Orientation radiograms for indexing and identification in image databases

S. Michel[1], B. Karoubi[2], J. Bigün[1] and S. Corsini[3]

[1] Signal Processing Laboratory, Swiss Federal Institute of Technology, CH-1015 Lausanne, Switzerland

[2] CREATIS,Research associated to CNRS (#1216) and affiliated to INSERM, Lyon, France

[3] Bibliothèque Cantonale et Universitaire Lausanne, CH-1015 Lausanne/Dorigny, Switzerland

ABSTRACT

Archival of images in databases, enabling further study with respect to their contents, is at our focus of attention. The major difficulties are i) the processing of a large number of images, ii) that the steadily growing number of images increase the complexity of the pattern recognition problems to be solved. We propose orientation radiograms, to be used as image signatures for shape based queries. These are the projections of a set of orientation decomposed images (here 6) to axes whose directions change synchronously with the orientation bands at hand. The peaks in the radiograms represent long edges or lines which are very important for the human when he recognizes or compares images. We present the results of experiments based on approximately 400 images in an application concerning typographic ornament images. Also is presented a comparative study comprising classical moment invariants.

1 Orientation radiograms

Image databases in general contain many images so that the discrimination task becomes very difficult as the global complexity of the shapes to be described is very high, meaning that low dimensional image signatures are not sufficiently powerful for shape description,[7, 8]. We decompose an image into its iso-orientation images (briefly orientation images) i.e. each component image of the decomposition corresponds to the response of a filter tuned to one orientation. All other orientations are suppressed which means that an "X-ray" in the direction of the "pass" orientation will ideally not *intersect* any edge element since otherwise that would mean the existence of other orientations than that of the tuned in a component image. Of course the X-ray will encounter edge elements, but it will "traverse" them, and not intersect them. Summing the responses of the component image along such parallel rays yields a one dimensional image, which we will call an *orientation radiogram.*

In order to illustrate the representation of shape by orientation radiograms, we consider the image of a rectangle, see the figure 1, and assume that we have 6 orientations in our decomposition. Only 2 orientation images have non-zero responses, corresponding to the orientations of 0^0 and 90^0. If we consider the projections of the orientation images along parallel rays corresponding to the tuning of the image, we obtain 6 1D signals, [Fig.2]. These signals allow to characterize the rectangle with respect to another object.

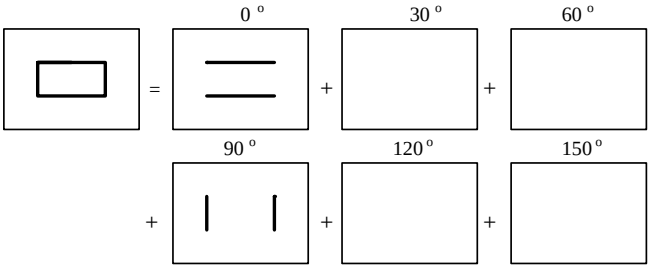

Figure 1: Schematic decomposition of a rectangular image based on 6 privileged orientations.

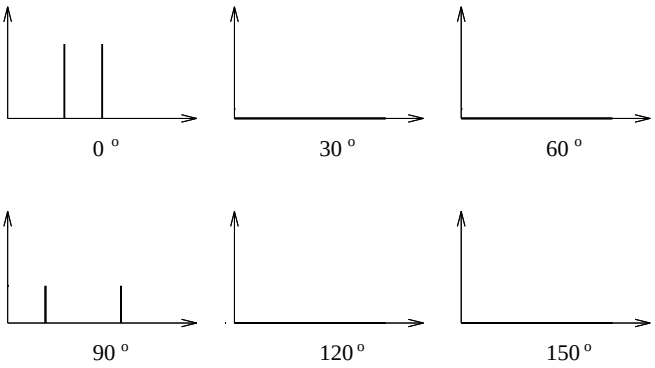

Figure 2: Schematic projections of the previous decomposition.

2 Orientation decomposition

Two practical tools to perform the orientation decompositions which can be applied to the database population

stage i.e. each time an image is added, are linear symmetry computation [1], and Gabor decomposition, [5]. We will describe the first approach as it is based on, but not fully described in [2].

An efficient method to get the local orientation information is to compute linear symmetry of the images. This method computes for each pixel a vector, linear symmetry, where its argument gives the dominant orientation of the local neighborhood and its magnitude stands for the certainty. By construction, only neighborhoods with a dominant orientation give high certainty. The argument should not be confused with that of the Gabor phase, which represents the edge type. It can be shown that, the linear symmetry computation fits an optimal line to the local power spectrum which is given by

$$z = \left(\nabla \vec{f}\right)^2 * m \tag{1}$$

where f, ∇f, $*m$ represent the original (grey level) image, the complex image $f_x + f_y i$, convolution with an averaging filter, respectively. The next step is to obtain the six orientation images, $r_l(\vec{x})$ which we do as:

$$r_l = |z| \cos^2(\theta_l - \varphi),\ \theta_l = \frac{l\pi}{6}, l = 0, ..., 5 \tag{2}$$

where θ_l corresponds to a tuned orientation, φ represents the argument of the linear symmetry vector, z. It can be shown that r_l corresponds to the inertia of the local power spectrum w.r.t. the direction represented by θ_l. It is maximum, $|z|$, when $\theta_l = \varphi + n\pi$, and it is minimum, 0 when $\theta_l = \varphi + \frac{pi}{2} + n\pi$. However, the decrease in the response when the linear symmetry orientation deviates from that of the tuned one θ_l, is not controllable above. Therefore we suggest:

$$r_l(\vec{x}) = |z|\ (\exp(\beta \cos^2(\theta_l - \varphi)) - 1) \tag{3}$$

in order to steer the orientation selection sensitivity through the factor β. $|z|$ is normalized linearly, so that it varies between 0 and 1 in *one* orientation image. Consequently, r_l can be rescaled linearly so that it is in $[0, 1]$, when a suitable β is determined experimentally. The implementation of projection computation is simplified if each image is suitably rotated followed by horizontal or vertical summation. However, oblique scanning and summation of orientation images is also possible, as an alternative.

3 Fourier Coefficients distance

The maxima in the radiograms are important since they correspond to long edges. Since a set of ten Fourier coefficients, leaving out the DC component, seem to preserve the peaks, we truncated the number of the complex Fourier coefficients at 10. We take the real part and the absolute value of the imaginary part of the Fourier coefficients in order to obtain invariance in a flip of the image (4). That is, since the radiograms are real functions we have,

$$\begin{aligned} \Re(F(\nu)) &= \Re(F(-\nu)) = \Re(G(\nu)) \\ \Im(F(\nu)) &= -\Im(F(-\nu)) = -\Im(G(\nu)) \end{aligned} \tag{4}$$

where $F(\nu)$ represents the Fourier coefficient of the ornament and $G(\nu)$ the Fourier Transform coefficient of the flipped image. The coefficients allow to build 6 feature vectors, each representing one radiogram, that will be used when comparing the ornaments.

The comparison between the feature vectors of a scanned image and those corresponding to the images in the database is made by computing a distance which is the sum of the Euclidean distances, here 6, between the corresponding feature vectors. The sorted distance allows to find the classes that match the scanned image. That is, we do not have to attribute a class but we have to display images that have similarities with the scanned image. False misses are not allowed, but a limited number of false matches is tolerated. An example of a query showing the 5 closest classes is shown in figure 3.

4 Results

The scanned images must be "recognized" even if they have variations in grey scale, translation and rotation. The Fourier Transform is invariant to translation (cyclical). The rotation invariance can be dealt with the rotation of the image before the start of the orientation decomposition. This can be done by using the first few orders of the moments. We note that a zero order moment represent the image surface, the first order gives the centroid of the picture, and the second order characterizes the main direction (orientation). Thus the global orientation variances can be compensated for.

The experimental database is made of 39 ornaments scanned from old books in the possession of the University Library of Lausanne, [3]. From this set, 37 classes can be extracted since two pairs represent the same class, but digitally differ. [OR36] and [OR25] are therefore missing from the misclassification tables below. Due to the fact that books are small and old, ornaments are photocopied and then scanned. Of course, this manipulation introduces further noise. To measure the ornament dissimilarities, for the purposes of this work, we have generated synthetic pictures from the original set because we do not dispose many images of the same class. A training set has been built in order to create reference features. The mean value of the radiogram Fourier coefficients of 5 geometric transformations are taken as reference feature vectors. These transformations are, 1) 2^o rotation 2) -4^o rotation and flip up-down 3) 90^o rotation 4) -2^o rotation and flip left-right 5) gamma correction with $\gamma = 3$. Thus there were 195 images in the training set.

To evaluate the validity of the class identification we have chosen to make 5 new transformations on the ori-

	bc	26	28	33	tot	rate
1	5	0	0	0	5	1.00
2	5	2	3	0	5	0.50
3	5	0	0	0	5	1.00
4	5	0	0	0	5	1.00
5	5	0	0	0	5	1.00
6	5	0	0	0	5	1.00
7	5	0	0	0	5	1.00
8	5	0	0	0	5	1.00
9	5	0	0	0	5	1.00
10	5	0	0	0	5	1.00
11	5	0	0	0	5	1.00
12	5	0	0	0	5	1.00
13	5	0	0	0	5	1.00
14	5	0	0	0	5	1.00
15	5	0	0	0	5	1.00
16	5	0	0	0	5	1.00
17	5	0	0	0	5	1.00
18	5	0	0	0	5	1.00
19	5	0	0	0	5	1.00
20	5	0	0	0	5	1.00
21	5	0	0	0	5	1.00
22	5	0	0	0	5	1.00
23	5	0	0	0	5	1.00
24	5	0	0	0	5	1.00
26	5	0	0	0	5	1.00
27	5	0	0	0	5	1.00
28	5	0	0	0	5	1.00
29	5	0	0	0	5	1.00
30	4	0	0	1	4	0.80
31	5	0	0	0	5	1.00
32	5	0	0	0	5	1.00
33	5	0	0	0	5	1.00
34	5	0	0	0	5	1.00
35	10	0	0	0	10	1.00
37	5	0	0	0	5	1.00
38	5	0	0	0	5	1.00
39	5	0	0	0	5	1.00

Table 1: Confusion matrix using Fourier coefficients, linear symmetry and difference vectors classification

	bc	01	03	10	11	13	16	18	19	20	21	23	26	28	34	38	39	tot	rate
1	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
2	5	0	0	0	0	0	0	0	0	0	3	0	0	0	1	0	1	5	0.50
3	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
4	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
5	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
6	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
8	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00
9	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00
10	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00
11	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
12	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00
13	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
14	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
15	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
16	0	0	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00
17	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00
18	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0.00
19	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0.00
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0.00
21	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0.00
22	0	0	0	0	0	0	0	1	0	0	0	0	4	0	0	0	0	0	0.00
23	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
24	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	1	0	0	0.00
26	0	0	0	0	0	0	0	0	0	1	4	0	0	0	0	0	0	0	0.00
27	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0.00
29	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
30	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
31	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
32	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
33	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	1.00
35	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.00
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0.00
38	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	4	0	0	0.00
39	0	0	0	0	0	0	5	0	0	0	0	1	0	0	0	0	0	0	0.00

Table 2: Confusion matrix using Reddi moment invariants and difference vectors classification

ginal set in order to use them as a test set. These transformations, resulting in 195 test images which are different than those in the training set, are, 1) 5^o rotation 2) -3^o rotation 3) -90^o rotation 4) flip left-right 5) flip up-down. The table 1 shows the misclassification of our method, using orientation radiograms computed with linear symmetry. Table 2 allows to compare our result with a classical method. In radiogram features, the recognition is fairly robust despite various synthetic (gray level as well as geometric transformations, including flipping) and natural disturbances, which in turn suggests that the used metric for image distance measurements captures well what the experts observe. We can see this in table 1 at line 35 where the transformed images of OR36 are attributed to the class 35 since the value 10 is in the diagonal cell. For the class 2 we do not have perfect merging but the attributed classes are very similar. The radiogram Fourier coefficient classification with the linear symmetry decomposition resulted in the rate of 97% correct classifications, which is obtained by dividing the trace by the sum of all elements of the confusion matrix, which should be compared to Reddi's moment invariants based features [9], 49 %, used for pattern recognition purposes. We chose Reddi's invariants for comparison since they performed slightly better than Hu's, [4], and Maitra's invariants, [6], on pattern recognition tasks involving ornament images. We could confirm the known deficiency of the moment features, namely that they over emphasize the periphery details on the cost of the central details in an image.

Figure 3 illustrates a query. An input image, on which 5^o rotation, flip left-right and gamma correction with $\gamma = 0.4$ has been performed, is used to look for the five nearest classes. The algorithm has found the real class as well as the only other 3 images belonging to the same family. The fifth image does not belong to the family but is displayed as the closest since 5 records were requested. The latter illustrates very well the difference of the database problem, where a false acceptance is tolerated while a false rejection is not accepted, from a pattern recognition of, for example, bank notes.

5 Conclusion

Our study suggests that orientation radiograms are suitable to efficiently describe the content of images for image database queries. Since radiograms are 1D signals with many peaks, a limited set of Fourier coefficients allows to characterize them well. Besides that the amount of the signature data is small, this representation is very interesting on its own, because invariance to flipping can be easily obtained.

The method tested on ornaments can be without modification be applied to general gray images as binarization of images is neither required nor carried out in

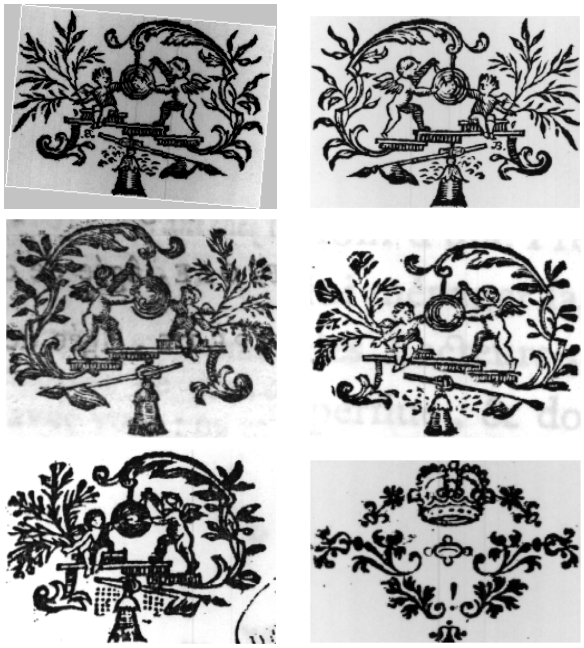

Figure 3: Query result made on the top image. Top left: input image, top right: nearest image and the following.

our particular test bed. However, experiments, in this respect need to be carried out.

An extension of the present concept, on which we are currently working, is queries on contents of sub images.

Acknowledgment

We want to thank M. Stefan Fischer at the Signal Processing Laboratory (LTS), for his wide help in computer science and software development.

References

[1] J. Bigün. Frequency and orientation sensitive texture measures using linear symmetry. *Signal Processing*, 29:1–16, 1992.

[2] J. Bigün and J. M. H. Du Buf. Symmetry interpretaton of complex moments and the local power spectrum. *Journal of Visual Communication and Image Representation*, 6(2):154–163, June 1995.

[3] S. Corsini. Vers un corpus des ornements typographiques lausannois du 18ème siècle. problèmes de définition et de méthode. In *Ornementation typographique et bibliographie historique, actes du Colloque de Mons*, pages 139–158, August 1988.

[4] M.K. Hu. Visual pattern recognition by moment invariants. *IRE Trans. Information Theory*, pages 179–187., 1962.

[5] A. K. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. *Pattern Recognition*, 24:1167–1186, 1991.

[6] S. Maitra. Moment invariants. *IEEE Proceedings*, 67(4), April 1979.

[7] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, and P. Yanker. The qbic project: Querying images by content using color, texture and shape. In *Proc. SPIE Electronic Imaging: Science and Technology*, San Jose, CA, February 1993.

[8] A. Pentland, R.W. Picard, and S. Sclaroff. Photobook: Tools for content-based manipulation of image databases. In *Proc. SPIE Storage and Retrieval Image and Video Database II*, San Jose, CA, February 1994.

[9] S. S. Reddi. Radial and angular moment invariants for image identification. *Pami*, 3(2):240 – 242, March 1981.