Application of l_1-norm minimization technique to image retrieval

C. S. Sastry, Saurabh Jain, and Ashish Mishra

Abstract—Image retrieval is a topic where scientific interest is currently high. The important steps associated with image retrieval system are the extraction of discriminative features and a feasible similarity metric for retrieving the database images that are similar in content with the search image. Gabor filtering is a widely adopted technique for feature extraction from the texture images. The recently proposed sparsity promoting l_1-norm minimization technique finds the sparsest solution of an under-determined system of linear equations. In the present paper, the l_1-norm minimization technique as a similarity metric is used in image retrieval. It is demonstrated through simulation results that the l_1-norm minimization technique provides a promising alternative to existing similarity metrics. In particular, the cases where the l_1-norm minimization technique works better than the Euclidean distance metric are singled out.

Keywords—l_1-norm minimization, content based retrieval, modified Gabor function.

I. Introduction

CONTENT Based Image Retrieval (CBIR) from large image databases has been an active area of research for long due to its applications in various fields like satellite imaging, medicine etc. CBIR systems extract features from the raw images and calculate an associative measure (similarity or dissimilarity) between a search image and database images based on these features. Hence feature extraction and selection of suitable similarity measure are very important steps. Several methods achieving effective feature extraction have been proposed in the literature [1] [7] [8], to name a few.

The Gabor filter is widely used to extract the texture features from images for image retrieval. This use is motivated by many factors [1] [7] [8]. Some of the methods [1] [7] proposed for feature extraction use direction dependent Gabor filters, and consequently the feature vectors become direction dependent. A new feature extraction technique is proposed in [8], which gets rid of direction dependence of Gabor filters. It is justified that the method in [8] is effectively used for CBIR application.

In various applications of numerical analysis, a common problem is to approximate or interpolate a complicated function using a short linear combination of more elementary functions. The problem of computing sparse linear representations with respect to redundant dictionary of basis elements has recently become a topic of high interest. The primary reason for this centers around the discovery that whenever the optimal representation is sufficiently sparse, it can be computed by l_1-norm minimization techniques [4] [5] [6]. As mentioned already, the suitable choice of similarity metric is inevitable for better performance of CBIR systems.

In this paper, a new similarity metric is derived in terms of recent sparsity promoting l_1-norm minimization technique. It is demonstrated experimentally that the similarity metric based on the l_1-norm minimization technique shows better retrieval performance than the standard metric based on the Euclidean distance. In particular, the cases where the l_1-norm based similarity metric works relatively better than the metric based on Euclidean distance are singled out.

The paper is organized as follows: In section 2, the Gabor function and the ways of feature extraction using it are discussed. Later on in section 3, the recent sparsity based l_1-norm minimization technique is presented. In sections 4 and 5, a procedure with l_1-norm minimization as a similarity metric in CBIR is presented. In the last section, simulation results and some concluding remarks are presented in some detail.

II. Gabor Transform

In this section, the Gabor based feature extraction technique is discussed. A 2D Gabor function is defined as

$$g_{\sigma_x,\sigma_y}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)} e^{2\pi i W x}, \qquad (1)$$

where, σ_x and σ_y are the scaling parameters of the filter, and W is central frequency. The function g_{σ_x,σ_y} acts as a local band-pass filter with certain optimal localization properties in both spatial and frequency domains. A class of self-similar Gabor filters can be obtained by appropriate dilations and rotations of g_{σ_x,σ_y} as follows:

$$\begin{aligned} g_{\sigma_x,\sigma_y,\theta}(x,y) \;&=\; \frac{1}{2\pi\sigma_x\sigma_y} e^{2\pi i W(x\cos\theta + y\sin\theta)} \\ &\quad . \;\; e^{-\frac{1}{2}\left(\frac{(x\cos\theta+y\sin\theta)^2}{\sigma_x^2}+\frac{(-x\sin\theta+y\cos\theta)^2}{\sigma_y^2}\right)}, \end{aligned} \qquad (2)$$

where θ determines the orientation of the filter. As the axes for $g_{\sigma_x,\sigma_y,\theta}$ are rotated by θ, the function $g_{\sigma_x,\sigma_y,\theta}$ has undulations along θ-direction. Given an input image f, the Gabor transform of it is the convolution with f of a set of Gabor filters of different preferred orientations and scales, that is, $F_{\sigma_x,\sigma_y,\theta}(x,y) := f \star g_{\sigma_x,\sigma_y,\theta}(x,y)$. Based on common observation [1] [7] [8], the mean and variance of the energy distribution of $F(x,y)_{\sigma_x,\sigma_y,\theta}$ are computed to identify a texture. The signature or compact representation of an image

C. S. Sastry, is with the Department of Electronics and Communication Engineering, Indian Institute of Information Technology, Design & Manufacturing (PDPM IIITDM), Dumna Airport Road, Jabalpur, M.P., 482011, India e-mail: (csastry@iiitdm.in).

Saurabh Jain is with German Automotive Electronics firm in Pune, India, e-mail: (saurabh.shanky@gmail.com).

Ashish Mishra is presently at IIT Bombay India, e-mail: (amishra.jsr@gmail.com).

Manuscript received June 26, 2009; revised July 31, 2009.

is, then, generated by considering $M_{\sigma_x,\sigma_y,\theta}$ and $V_{\sigma_x,\sigma_y,\theta}$ in vector form for different σ_x, σ_y and θ.

In order to generate angle independent Gabor filters and thereby to generate rotation invariant features, a modified Gabor function G_{σ_x,σ_y} is defined in [8] as

$$G_{\sigma_x,\sigma_y}(x,y) = \frac{1}{2\pi}\int_0^{2\pi} g_{\sigma_x,\sigma_y,\theta}(x,y)d\theta. \tag{3}$$

Then feature vectors are designed by taking $M_{\sigma_x,\sigma_y,f,i}$ and $V_{\sigma_x,\sigma_y,\eta,i}$ as components with

$$\begin{aligned} M_{\sigma_x,\sigma_y,f,i} &= \int\int_{D_i} \left|G_{\sigma_x,\sigma_y} \star f(x,y)\right| dxdy, \\ V_{\sigma_x,\sigma_y,f,i} &= \int\int_{D_i} \Big(|f \star G_{\sigma_x,\sigma_y}(x,y)| \\ &- M_{\sigma_x,\sigma_y,f,i} \Big)^2 dxdy, \end{aligned} \tag{4}$$

where, $D_i = \{(x,y)|R_i^2 \leq x^2+y^2 < R_{i+1}^2\}$ is the region over which the double integration is performed in the above equations. It is demonstrated in [8] that the modified Gabor function shows better performance than the standard Gabor function for texture classification. In view of this, the present work uses the feature extraction procedure that is based on the modified Gabor function.

III. Sparsity promoting optimization techniques

In this section, the powerful ideas proposed for solving underdetermined systems are reviewed briefly.

Consider the matrix equation

$$y = \Phi x_0, \tag{5}$$

where Φ is a matrix of size $n \times N$ and x_0 is the solution of the above system. The case where the system (5) is underdetermined (i.e., $n < N$) is of particular interest in sparsity promoting methods. In this case, the system admits infinitely many solutions.

A. Convex optimization techniques

One way [5] [6] of choosing a solution of (5) involves taking 'shortest' vector. More precisely, assume that the vector $x_0 \in \Re^N$ is sparse. That is, the set $Sparsity(x_0)$ defined to be $Sparsity(x_0) := \{i : x_{0,i} \neq 0\}$ has the cardinality $\|x_0\|_0 := |Sparsity(x_0)| = k$ smaller than N. Additionally, assume that the columns of Φ have unit norm. The sparse solution of (5) is obtained intuitionally from

$$\min_\alpha \|\alpha\|_0 \quad subject \quad to \quad \Phi\alpha = y, \tag{6}$$

which is a non-convex and an NP-hard problem. *Basis pursuit* (BP) [5] convexifies the algorithm and proposes the solution in terms of l_1-norm minimization as follows

$$\min_\alpha \|\alpha\|_1 \quad subject \quad to \quad \Phi\alpha = y. \tag{7}$$

The above optimization problem can be recast as a linear programming problem (LPP), for which solutions are available even in large scale problems owing to modern interior point linear programming methods. The results in [5] have empirically observed that the convexification of l_0 optimization works well, implying that if indeed a (sparse) solution of (6) exists, (7) finds it. Several recent works (for example, [2] and the references therein) have studied theoretically this phenomenon and have found that under certain conditions, (6) and (7) lead to the same solution. One common theme in these approaches that analyze BP is the use of the *coherence parameter* as a way to characterize the dictionary Φ. The coherence parameter μ is defined as the maximal inner product between the pairs of dictionary columns, i.e.,

$$\mu = max_{j\neq k}\left| < \phi_j, \phi_k > \right|, \tag{8}$$

where ϕ_i stands for the i^{th} column of Φ. For a full rank general matrix, the coherence parameter μ satisfies [5][6]

$$\mu \geq \sqrt{\frac{N-n}{n(N-1)}}. \tag{9}$$

For an orthogonal matrix Φ, μ is zero, while for an over complete dictionary it is necessarily nonzero. The parameter μ is desired to be small. Then Φ almost behaves like an unitary matrix. When μ is small, the dictionary Φ is said to be *incoherent*, and in this case the solution of the system (5) is highly sparse.

B. Recovery of sparsest solution from noisy measurements

In many practical situations, it is not sensible to assume [4] [5] that the available data (y) obey precise $y = \Phi x_0$ equality. A more realistic scenario assumes noise in the measurements as

$$y = \Phi x_0 + e; \quad \|e\|_2 \leq \epsilon. \tag{10}$$

The solution [4] of above problem can be obtained by solving

$$\min_\alpha \|\alpha\|_1 \quad subject \quad to \quad \|\Phi\alpha - y\|_2 \leq \epsilon, \tag{11}$$

In the framework of Uniform Uncertainty Principle (UUP), it is proved in [4] that the stable recovery of solution of (10) is possible and (11) recovers an unknown sparse object with an error at most proportional to the noise level, as follows

Theorem [4]: Let k be such that $\delta_{3k}+3\delta_{4k} < 2$. Then, for any signal x_1 supported on T with $|T| \leq k$ and any perturbation e with $\|e\|_2 \leq \epsilon$, the solution $\tilde{x}_0$ to (11) obeys

$$\|x_1 - \tilde{x}_0\|_2 \leq C_k\epsilon, \tag{12}$$

where, the constant C_k depends only on δ_{4k} and behaves well for reasonable values of k.

If small changes in the observations result in small changes in the recovery, the solution is said to be stable.

IV. CBIR algorithm based on l_1-norm minimization

In this section, a new image retrieval method is presented with the l_1-norm minimization method as similarity metric.

As stated already, the feature vectors of both search and database images are generated in the present work using the modified Gabor function. Then the matrix Φ is defined to be the feature matrix consisting of feature vectors of database

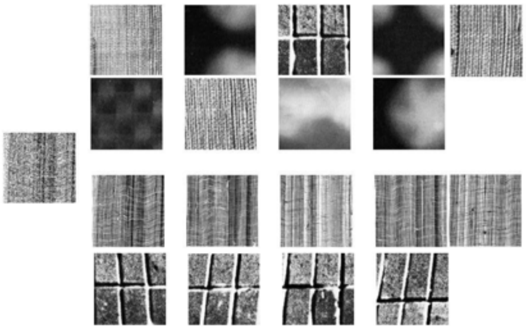

Fig. 1. Retrieval performances of l_2 and l_1 metrics with noisy search images. The images on the first two rows are those retrieved by l_2 metric, while the images on the last two rows are those retrieved by l_1 metric. The image on the first column is search image.

elements as columns. The vector y is taken to be the feature vector of search image. In order to retrieve the database images that are similar in content with the search image, the matrix equation $y = \Phi x_0$ is solved for its sparsest solution using the l_1-norm minimization technique discussed in the previous section. The reason for seeking sparsest solution is that the nonzero components in the solution vector helps in identifying the columns of Φ that are most relevant to y. Suppose i^{th} component of x_0 is maximum in magnitude. Then the image corresponding to the i^{th} column of Φ is labeled to be the most relevant retrieved image. In order to identify the next relevant image, the i^{th} column of Φ is removed and the resultant matrix (say, Φ again) is used in place of initial Φ. Then the matrix equation $y = \Phi x_0$ is solved to get sparsest solution. As the first relevant image is identified, the second most relevant is labeled to be the image corresponding to the column of Φ whose respective component in the solution has maximum absolute value. This way the first few most relevant images may be identified.

It is observed experimentally that the coherence parameter μ of the feature matrix Φ lies closer to the lower bound. As the database is very big, the inequality $n << N$ is an obvious consequence and hence the system admits highly sparse solution. If the number of images that are intended to be retrieved from database is more than the number of significant components of sparse solution, a problem in retrieving the prescribed number of images may arise. This is the reason for removing a column from Φ each time a relevant image is identified. The stated way of retrieving similar images is computationally somewhat more involved than the similarity metric based on the standard Euclidean distance. But in some cases, the stated procedure outperforms Euclidean based similarity metric, which is discussed in the next section. To accommodate the errors arising out of discrete domain implementation of Gabor function, some error in the computation of feature vectors is allowed and the matrix equation is taken to be $y = \Phi x_0 + e$ with $\|e\| \leq \epsilon$. The

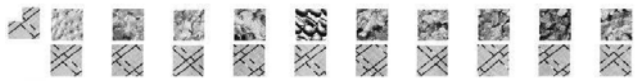

Fig. 2. Retrieval performances of l_2 and l_1 metrics with cropped search images. The images on the first row are those retrieved by l_2 metric, while the images on the second row are those retrieved by l_1 metric. The image on the first column is search image.

pseudo code of the proposed method is given below:

- Input: Feature matrix Φ, feature vector y of search image. L : Number of images to be retrieved. Set $\Phi_1 = \Phi$.
- For each of $i = 1, 2, \ldots, L$
  Solve $y = \Phi_i x_0 + e$ with $\|e\| \leq \epsilon$.
  Let j be such that $|x_j| \geq |x_l| \; \forall \; l \neq j$.
  Let Φ_{i+1} be Φ_i without i^{th} column.
  i^{th} most relevant image of search image is the image corresponding to j^{th} column of Φ_i.
- Output: L number of images of database that are similar to search image.

V. Simulation Results

As stated already in previous sections, the goal of the present work is to study and identify whether or not the similarity metric provided by the recent l_1-norm minimization technique is useful in CBIR applications.

Using the modified Gabor function, the feature vectors of size 32 have been generated with 4 choices of (σ_x, σ_y), viz $\sigma_x = \sigma_y = 1, 2, 3, 4$. The components in feature vectors are means and variances over concentric circular regions as suggested in [8]. Hereafter, the similarity metrics provided by the Euclidean distance and the l_1-norm minimization techniques are referred to as l_2 and l_1 metrics respectively. The present simulation work has been carried out with the software developed in l_1 magic [3], by setting ϵ in (10) to 0.001. It has been noted that the overall retrieval performance of l_1 similarity metric is at least same as that of l_2 metric on Brodatz image database [1]. Simulation work, however, suggests that: 1). Low contrast images are better retrieved by l_1 similarity metric than with l_2 metric, 2). l_1 metric has consistent performance within the same class, that is, when several images of same class are taken as search images, l_1 metric shows similar retrieval performance. On the otherhand, the retrieval performance of l_2 metric within the same class is highly inconsistent. 3). Even when both l_1 and l_2 metrics show similar retrieval performance, l_1 metric catches more relevant images first. This aspect may be seen in Figure 1.

Intending to study the performance of both metrics on the noisy search images, simulation work considered the search image that is corrupted by Gaussian white noise. Experimental results have shown that the Gaussian white noise of mean 0.1 and variance up to 0.05 does not affect the retrieval performance of l_1 similarity metric. While in the same noisy setting, the performance of l_2 metric gets affected. When the search images are corrupted with salt & pepper noise, experimental results have indicated that the salt & pepper noise of density up to 0.12 does not affect the retrieval performance of l_1 metric. Figure 1 shows the retrieval performance of l_1

metric on noisy search image. In Figure 1, the first two rows are of l_2 metric and while the next two rows are of l_1 metric. The single image on first column is noisy search image.

Finally, partially cropped images have been considered as search images, which are shown for example on the first column of Figure 2 . Simulation work with l_1 metric has indicated that for some search images, the retrieval performance is not affected significantly, when the window of size up to 56×56 is cropped in a search image of size 128×128. The retrieval performance of a cropped image is shown in Figure 2.

VI. Conclusion

In the present work, a new similarity metric is proposed for CBIR in terms of the recent powerful l_1-norm minimization technique, which is developed for solving the underdetermined system of linear equations. It is demonstrated through the simulation work that the new similarity metric has the potential to perform well as compared to the metric based on the standard Euclidean distance for the CBIR of texture images.

References

[1] P. Brodatz, *Textures: A photographic album for artists and designers*, Dover Publication, New York, 1996.

[2] A. M. Bruckstein, D. L. Donoho and M. Elad, *From sparse solutions of systems of equations to sparse modeling of signals and images*, SIAM Review, vol. 51, No. 1, pp: 34-81, 2009.

[3] E. Candes and J. Romberg, l_1 *magic: Recovery of sparse signals via convex programming*, http://www.acm.caltech.edu/l1magic/, 2005.

[4] E. Candes and J. Romberg and T. Tao, *Stable signal recovery from incomplete and inaccurate measurements*, Comm. Pure and Applied Maths, 59, 1207–1223, 2006.

[5] S. Chen and D. Donoho and M. Saunders, *Atomic decomposition by basis pursuit*, SIAM Review, 43(1), 129–159, 2001.

[6] D. Donoho, *For most large underdetermined systems of linear equations the minimal* l_1*-norm near solution approximates the sparsest solution*, Comm. Pure and Applied Maths, 59(10), 907–34, 2006.

[7] B. S. Manjunath and Y. S. Ma, *Texture features for browsing and retrieval of image data*, IEEE T. on Pattern Analysis and Machine Intelligence, 18(8):837–842, 1996.

[8] C. S. Sastry and M. Ravindranath and A. K. Pujari B. L. Deekshatulu, *A modified Gabor method for content based image retrieval*, Pattern Recognition Letters, 28, 293–300, 2007.

C. S. Sastry obtained his Ph.D from IIT Kanpur. He is presently working as an Assistant Professor at IIITDM, Jabalpur. His research interests include wavelets, recent sparsity promoting techniques and tomography.

Saurabh Jain obtained his bachelor's degree in Electronics and Communication Engineering from IIITDM Jabalpur in 2009. He is presently associated with German Automotive Electronics firm in Pune, India.

Ashish Mishra obtained his bachelor's degree in Electronics and Communication Engineering from IIITDM Jabalpur in 2009. He is presently working as a project associate at IIT Bombay, India.