf{a}$ and $\mathbf{b}$.

The motivation for the inner product step is to reduce the size of the hash vector. Consider for example, applying the NMF-NMF hashing algorithm to a $512 \times 512$ image, with $p = 10$, $k = 200$, $r_1 = 5$ and $r_2 = 5$. This would result in a hash vector of length $N = 1500$. With floating point storage for each entry, such hash lengths are impractical. This is also a problem with the SVD based hash [10].

We must emphasize however, that the design of the weight vectors $\mathbf{t}_i$ should be done carefully so that the perceptual qualities of the hash are retained. Under a given attack $\mathcal{A}$ on the image $I$, we may model the NMF-NMF hash vector as $h_K^{NMF-NMF}(\mathcal{A}(I)) = h_K^{NMF-NMF}(I) + \mathbf{n}_A$. In particular, we observed that the components of $\mathbf{n}_A$ are approximately i.i.d under a large class of geometric distortions/attacks on the image. For our purposes hence, we picked each $\mathbf{t}_i$ to have i.i.d Gaussian components of zero mean and unit variance. If the noise were to be highly correlated (as is the case with other representations such as wavelets, SVD vectors), the design of the weight vectors would be much harder. Picking weight vectors pseudo-randomly with i.i.d components also enhances the security of the hash. Further, they were chosen to be Gaussian because for a given variance, the Gaussian random variable has the maximum differential entropy.

(a) (b)

Figure 3: Attacked versions of the lena image. (a) rotation by $15^o$, 15% cropping, resizing and JPEG QF = 10, (b) rotation by $25^o$, 30% cropping, resizing, JPEG QF = 10, contrast adjustment and strong Stirmark random bending.

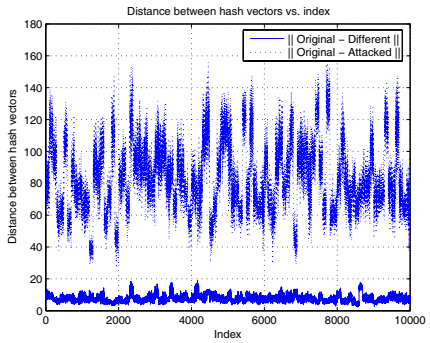
## 4. EXPERIMENTAL RESULTS

For results presented next, the NMF-NMF-SQ hash algorithm was employed with $p = 25$, $k = 100$, $r_1 = 2$, $r_2 = 2$, and hash length $M = 64$. Further, all images were resized to $256 \times 256$ via bicubic interpolation prior to hashing.

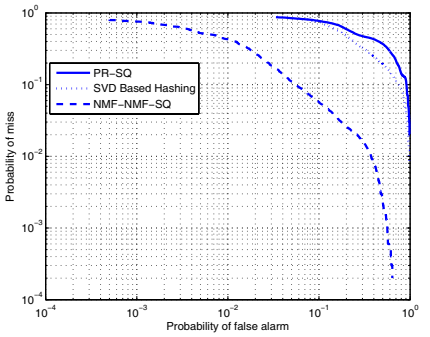
Fig. 3 shows two geometrically attacked versions of the *lena* image. The attack in Fig. 3 (a) is rotation by $15^o$, 15% cropping, resizing and JPEG compression with QF = 10, and in Fig. 3 (b) is rotation by $25^o$, 30% cropping, resizing, JPEG with QF = 10, random contrast enhancement and strong Stirmark [18] random bending. For each attack, we generated hash vectors for 100 different images picked randomly from an image database of 4000 images, and 100 different secret keys. This would result in 10000 pairs of hash vectors corresponding to original and attacked images, and another 10000 pairs corresponding to completely different images.

A robust hash is desired to have the property that when $I$ is perturbed by an attack, the hash is not changed much, or in other words $\| h_K(I) - h_K(\mathcal{A}(I)) \| < \tau$ with high probability. Similarly, if two distinct image $I$ and $I'$ are compared then, it is desired that $\| h_K(I) - h_K(I') \| > \tau$ with high probability. Here, $\| \cdot \|$ denotes a meaningful notion of distance on the difference of hash vectors. We utilized the $L2$ norm, but other measures are possible. Two types of errors are hence possible: the event of $\| h_K(I) - h_K(\mathcal{A}(I)) \| > \tau$ denoted by "miss", and $\| h_K(I) - h_K(I^{'}) \| < \tau$ denoted by "false alarm".

Fig. 4 (a) then plots for the attack in Fig. 3 (a), the distance between hash vectors of altogether different images, and the distance between hash vectors of original and attacked images. It is evident from Fig. 4 (a) that there is no overlap between the solid and dotted plots which means that no errors were observed. Fig. 4 (b) then shows ROC curves for the second attack in Fig. 3 (b). We compare with 1.) hash algorithm based on quantization of pseudo-random statistics of wavelet coefficients [2] which we term




(a)




(b)

Figure 4: (a) Vertical axis is the distance ($L2$ norm) between hash vectors of original and attacked images (solid), and altogether different images (dotted); Fig. 3 (a) shows an example attacked image. Horizontal axis represents the sample index. For this experiment, 100 randomly picked images and 100 secret keys were used resulting in a total 10000 samples. (b) ROC curves for the attack in Fig. 3 (b).

as PR-SQ hashing, and 2.) SVD based image hashing [10]. For both of these schemes, the hash vector was designed to be of length 150 (we did this to give more advantage to the two hash algorithms we compare against).

It is clear from Fig. 4 (b) that both the miss ($P_M$) and false alarm ($P_F$) probabilities are orders of magnitude lower for the proposed NMF-NMF-SQ hash algorithm. Note further that the geometric attack in Fig. 3 (b) is in fact not even perceptually acceptable. The reason we show ROC curves for such an attack is because for most perceptually-acceptable attacks such as the one in Fig. 3 (a) our hash algorithm did not incur any errors.

## 5. CONCLUSIONS AND DISCUSSIONS

We introduce a new pseudo-random signal representation based on non-negative matrix factorizations (NMF) and apply it to the robust perceptual image hashing problem. We make the following crucial observation: the success of hash algorithms based on low-rank matrix approximations depends largely on the constraints employed in obtaining the approximation. We establish that the non-negativity constraints in NMF are much better suited for perceptual hashing than the orthogonality constraints in traditional decompositions, like SVD. Future work may explore applications of our current hash algorithm in anti-piracy search, image authentication, and watermarking.

## 6. REFERENCES


[1] A. Menezes, V. Oorschot, and S. Vanstone, *Handbook of Applied Cryptography*, CRC Press, 1998.

[2] K. Mihcak, R. Venkatesan, and T. Liu, "Watermarking via optimization algorithms for quantizing randomized semi-global image statistics," *ACM Multimedia Systems Journal*, Apr. 2005.

[3] M. Schneider and S. F. Chang, "A robust content based digital signature for image authentication," *Proc. IEEE Conf. on Image Processing*, vol. 3, pp. 227–230, Sept. 1996.

[4] C. Kailasanathan and R. Safavi Naini, "Image authentication surviving acceptable modifications using statistical measures and k-mean segmentation," *Proc. IEEE-EURASIP Work. Nonlinear Sig. and Image*, June 2001.

[5] R. Venkatesan, S. M. Koon, M. H. Jakubowski, and P. Moulin, "Robust image hashing," *Proc. IEEE Conf. on Image Processing*, pp. 664–666, Sept. 2000.

[6] C. Y. Lin and S. F. Chang, "A robust image authentication system distingushing JPEG compression from malicious manipulation," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 2, pp. 153–168, Feb. 2001.

[7] C.-S. Lu and H.-Y. M. Liao, "Structural digital signature for image authentication," *IEEE Transactions on Multimedia*, pp. 161–173, June 2003.

[8] J. Fridrich and M. Goljan, "Robust hash functions for digital watermarking," *Proc. IEEE International Conf. on Information Technology: Coding and Computing*, pp. 178–183, Mar. 2000.

[9] K. Mihcak and R. Venkatesan, "New iterative geometric techniques for robust image hashing," *Proc. ACM Workshop on Security and Privacy in Digital Rights Management Workshop*, pp. 13–21, Nov. 2001.

[10] S. S. Kozat, K. Mihcak, and R. Venkatesan, "Robust perceptual image hashing via matrix invariances," *Proc. IEEE Conf. on Image Processing*, pp. 3443–3446, Oct. 2004.

[11] J. Dittman, A. Steinmetz, and R. Steinmetz, "Content based digital signature for motion picture authentication and content-fragile watermarking," *Proc. IEEE Int. Conf. on Multimedia Computing and Systems*, pp. 209–213, 1999.

[12] V. Monga and B. L. Evans, "Robust perceptual image hashing using feature points," *Proc. IEEE Conf. on Image Processing*, Oct. 2004.

[13] Daniel D. Lee and H. Sebastian Seung, "Algorithms for non-negative matrix factorization," *Advances in Neural Information Processing Systems*, 2001.

[14] Daniel D. Lee and H. Sebastian Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, Oct. 1999.

[15] B. Schiele D. Guillamet and J. Vitria, "Analyzing non-negative matrix factorization for image classification," *IEEE Int. Conf. Pattern Recognition*, vol. 2, pp. 116–119, Aug. 2002.

[16] L. K. Saul and Daniel D. Lee, "Multiplicative updates for classification by mixture models," *Advances in Neural Information Processing Systems*, 2002.

[17] T. Kawamoto, K. Hotta, T. Mishmima, J. Fujiki, M. Tanaka, and T. Kurita, "Estimation of single tones from chord sounds using non-negative matrix factorization," *Neural Network World*, July 2000.

[18] "Fair evaluation procedures for watermarking systems," http://www.petitcolas.net/fabien/watermarking/stirmark, 2000.