Texture Features for Browsing and Retrieval of Image Data

B.S. Manjunath and W.Y. Ma

Abstract—Image content based retrieval is emerging as an important research area with application to digital libraries and multimedia databases. The focus of this paper is on the image processing aspects and in particular using texture information for browsing and retrieval of large image data. We propose the use of Gabor wavelet features for texture analysis and provide a comprehensive experimental evaluation. Comparisons with other multiresolution texture features using the Brodatz texture database indicate that the Gabor features provide the best pattern retrieval accuracy. An application to browsing large air photos is illustrated.

Index Terms—Digital libraries, image database, content-based image retrieval, texture analysis, Gabor wavelets.

1 Introduction

Retrieval of image data based on pictorial queries is an interesting and challenging problem. The recent emergence of multimedia databases and digital libraries makes this problem all the more important. While manual image annotations can be used to a certain extent to help image search, the feasibility of such an approach to large databases is a questionable issue. In some cases, such as face or texture patterns, simple textual descriptions can be ambiguous and often inadequate for database search.

The objective of this paper is to study the use of texture as an image feature for pattern retrieval. An image can be considered as a mosaic of different texture regions, and the image features associated with these regions can be used for search and retrieval. A typical query could be a region of interest provided by the user, such as outlining a vegetation patch in a satellite image. The input information in such cases is an intensity pattern or texture within a rectangular window. See Fig. 6 for an example of a texture based browsing application.

Texture analysis has a long history and texture analysis algorithms range from using random field models to multiresolution filtering techniques such as the wavelet transform. Several researchers have considered the use of such texture features for pattern retrieval [18], [19]. This paper focuses on a multiresolution representation based on Gabor filters. The use of Gabor filters in extracting textured image features is motivated by various factors. The Gabor representation has been shown to be optimal in the sense of minimizing the joint two-dimensional uncertainty in space and frequency [4]. These filters can be considered as orientation and scale tunable edge and line (bar) detectors, and the statistics of these microfeatures in a given region are often used to characterize the underlying texture information. Gabor features have been used in several image analysis applications including texture classification and segmentation [1], [14], image recognition [5], [8], [13], image registration, and motion tracking [15].

The main contributions of this paper are summarized below:

1) A simple texture feature representation based on Gabor features is proposed, and a filter design strategy is suggested to reduce the redundancy in the representation.
2) An adaptive filter selection algorithm is proposed which can facilitate fast image browsing.
3) A detailed comparison with the performance of three other multiscale texture features is provided. Our proposed representation compares favorably in terms of feature computations and retrieval accuracy.
4) Finally, an application to browsing large air photos is demonstrated.

• *The authors are with the Department of Electrical and Computer Engineering, University of California at Santa Barbara, Santa Barbara, CA 93106-9560. E-mail: manj@ece.ucsb.edu, wei@iplab.ece.ucsb.edu.*

Manuscript received Dec. 16, 1994. Recommended for acceptance by R. Picard. For information on obtaining reprints of this article, please send e-mail to: transpami@computer.org, and reference IEEECS Log Number P96055.

2 Texture Feature Extraction

2.1 Gabor Functions and Wavelets

A two dimensional Gabor function $g(x, y)$ and its Fourier transform $G(u, v)$ can be written as:

$$g(x,y) = \left(\frac{1}{2\pi\sigma_x\sigma_y}\right)\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jWx\right] \quad (1),$$

$$G(u,v) = \exp\left\{-\frac{1}{2}\left[\frac{(u-W)^2}{\sigma_u^2}+\frac{v^2}{\sigma_v^2}\right]\right\} \quad (2),$$

where $\sigma_u = 1/2\pi\sigma_x$ and $\sigma_v = 1/2\pi\sigma_y$. Gabor functions form a complete but nonorthogonal basis set. Expanding a signal using this basis provides a localized frequency description. A class of self-similar functions, referred to as *Gabor wavelets* in the following discussion, is now considered. Let $g(x, y)$ be the mother Gabor wavelet, then this self-similar filter dictionary can be obtained by appropriate dilations and rotations of $g(x, y)$ through the generating function:

$$g_{mn}(x,y) = a^{-m}G(x',y'), \quad a>1, \quad m,n = \text{integer}$$
$$x' = a^{-m}(x\cos\theta + y\sin\theta), \text{ and } y' = a^{-m}(-x\sin\theta + y\cos\theta), \quad (3)$$

where $\theta = n\pi/K$ and K is the total number of orientations. The scale factor a^{-m} in (3) is meant to ensure that the energy is independent of m.

2.2 Gabor Filter Dictionary Design

The nonorthogonality of the Gabor wavelets implies that there is redundant information in the filtered images, and the following strategy is used to reduce this redundancy. Let U_l and U_h denote the lower and upper center frequencies of interest. Let K be the number of orientations and S be the number of scales in the multiresolution decomposition. Then the design strategy is to ensure that the half-peak magnitude support of the filter responses in the frequency spectrum touch each other as shown in Fig. 1. This results in the following formulas for computing the filter parameters σ_u and σ_v (and thus σ_x and σ_y).

$$a = (U_h/U_l)^{-\frac{1}{S-1}}, \quad \sigma_u = \frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}},$$

$$\sigma_v = \tan\left(\frac{\pi}{2k}\right)\left[U_h - 2\ln\left(\frac{\sigma_u^2}{U_h}\right)\right]\left[2\ln 2 - \frac{(2\ln 2)^2\sigma_u^2}{U_h^2}\right]^{-\frac{1}{2}}, \quad (4)$$

where $W = U_h$ and $m = 0, 1, \ldots, S-1$. In order to eliminate sensitivity of the filter response to absolute intensity values, the real (even) components of the 2D Gabor filters are biased by adding a constant to make them zero mean (This can also be done by setting $G(0, 0)$ in (2) to zero).

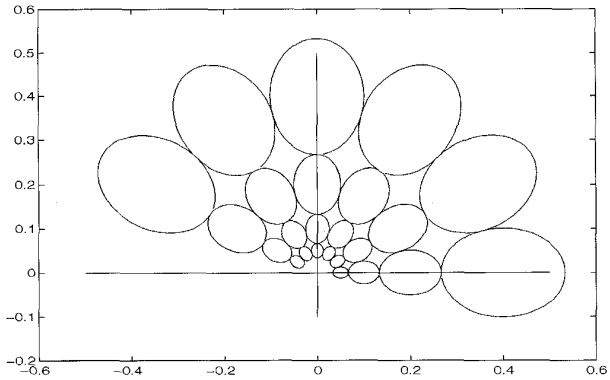


Fig. 1. The contours indicate the half-peak magnitude of the filter responses in the Gabor filter dictionary. The filter parameters used are $U_h = 0.4$, $U_l = 0.05$, $K = 6$, and $S = 4$.

2.3 Feature Representation

Given an image $I(x, y)$, its Gabor wavelet transform is then defined to be

$$W_{mn}(x,y) = \int I(x_1, y_1) g_{mn}^*(x - x_1, y - y_1) dx_1 dy_1, \quad (5)$$

where * indicates the complex conjugate. It is assumed that the local texture regions are spatially homogeneous, and the mean μ_{mn} and the standard deviation σ_{mn} of the magnitude of the transform coefficients are used to represent the region for classification and retrieval purposes:

$$\mu_{mn} = \int\int |W_{mn}(xy)| dxdy, \text{ and } \sigma_{mn} = \sqrt{\int\int (|W_{mn}(x,y)| - \mu_{mn})^2 dxdy}. \quad (6)$$

A feature vector is now constructed using μ_{mn} and σ_{mn} as feature components. In the experiments, we use four scales $S = 4$ and six orientations $K = 6$, resulting in a feature vector

$$\bar{f} = [\mu_{00}\ \sigma_{00}\ \mu_{01}\ \cdots\ \mu_{35}\ \sigma_{35}]. \quad (7)$$

2.3.1 Distance Measure:

Consider two image patterns i and j, and let $\bar{f}^{(i)}$ and $\bar{f}^{(j)}$ represent the corresponding feature vectors. Then the distance between the two patterns in the feature space is defined to be

$$d(i,j) = \sum_m \sum_n d_{mn}(i,j),$$

where

$$d_{mn}(i,j) = \left| \frac{\mu_{mn}^{(i)} - \mu_{mn}^{(j)}}{\alpha(\mu_{mn})} \right| + \left| \frac{\sigma_{mn}^{(i)} - \sigma_{mn}^{(j)}}{\alpha(\sigma_{mn})} \right|. \quad (8)$$

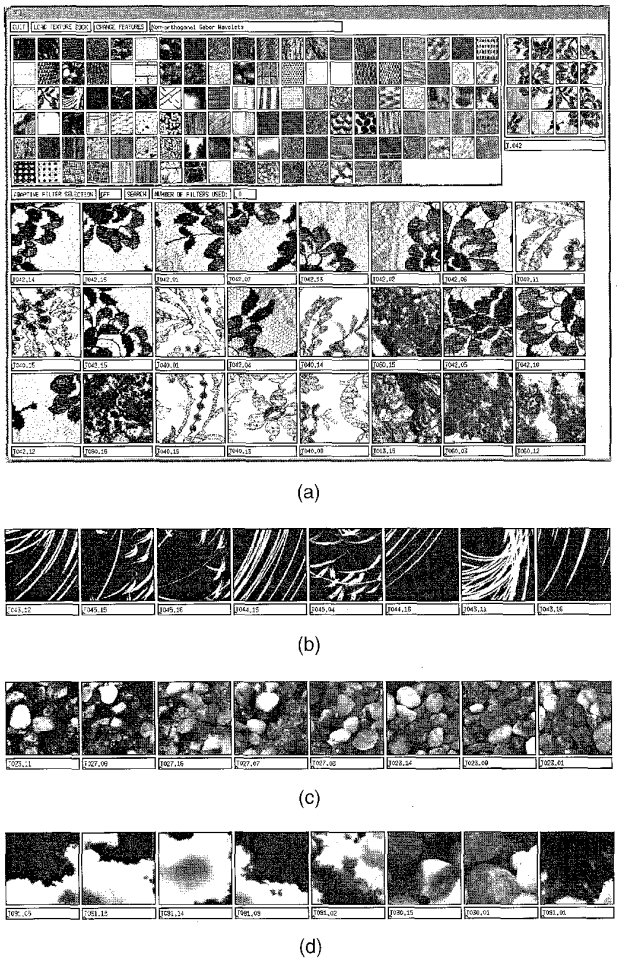
$\alpha(\mu_{mn})$ and $\alpha(\sigma_{mn})$ are the standard deviations of the respective features over the entire database, and are used to normalize the individual feature components.

2.4 Retrieval Performance

2.4.1 Texture Database

The texture database used in the experiments consists of 116 different texture classes. Each of the 512 × 512 images is divided into 16 128 × 128 nonoverlapping subimages, thus creating a database of 1,856 texture images. A query pattern in the following is any one of the 1,856 patterns in the database. This pattern is then processed to compute the feature vector as in (7). The distance $d(i, j)$, where i is the query pattern and j is a pattern from the database, is computed. The distances are then sorted in increasing order and the closest set of patterns are then retrieved. In the ideal case all the top 15 retrievals are from the same large image. The performance is measured in terms of the average retrieval rate which is defined as the average percentage number of patterns belonging to the same image as the query pattern in the top 15 matches.

We observe that the use of σ_{mn} feature in addition to the mean improves the retrieval performance considerably. This perhaps explains the low classification rate of the Gabor filters reported in [3] where only the mean value was used. On the average 74.37% of the correct patterns are in the top 15 retrieved images. The performance increases to 92% if the top 100 (about 6% of the entire database) retrievals are considered instead (i.e., more than 13 of the 15 correct patterns are present). Some retrieval examples are shown in Fig. 2. A detailed comparison with other texture features is given in Section 4.



(a) (b) (c) (d)

Fig. 2. Texture browsing using Gabor features. Examples shown are some of the difficult patterns to analyze. The average retrieval rate is shown in the parentheses: (a) Texture browsing interface with all 116 textures and retrievals for D42, lace (50%), (b) D43, swinging light bulb (54%), (c) D23, beach pebbles (54%), and (d) D91, clouds (25%). The first image in each of (b)-(d) represents the query image. Note that some of the incorrect matches actually look quite similar to the query pattern.

3 Adaptive Filter Selection

In many cases it is desirable to reduce the image processing time to the extent possible while not seriously affecting the overall performance. One way to reduce the computations is to select the Gabor filters in a pattern dependent way.

An insight into the discrimination quality of individual features can be obtained by considering the average intraclass to inter-class distance ratio for each of the components. This ratio is about 0.25 on the average, with higher frequency components providing a better discrimination [16]. In database search one is often interested in finding out how much of the search space can be eliminated by using a particular feature. Suppose we want to keep all the 15 correct textures belonging to the same pattern in the set of retrieved images. The average threshold is about 30%, and as in the previous case, high frequency components have a better discrimination. More details can be found in [16]. These experiments indicate that each feature component can individually provide useful discrimination measure which can be further improved by selectively choosing these filters based on the query pattern.

The selection scheme described here uses the spectral information in conjunction with the *average database image* properties to select a subset of filters. The purpose is to identify salient query image properties which best distinguish it from the database items. Fig. 3 shows a schematic diagram of this method. The difference between the spectrum of the input image pattern and the average spectrum provides information about salient spectral characteristics of the given image. Let

$$D(u,v) = \left[\left|F_{input}(u,v)\right| - \left|F_{mean}(u,v)\right|\right]^2 \Big/ \left(F_{var}(u,v)\right),$$

where $F_{input}(u,v)$ is the Fourier transform of the input image pattern. $F_{mean}(u,v)$ and $F_{var}(u,v)$ are the mean and variance associated with the distribution of Fourier transforms of all image patterns in the database. $D(u, v)$ basically is the energy of the *difference* normalized by the variance associated with each frequency component (u, v). Each filter is evaluated based on the total *difference* energy within its spectral coverage:

$$\zeta_{mn} = \sum_u \sum_v D(u,v) \cdot \left|G_{mn}(u,v)\right|^2,$$

where $G_{mn}(u,v)$ is the frequency response of the filter $g_{mn}(x,y)$. The larger the value ζ_{mn} is, the better the performance of the filter will be for pattern retrieval. Thus the filters can be ordered based on ζ_{mn}. Fig. 3 shows an example where the input pattern has strong orientation preference which distinguishes the pattern from much of the database images. Using only the top four filters, one can retrieve on the average about 50% of the correct patterns in the top 15 retrieved patterns [16]. Such a strategy may be reasonable in many image database applications and amounts to a significant savings in image processing computations. Fig. 4 shows some examples wherein different number of filters are used.

Computing the entire Gabor feature vector of (7) takes 9.3 seconds of CPU time (in MATLAB on a SUN-Sparc20, see Table 2). Search and retrieval takes about 1.02 seconds. Using only four feature components requires 2.3 seconds of feature extraction time including the adaptive filter selection (0.7 seconds), and about 0.1 seconds for search and retrieval.

4 Experimental Results

A detailed comparison with some of the other recently proposed multiresolution texture image features are made in this section. For the Gabor feature case, the 24 × 2 component feature vectors are used. An application to browsing large air-photos is illustrated.

4.1 Comparison with Other Texture Features

The comparisons are made with the conventional pyramid-structured wavelet transform (PWT) features, tree-structured wavelet transform (TWT) features, and the multiresolution simultaneous autoregressive model (MR-SAR) features. The filter coefficients used for computing PWT are the 16-tap Daubechies orthogonal wavelets [6] (same as the ones used in [3] for the TWT). The 128 × 128 image pattern is decomposed into three levels (4 × 3 = 12 bands) of the wavelet transform. The mean and standard deviation of the energy distribution corresponding to each of the subbands at each decomposition level are used to construct a (12 × 2) feature vector.

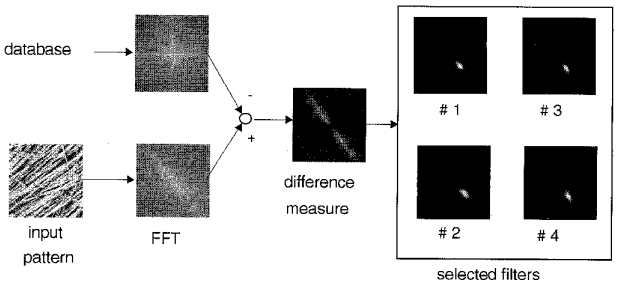


Fig. 3. An example of the filter selection strategy.

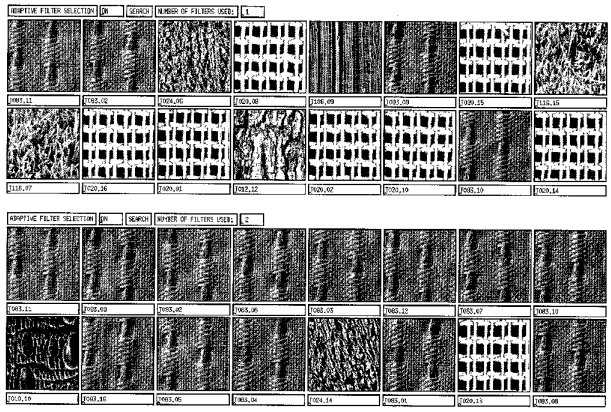


Fig. 4. Examples of retrieval using a different number of filters selected by the adaptive filter selection strategy. The upper-left image in each block is the query image (D83: woven matting), and the top 15 matches are shown in row scan order with increasing distance. Top: Results using one filter; bottom: results using two filters. With four filters, all the correct 15 patterns are retrieved for this case.

In [3], decomposition of image subbands at each level is based on energy considerations and this results in a tree structured decomposition where different patterns have different structures. For pattern retrieval applications, it is convenient to have a fixed structure. A fixed decomposition tree can be obtained by sequentially decomposing the LL, LH, and HL subbands. The HH band is not decomposed as this often does not lead to stable features. A three level decomposition results in 52 (4(1 + 3 + 9)) subbands. As in the PWT, the mean and standard deviation in each subband are used to construct a 52 × 2 component feature vector.

The third set of feature used are the MR-SAR model features [17]. Previous work [9], [20] indicate that the MR-SAR features at levels 2, 3, and 4, provide the best overall performance. At each level, five parameters are computed to represent the texture, thus requiring a total of 15 feature components. The Mahalanobis distance is used to compare the feature vectors.

4.1.1 Summary of Comparisons

Table 1 provides a summary of the experimental results. It shows the retrieval accuracy of the different texture features for each of the 116 texture classes in the database. The Gabor features give the best performance at close to 74% retrieval. This is closely followed by the MR-SAR features at 73%. The TWT features perform mar-

TABLE 1
AVERAGE RECOGNITION RATE FOR THE 116 TEXTURE IMAGES IN THE DATABASE.
THE D* LABELS INDICATE TEXTURES FROM THE BORDATZ ALBUM [2] AND O* LABELS INDICATE TEXTURES FROM THE USC DATABASE.

	Average Retrieval Rate%					Average Retrieval Rate%					Average Retrieval Rate%			
	Gabor	PWT	TWT	MRSAR		Gabor	PWT	TWT	MRSAR		Gabor	PWT	TWT	MRSAR
D1	99.17	97.08	97.50	98.75	D42	50.00	59.17	56.67	38.75	D81	100.00	90.83	95.83	94.17
D2	52.92	32.50	36.25	67.92	D43	11.25	13.75	11.25	9.58	D82	100.00	100.00	100.00	100.00
D3	94.58	75.42	58.33	49.17	D44	12.50	13.33	12.92	15.00	D83	100.00	98.75	99.58	99.17
D4	100.00	90.83	67.92	88.33	D45	14.58	22.08	25.83	3.33	D84	100.00	100.00	100.00	99.58
D5	72.92	52.92	52.08	63.33	D46	94.17	70.42	80.42	92.50	D85	99.58	96.67	100.00	99.17
D6	100.00	100.00	100.00	100.00	D47	100.00	100.00	100.00	97.50	D86	91.67	60.83	71.67	95.00
D7	35.42	21.25	19.58	52.08	D48	49.17	72.08	77.08	86.67	D87	99.58	92.08	82.92	92.92
D8	95.00	79.58	74.58	95.42	D49	100.00	100.00	100.00	100.00	D88	41.67	48.75	51.67	45.83
D9	93.75	84.58	77.50	64.58	D50	87.92	56.25	75.83	87.08	D89	21.25	22.08	26.25	10.00
D10	85.83	78.75	68.75	79.17	D51	83.75	91.25	93.33	90.83	D90	34.58	19.58	31.67	47.08
D11	100.00	73.75	80.00	98.33	D52	72.08	55.42	61.25	70.00	D91	25.42	12.92	16.25	27.50
D12	86.25	79.58	70.42	75.00	D53	100.00	100.00	100.00	100.00	D92	87.50	92.50	92.08	68.33
D13	42.92	38.75	30.83	43.75	D54	50.83	56.67	47.50	57.92	D93	72.92	38.75	30.00	62.08
D14	100.00	100.00	100.00	100.00	D55	100.00	97.08	99.17	100.00	D94	100.00	91.67	92.08	90.83
D15	69.17	79.58	57.92	80.00	D56	100.00	100.00	100.00	100.00	D95	87.50	65.00	92.50	87.50
D16	100.00	100.00	100.00	100.00	D57	100.00	94.17	100.00	100.00	D96	98.33	77.50	94.17	99.17
D17	100.00	100.00	100.00	100.00	D58	29.58	18.33	25.42	27.92	D97	37.08	29.17	39.58	37.92
D18	79.17	79.17	92.50	69.58	D59	20.42	10.83	14.17	23.33	D98	52.50	52.08	65.00	72.92
D19	80.42	73.33	62.50	86.67	D60	52.50	30.00	37.92	50.00	D100	87.08	71.67	53.33	74.58
D20	100.00	87.50	94.58	100.00	D61	43.75	47.92	45.83	41.25	D101	58.75	65.00	63.75	62.92
D21	100.00	100.00	100.00	100.00	D62	35.83	45.00	50.83	43.75	D102	53.33	51.25	51.25	56.25
D22	75.00	82.50	70.83	56.67	D63	34.17	24.17	27.08	35.00	D103	56.67	72.50	73.75	66.67
D23	53.75	41.25	41.67	39.58	D64	94.58	90.00	79.17	97.92	D104	54.58	59.17	66.67	48.75
D24	85.83	95.83	82.92	90.00	D65	100.00	100.00	98.75	100.00	D105	63.33	50.00	44.58	53.33
D25	88.75	53.75	48.33	75.83	D66	96.67	90.00	95.00	68.33	D106	44.17	55.83	54.58	51.67
D26	100.00	88.75	99.17	100.00	D67	70.00	53.75	63.75	62.08	D107	52.50	59.58	60.83	55.42
D27	36.67	34.58	34.58	40.42	D68	100.00	99.58	100.00	100.00	D108	37.50	28.75	29.58	32.92
D28	95.42	86.67	97.50	88.75	D69	42.50	39.17	44.17	36.25	D109	78.75	73.75	76.67	76.25
D29	72.08	60.00	64.58	67.50	D70	49.17	45.42	57.50	79.17	D110	87.92	78.75	75.42	62.92
D30	33.75	23.75	33.75	50.42	D71	42.92	45.83	68.75	72.50	D111	90.83	90.42	62.92	91.67
D33	77.92	72.50	57.50	92.50	D72	47.50	48.75	75.83	79.17	D112	61.67	50.42	58.33	63.33
D34	99.17	92.92	98.33	100.00	D73	66.67	51.67	57.92	57.08	O1	100.00	100.00	100.00	92.08
D35	98.33	82.92	68.33	89.17	D74	78.75	85.00	85.00	94.58	O2	100.00	100.00	98.75	100.00
D36	49.17	57.08	30.00	31.67	D75	95.42	86.67	93.75	98.75	O3	98.75	75.42	93.33	97.50
D37	100.00	78.75	79.58	99.58	D76	99.17	96.25	95.83	92.08	O4	100.00	100.00	100.00	100.00
D38	46.67	31.67	29.17	42.08	D77	100.00	100.00	100.00	100.00	O5	86.25	83.33	100.00	93.33
D39	39.58	24.17	27.92	60.00	D78	97.50	93.33	85.42	88.33	O6	100.00	100.00	100.00	100.00
D40	52.08	56.67	67.50	41.67	D79	100.00	100.00	100.00	97.92	O7	71.25	68.75	52.08	57.50
D41	78.75	68.33	52.08	58.33	D80	100.00	85.83	85.83	91.67	**Avg.**	**74.37**	**68.70**	**69.41**	**73.18**

TABLE 2
CPU TIMES (ON A SUN SPARC20 WITH ONE PROCESSOR) AND FEATURE VECTOR LENGTH FOR THE VARIOUS TEXTURE FEATURES.
GABOR FEATURES ARE COMPUTED IN MATLAB AND ALL THE OTHERS ARE WRITTEN IN C LANGUAGE.

	Gabor Features		PWT	TWT	MRSAR
	Full feature	Adaptive (4 filters)			
Feature Extraction Time	9.3 sec.	2.3 sec	1.3 sec	2.3 sec	34.0 sec
Feature Vector Length	48 (24 x 2)	8 (4 x 2)	24 (12 x 2)	104 (52 x 2)	15
Searching and Sorting Time	1.02 sec	0.1 sec	0.98 sec	1.70 sec	0.70 sec

ginally better (69.4%) than the PWT features (68.7%). Fig. 5 shows a graph illustrating this retrieval performance as a function of number of top matches considered. In summary,

- In general, feature components corresponding to higher frequencies have better discriminating performance. However, decomposing the HH band in the tree-structured representation often leads to a decrease in performance, indicating that these features are not very robust.
- Experiments with different wavelet transforms indicate very little variation in performance with respect to the choice of filters [10].
- The marginal improvement of the TWT features comes at the expense of having a much larger feature vector, which adds to the overhead associated with indexing and search.
- It is important to explore different similarity measures for each of the different sets of features. For example, using the Mahalanobis distance instead of the Euclidean distance improved the performance from 64% to 73% for the MR-SAR features. Normalized Euclidean distance worked better for all the others.
- For Brodatz images, best results using the Gabor features were obtained using four scales and six orientations within each scale. Note that rotation and scale invariance is not addressed.

Table 2 provides the CPU times for feature extraction and sequential search of the database. The computations involving Gabor features are performed in MATLAB where as the other feature computations are implemented in C language. In terms of feature computation time, the MR-SAR is the most expensive.

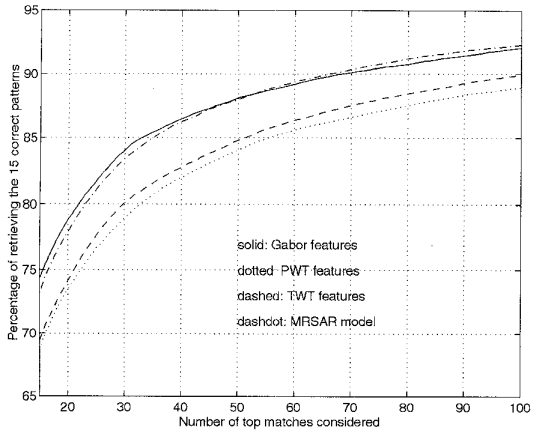


Fig. 5. Retrieval performance according to the number of top matches considered.

4.2 An Image Browsing Example

Query based on texture properties will have many applications in image and multimedia databases. Here, we describe with an example our current work on incorporating these features for browsing large satellite images and air photos. This work relates to the UCSB Alexandria digital library project [21] whose goal is to create a digital library of spatially indexed data such as maps and satellite images. Typical images in such a database range from few megabytes to hundreds of megabytes, posing challenging problems in image analysis and visualization of data. Content based retrieval will be very useful in this context in answering queries such as "Retrieve all LANDSAT images of Santa Barbara which have less than 20% cloud cover," or "Find a vegetation patch that looks like this region."

We are currently investigating the use of texture primitives to accomplish rapid content based browsing within an image or across similar images. Fig. 6 shows an example of browsing 5,248 × 5,248 air photos. The original image is analyzed in blocks 128 × 128 pixels and the texture features are computed and stored as image "meta-data." The user can select any position and use that pattern to search for similar looking regions. Our current work is on incorporating simple texture based segmentation schemes into this browsing thus allowing arbitrarily shaped regions into the analysis.

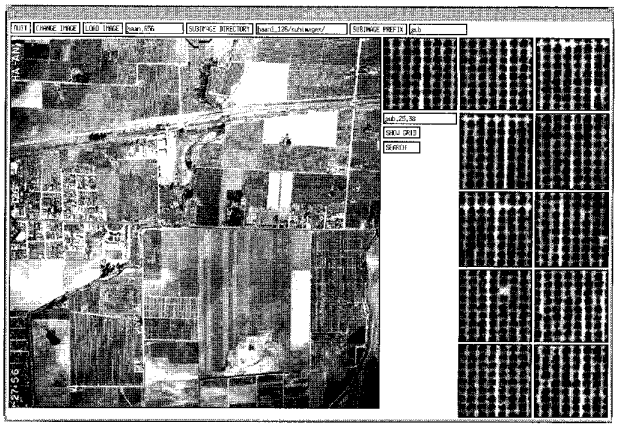
4.3 Discussions

A Gabor wavelet based texture analysis scheme is proposed and its application to image databases is demonstrated. A comprehensive performance evaluation of the method is given using a large number of textures and a comparison with some of the well known multiresolution texture classification algorithms is made. Further, a novel adaptive filter selection strategy is suggested to reduce the image processing computations while maintaining a reasonable level of retrieval performance. The experimental results indicate that these Gabor feature are quite robust. Rotation and scale invariance is important in many applications and our preliminary results on rotation invariant classification [7] using Gabor features are very encouraging.

Finally, a note on similarity measures. It is widely acknowledged that this is an important but a difficult problem. Our initial results using simple hybrid neural network learning algorithms appear very promising in the context of learning similarity [11], [12].

ACKNOWLEDGMENTS

This research was partially supported by National Science Foundation grant IRI-9411330 and by NASA under grant number NAGW-3951. We thank Professor Picard and F. Liu at MIT for providing the software for the MR-SAR features and C. Fischer at UCSB for the air photos.



(a)

(b) (c) (d)

Fig. 6. An example of browsing a large air photo using Gabor texture features. Part (a) shows the down-sampled version of the image and the retrieval results on a vegetation pattern, (b) the query pattern contains a part of the highway and retrievals are all from highway segments, (c) shows the region containing some buildings (center of the image toward the left), (d) shows one interesting result where the query contains a number marked on the image (lower left corner), and the top matches also contain similar patterns.

REFERENCES

[1] A.C. Bovic, M. Clark, and W.S. Geisler, "Multichannel Texture Analysis Using Localized Spatial Filters," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 55-73, Jan. 1990.

[2] P. Brodatz, *Textures: A Photographic Album for Artists & Designers*. New York: Dover, 1966.

[3] T. Chang and C.-C.J. Kuo, "Texture Analysis and Classification with Tree-Structured Wavelet Transform," *IEEE Trans. Image Processing*, vol. 2, no. 4, pp. 429-441, Oct. 1993.

[4] J.G. Daugman, "Complete Discrete 2D Gabor Transforms by Neural Networks for Image Analysis and Compression," *IEEE Trans. ASSP*, vol. 36, pp. 1,169-1,179, July 1988.

[5] J.G. Daugman, "High Confidence Visual Recognition of Persons by a Test of Statistical Independence," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1,148-1,161, Nov. 1993.

[6] I. Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," *IEEE Trans. Information Theory*, vol. 36, pp. 961-1,005, Sept. 1990.

[7] G.M. Haley and B.S. Manjunath, "Rotation Invariant Texture Classification Using the Modified Gabor Filters," *Proc. IEEE ICIP '95*, vol. I, pp. 262-265, Washington D.C., Oct. 1995.

[8] M. Lades et al., "Distortion Invariant Object Recognition in the Dynamic Link Architecture," *IEEE Trans. Computers*, vol. 42, no. 3, pp. 300-311, Mar. 1993.
[9] F. Liu and R.W. Picard, "Periodicity, Directionality, and Randomness: Wold Features for Image Modeling and Retrieval," MIT Media Lab Technical Report no. 320, Mar. 1995.
[10] W.Y. Ma and B.S. Manjunath, "A Comparison of Wavelet Features for Texture Annotation," *Proc. IEEE Int'l Conf. Image Processing '95* , vol. II, pp. 256-259, Washington D.C., Oct. 1995.
[11] W.Y. Ma and B.S. Manjunath, "Image Indexing Using a Texture Dictionary," *Proc. SPIE*, vol. 2,606, pp. 288-296, Philadelphia, Oct. 1995.
[12] W.Y. Ma and B.S. Manjunath, "Texture Features and Learning Similarity," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, San Francisco, June 1996.
[13] B.S. Manjunath and R. Chellappa, "A Feature Based Approach to Face Recognition," *Proc. IEEE Conf. CVPR '92*, pp. 373-378, Champaign, Ill., June 1992.
[14] B.S. Manjunath and R. Chellappa, "A Unified Approach to Boundary Detection," *IEEE Trans. Neural Networks*, vol. 4, no. 1, pp. 96-108, Jan. 1993.
[15] B.S. Manjunath, C. Shekhar, and R. Chellappa, "A New Approach to Image Feature Detection with Applications," *Pattern Recognition*, Apr. 1996.
[16] B.S. Manjunath and W.Y. Ma, "Texture Features for Browsing and Retrieval of Image Data," CIPR TR 95-06, July 1995 (available on www at http://vivaldi.ece.ucsb.edu).
[17] J. Mao and A. Jain, "Texture Classification and Segmentation Using Multiresolution Simultaneous Autoregressive Models," *Pattern Recognition J.*, vol. 25, no. 2, pp. 173-188, 1992.
[18] W. Niblack et al., "The QBIC Project," *Proc. SPIE*, vol. 1,908, pp. 173-181, Feb. 1993.
[19] A. Pentland, R.W. Picard, and S. Sclaroff, "Photobook: Tools for Content Based Manipulation of Image Databases," *Proc. ICASSP '93*, vol. V, pp. 161-164, Minneapolis, Apr. 1993.
[20] R.W. Picard, T. Kabir, and F. Liu, "Real-Time Recognition with the Entire Brodatz Texture Database," *Proc. IEEE Conf. CVPR '93*, pp. 638-639, New York, June 1993.
[21] The Alexandria Digital Library Project, http://alexandria.sdc.ucsb.edu.

Retrieving Multispectral Satellite Images Using Physics-Based Invariant Representations

Glenn Healey and Amit Jain

Abstract—We present a set of algorithms and a search strategy for the robust content-based retrieval of multispectral satellite images. Since the property of interest in these images is usually the physical characteristics of ground cover, we use representations and methods that are invariant to illumination and atmospheric conditions. The representations and algorithms are derived for this application from a physical model for the formation of multispectral satellite images. The use of several representations and algorithms is necessary to interpret the diversity of physical and geometric structure in these images. Algorithms are used that exploit multispectral distributions, multispectral spatial structure, and labeled classes. The performance of the system is demonstrated on a large set of multispectral satellite images taken over different areas of the United States under different illumination and atmospheric conditions.

Index Terms—Image database,image retrieval, color constancy, satellite images, color, machine vision, texture, computer vision, recognition.

1 Introduction

MULTISPECTRAL remote-sensing data are being used for an increasing number of applications in a diverse set of fields including agriculture, geology, mapping, water resources, and environmental science. The volume of satellite data that is available for such applications is staggering. A Landsat thematic mapper, for example, generates seven band images using three visible and four infrared regions of the spectrum. Even though the Landsat provides coarser spatial resolution than many other remote-sensing satellites, a single image corresponding to a 170 km by 185 km region of earth requires over 200 Mbytes of storage and the satellite generates about 5,000 images per week. It has been predicted that in a few years the amount of data originating from remote-sensing satellites will reach a terabyte per day [2].

Given this large volume of data, effective tools for image access are essential to allow the information in the database to be fully exploited. Image database systems traditionally access images using keywords or text associated with the images. Unfortunately, it is often difficult to assign textual descriptions to images and consequently text-based queries often fail. In addition, the task of manually annotating the current volume of satellite imagery would involve a large amount of time and expense.

A recent trend in retrieval from image databases has been to allow queries based on image content [12], [11], [15], [1]. Using this paradigm, a user can specify a search using image properties such as shape, color, or texture. In most cases, the user will not specify numerical values of these properties, but rather will present the system with example images. The system will compute features from the example and use these features for database indexing. In most cases, a metric is defined on these features that is intended to model perceptual similarity.

• *The authors are with the Department of Electrical and Computer Engineering, University of California, Irvine, CA 92717. E-mail: {healey, amit}@ece.uci.edu.*

Manuscript received July 5, 1995.
Recommended for acceptance by R.W. Picard.
For information on obtaining reprints of this article, please send e-mail to: transpami@computer.org, and reference IEEECS Log Number P96052.