A Case-Based Approach to Image Recognition

Alessandro Micarelli[1], Alessandro Neri[2], and Giuseppe Sansonetti[1]

[1] Dipartimento di Informatica e Automazione, Università degli Studi "Roma Tre"
Via della Vasca Navale, 79, I-00146 Roma, Italia
{micarel,gsansone}@dia.uniroma3.it
[2] Dipartimento di Ingegneria Elettronica, Università degli Studi "Roma Tre"
Via della Vasca Navale, 84, I-00146 Roma, Italia
neri@ele.uniroma3.it

Abstract. In this paper we present a case-based approach to the recognition of digital images. The architecture we propose is based on the "wavelet transform" that has been used for the representation, in the form of old cases, of images already known to the system. The paper also presents our report on a case study in the field of "mobile robots". The described system is capable of analyzing maps obtained from the sensors of a robot, and classifing them as one of the possible "objects" present in the environment in which the robot navigates. The first results we have obtained are encouraging and support the choice of the case-based approach to image recognition using the wavelet transform as a tool for image representation and analysis.

1 Introduction

The purpose of this study is to show how the case-based philosophy may be successfully applied to fields such as the one of image recognition, where, with the exception of rare but significant instances [5,6], the limited availability of literature concerning this subject-matter could misleadingly suggest that CBR is not suitable for possible application in this field. Through our work, we wish to emphasize the advantages obtained in tackling image recognition problems by exploiting a number of already resolved cases, appropriately represented.

In principle, the ideal situation in respect to an image representation would be to emphasize – inside the image-signal – those very zones that are the focus of human vision interest. In this respect, extensive studies have demonstrated that the first elements that attract our attention at a perceptive level are lines and edges, i.e., elements that can be defined by way of measurable quantities on the basis of values given to image signals. Our choice in respect to the representation of the images to be subsequently analyzed, is based on the theory of wavelet transform. During the past decade, the wavelet transform theory has obtained considerable success in various application fields. In fact, numerous signal processing applications relate to this theory, such as image multi-resolution encoding, sub-frequency band encoding for video signal compression and the analysis of non-stationary signals such as geophysical and bio-medical ones.

E. Blanzieri and L. Portinale (Eds.): EWCBR 2000, LNAI 1898, pp. 443–454, 2000.

Our paper is organized as follows: in Section 2, we illustrate a generic type of case-based architecture that we propose for the recognition of a vast range of images. In particular, we describe the representation technique for images in the library of cases and the retrieval algorithm operating on such library in respect to the analysis of new images. In Section 3, we describe a case study relating to the application of our recognition approach to objects of interest in the field of mobile robots. Our final remarks are given in the concluding section.

2 The Wavelet-Based CBR Architecture

The problem we intend to address concerns the recognition of an image represented in digital format and its classification under one of the categories of a set of predetermined categories, each of which is linked to an "object" of interest. As we already mentioned earlier, the idea of the approach presented herein consists in taking advantage of the tools officered by the wavelet transform theory, appropriately adapted for use in a case-based architecture. Figure 1 illustrates a diagram that summarizes the mechanism we propose. Let's assume we have a new image to classify, in bitmap format (termed "new digital image" in the figure), the initial operation performed relates to the extraction of "features" relevant for the purposes of the image recognition. The result of the pre-processing will be the representation of the image based on the wavelet transform. This representation constitutes the "new case" of the proposed CBR system. The retrieval module shown in the figure will effect a research in the case library containing the old cases, with a <*problem representation*, *solution*> structure, which in this specific case will be <*image representation*, *object represented by the image*>. The solution given in the old case can therefore be seen as a pointer to a new library, which we can name "Library of Objects", containing the objects (i.e., "categories") that could be present in the images to be analyzed. The "Recognized Object" is at this point taken into consideration by the environment where the system is operating.[1] This object, which constitutes the old solution of the case retrieved from the Case Library, will also be considered as the solution of the new problem (basically, there is no need for an adaptation of the old solution to suit the new case) and if the environment accepts it, the pair <*New Image*, *Recognized Object*> can be inserted as a new case in the Case Library. The most problematic aspects to be worked out are therefore the actual modalities for the representation of images as cases and the definition of a similarity metric to be used for the retrieval phase in respect to the Case Library.

In principle, given a set of known cases, a simple approach to retrieve the one that better matches the situation under examination could be based on the computation of the cross-correlation between the new image and templates

[1] For a better understanding of this idea, the environment of the case-study presented in Section 3, concerning a vision problem in respect to mobile robots, can be seen as constituted by the robot and the human expert who is supervising the training of the robot itself.

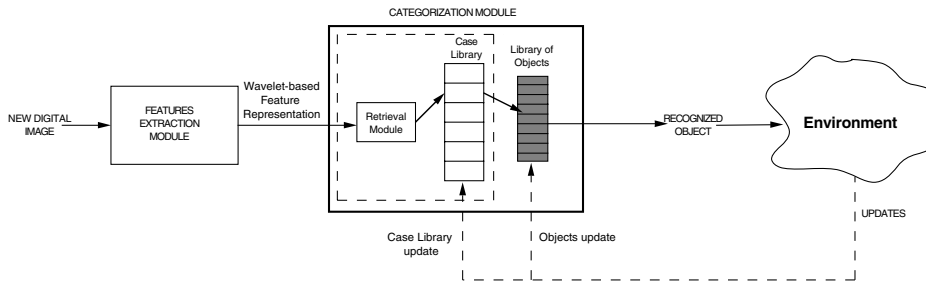

Fig. 1. CBR Approach to Image Recognition

extracted from the old cases. However, although conceptually simple, this approach is rather ineffective. In fact, the ability of handling cases derived from those contained in the library by applying to them simple geometric transformations, such as translation, rotation or scaling, would require the computation of the said metric for infinitely many variants. In addition, direct evaluation of the cross-correlation function would imply storage and on-line processing of the related images. To manage this computational burden, a viable approach for case indexing could rely on the use of some kinds of "invariant" signatures, investigated in the past decades in the context of classical pattern recognition [8,12]. The major class of rotation invariant image classification techniques, originally introduced in the field of optical processing (see for instance [8,23]), is based on the extraction of the dominant circular harmonic components obtained expanding a given image in its angular Fourier's series [12,15,18,9,10,11,24]. This circular harmonic decomposition can be further combined with scale invariant representations like the Fourier-Mellin Transform (FMT) to devise rotation and scale invariant pattern recognition algorithms [22]. To preserve the ability of the system of inferring the current situation based on the whole set of collected cases, at the expenses of a reasonable computational complexity, in our architecture we resort to a general purpose rotation and scale invariant indexing tool based on the cited mathematical representations, proposed in [11], and [12]. In essence, the index associated to each pattern stored in the case dictionary is constituted by the dominant components of the Laguerre-Gauss Circular Harmonic Wavelet (CHW) decomposition, computed by means of a bank of linear operators.

Since retrieval requires the evaluation of the distance between the new and the old case, a proper selection of the color space employed for the image representation is crucial with respect to the effectiveness of the metric, in terms of discrimination capability. Here, inspired by the behavior of the human visual system, we propose to use the CIE[2] $L^*a^*b^*$ perceptually uniform color space. In fact, at a first glance, the measure Δs of the difference perceived by a human observed between the stimuli produced by two uniform images satisfies the relationship

[2] CIE: Comitat International d'Éclarage.

$$(\Delta s)^2 = (\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2$$

Then, given the i-th image pattern related to the i-th case, let

$$\mathbf{g}^{(i)}(\mathbf{x}) = \begin{bmatrix} L^{*(i)}(\mathbf{x}) \\ a^{*(i)}(\mathbf{x}) \\ b^{*(i)}(\mathbf{x}) \end{bmatrix}$$

be its representation in the CIE $L^*a^*b^*$ color space, where $\mathbf{x} = [x_1\ x_2]^T \in \Re^2$ denotes the coordinates of a point of the real plane. To build the index associated to it, we represent $\mathbf{g}^{(i)}(\mathbf{x})$ by means of its Laguerre-Gauss transform (LGT), that constitutes the polar counter-part of the Hermite Transform proposed in [15]. Let us observe that the elements of the basis employed for the image representation are tuned to specific, visually relevant, image features. In fact, LG functions of the first order are matched to edges, while second order LGF are matched to lines and third and fourth order LGF to forks and crosses respectively. Thus, the current case base indexes are build upon the kind of features that the human visual system is supposed to use too. In addition, due to this inherent capability, the representation of $\mathbf{g}^{(i)}(\mathbf{x})$ is usually rather sparse. But, use of the LGT is not only effective for the compactness of the associated index. One of the most relevant benefits of this representation is the availability of fast algorithms for the evaluation of the L^2-norm of the difference between the image at hand and the patterns stored in the case dictionary, irrespective of their location and orientation [12]. As detailed in appendix, given a digital image, its representation stored in the Case Library is constituted by a sparse matrix $\mathbf{C}_i$ whose elements are obtained by shrinking the LGT coefficients given by (1).

Then, as detailed in [16], the "rotation invariant" similarity metrics between the actual image $\mathbf{f}(\mathbf{x})$ with LGT shrinked coefficients $\mathbf{D}_{n,k}(\mathbf{m})$ and the i-th pattern with LGT coefficients $\mathbf{C}^{(i)}_{n,k}(\mathbf{x}_0)$, associated with the L^2-norm is

$$\mathbf{\Delta}(\mathbf{f}, i) = \min_{\mathbf{b}_i \in \mathbf{M}} \sum_{n=-\infty}^{+\infty} \mathbf{M}^{(i)}_n(\mathbf{b}_i) cos\left[n\hat{\varphi}_i - \eta^{(i)}_n(\mathbf{b}_i)\right] \tag{1}$$

where

$$M^{(i)}_n(\mathbf{b}_i) e^{j\eta^{(i)}_n(\mathbf{b}_i)} \cong \sum_k \overline{\mathbf{C}}^{(i)}_{n,k}(\tilde{\mathbf{x}}_0 + \mathbf{m}) \mathbf{D}_{n,k}(\tilde{\mathbf{x}}_0 + \mathbf{b_i} + R_{-\hat{\varphi}_i}\mathbf{m})$$

and the actual $\hat{\varphi}_i$ image orientation is estimated as specified in ther appendix. In addition we shrink the expansion coefficients, so that:

$$C_i(n, k, \mathbf{m}) = \begin{cases} 0 & \text{if } C^{(i)}(n, k, \mathbf{m}) < \gamma \\ C^{(i)}(n, k, \mathbf{m}) & \text{otherwise} \end{cases}$$

In Figure 2 we only sketch the algorithm for the similarity metrics in pseudocode.

```
Function Distance-Δ(C_i , D) returns Δ
    inputs : C_i ; the i-th old case
             D; the new case
    for each b in lattice-M do
       begin
          for each n in [0, N] do
             begin
                x_n = 0
                for each k in [0, K] do
                   x_n = x_n + inner-product(C_i(n, k, 0), D(n, k, b))
             end
             φ̂ ← best-estimate(x)
             R_−φ̂ ← rotation-matrix(φ̂)
             Δ(f, i, b) ← 0
             for each n in [0, N] do
                begin
                   for each k in [0, K] do
                      for each m in lattice-M do
                         Δ(f, i, b) = Δ(f, i, b) +
                         inner-product(C_i(n, k, m, D(n, k, b + R_−φ̂_i m))
                end
       end
    Δ ← min_{b∈M} Re[Δ(f, i, b)]
    returns Δ
```

Fig. 2. The Distance Metrics Algorithm

3 A Case Study: Image Recognition for an Autonomous Mobile Robot

In this section we present an application of the approach proposed in respect to the problem of a robot autonomous navigation in an environment of which it has no prior knowledge [19,20,3].

The model we consider (NOMAD 200 by Nomadic Technologies) was equipped with a ring comprising 16 ultrasonic rangefinders (POLAROID), with an independent rotation ability; in order to increase measured accuracy, at each position of the environment the ring is rotated twice by an angle of 7.5°: as a result, 16 x 3 = 48 range readings are obtained at each measure point. The working principle of ultrasonic sensors is simple: they beam a packet of ultrasonic waves, identify the resulting echo and, based on the measurement of the time elapsed from the moment of emission to the resulting echo, provide the range between the robot and the detected obstacle [21].

In addition these devices are particularly common since – besides their decidedly easy-to-use property – they guarantee a satisfactory compromise between cost and accuracy, an important aspect that cannot be underestimated, espe-

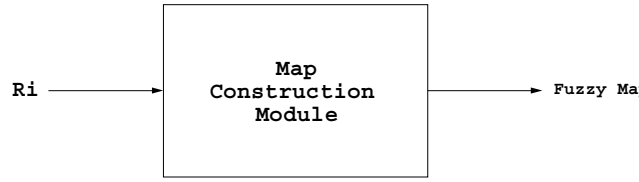

Fig. 3. Map Building Module

cially if one considers the idea of subsequently marketing the prototype. This is the reason behind the frequent choice of cheaper sensing devices instead of more sophisticated sensors such as laser or infrared ones, or even video cameras. Obviously this choice implies other consequences, indeed, by using ultrasonic sensors one must be prepared to face a sizeable amount of uncertainty that they entail in the measurement process [14]. In fact, in principle, the map building phase of the environment surrounding the robot could be viewed as a decision problem, which consists in establishing, for each point in the area of interest (the universal set U), which of the complementary subsets (the empty space or the occupied space, two crisp sets giving a partition of U) it belongs to. The significant amount of uncertainty given by an ultrasonic measurement process somewhat compromises its support of such a clear-cut separation. Therefore the map building module could be seen more efficient if developed as shown in Figure 3.

In other words, having as input the sequence of range readings R_i, collected at known sensor locations (obviously the achieved representation of the environment is less coarse if the measurement points are numerous and well distributed), and as output a fuzzy set, so-called the "fuzzy map", which supplies information on the collision risk in respect to each point in the environment.

Previous studies in this field [19,20] have emphasized the advantages gained from the use of Fuzzy Logic techniques and, in particular, of the Fuzzy Set theory, in respect to sensor-based navigation architectures. In fact, the origin of this theory, stemming from less conservative axioms as opposed to more conventional approaches (Probability Theory and Belief Measures), makes it a flexible and thoroughly efficient tool for the development of an environment representation on the basis of uncertain data [7]. Furthermore, its contribution is particularly relevant in those cases, far from remote, where the measurement is polluted by random and irregular values (outliers) [2,25,13].

We therefore opted for an intrinsically uncertain representation of the environment thus obtaining a gray level map such as the one shown in the figure, where the membership function m (which may assume any real value within the interval [0,1]) quantifies the possibility of each position being taken up by an obstacle; in this particular case the environment surrounding the robot was subdivided in a matrix with 40 x 40 square cells, their side $d = 0.1$ m, as shown in Figure 4 (representing a crossing to the right).

Interestingly, the result is a representation similar to an "occupancy grid" usually obtained using stochastic techniques [4]; inside this grid, the darker zones

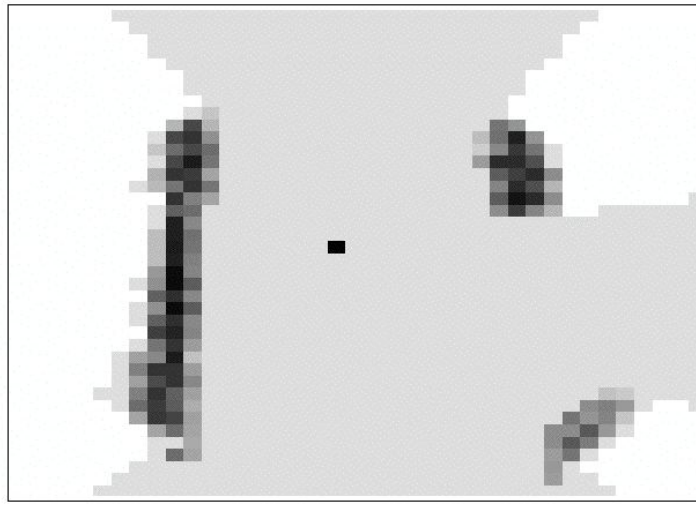

Fig. 4. Fuzzy Map Processed by the System

are the ones where the respective value m is higher, i.e., they reflect a greater level of membership to the fuzzy set of the dangerous cells, whereas the white areas correspond to the cells belonging to unexplored zones, i.e., for which the sensors have not provided any data.

In specific applications, as the robot mission control, used as test bed of the proposed CBR application, the description of the occupancy grid given by the LGT is still too rich, and further simplifications are possible. More in detail, the strategy selection for the short term robot guidance requires a description of the location of obstacles and constraints of the actual scenario. On the other hand, as shown in Figure 4, the low cost sensors actually employed for the acoustic imaging task, produce very rough, noisy images. As enlightened by the experimental activity, a viable approach to a sub-optimal, low complexity guidance and control system consists in replacing the full 2-D description with the polar map of the angular location of obstacles which are in the robot vicinity, i.e., whose distance is less than a predefined value [19,20]. Due to the mentioned sensor limitations, we are able to draw just a fuzzy polar map (see Figure 6). Nevertheless, we still require the ability of the CBR system to retrieve a proper strategy even when the actual situation and a case already known differ just for a rotation of the robot attitude. Since rotations produce shift in the polar map, to design an index system and its associated metric, independent from rotations and scaling we resort to a wavelet transform representation, as in the full resolution 2D situation. Once again, the properties of clustering within the same scale and persistence across different scales of the wavelet transform allows to create extremely compact indexes. In addition, when bi-orthogonal wavelets are

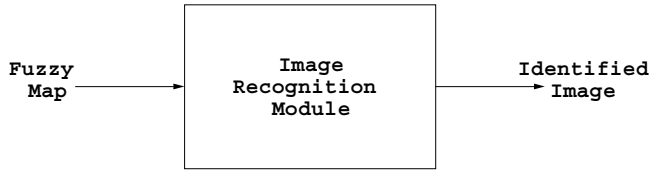

Fig. 5. Image Recognition Module

employed, the metric associated to the L^2-norm can be very efficiently evaluated. In addition, we expect that the differences between new and old cases will affect mainly the details of the fuzzy polar map, related to the finest scales.

Accordingly, the module for image recognition may be given as represented in Figure 5.

As far as the output is concerned, the case has been carefully assessed and deemed it absolutely not reductive to group the infinite situations that the robot could encounter during its navigation essentially into five different "objects":

- Passage
- Corner
- Crossing
- End Passage
- Open space

thus forming a finite set.

The library of the 1-D wavelets transforms of the fuzzy polar map represents the history of the robot's experiences gained until that moment.

4 Conclusions

Traditional methodologies of pattern recognition usually require the availability of templates of the objects we want to classify. This template collection reflects the a priori knowledge we have about the problem to be solved by the image classifier. However, in practical cases, like the cited robot autonomous navigation, or the browsing of large multimedia databases, the prior knowledge could be rather poor, thus leading to a rise of misclassifications. In our contribution, we encapsulated traditional feature extraction and comparison algorithms into a CBR shell that allows constant update of the environment knowledge, or, in other words, of the library of templates as well as of the library of objects to be recognized. We remark that, in principle, there is no limit to the kind and number of features that can be employed for case retrieval. Nevertheless features have to be selected according to their effectiveness in gathering the information the users is interested to. Presently an extensive test of our system is in progress. Further developments of this approach could consist in: the introduction in the features extraction module of data fusion algorithms for the combined use of

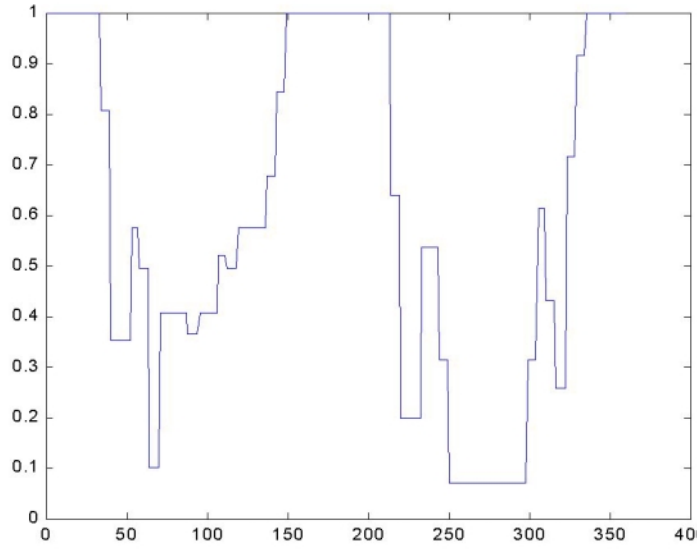

Fig. 6. Polar Representation of the Map of Figure 4

different features, as well as, in the refinement of the case library and object update procedure, to incorporate the evaluation of the effects produced on the environment by the current classification. For instance, in the robot guidance, the update could account for the effectiveness of the move induced by the classification.

References

1. Androutsos, D. A., Plataniotis, K. N. and Venetsanopulos, A. N.: Directional Detail Histogram for Content Based Image Retrieval *DSP97 13th Int. Conf. On Digital Signal Processing*, Proc. July 2-4, 1997, Santorini, pp. 225-228.
2. Bandemer, H. and Nather, W.: *Fuzzy Data Analysis*, Kluwer Academic Publishers, 1992. 448
3. Cox, I. J. and Willfong, G. T.: *Autonomous Robot Vehicles*, Springer-Verlag, 1990. 447
4. Elfes, A.: Occupancy Grids: A Stochastic Spatial Representation for Active Robot Perception. In: Iyengar, S. S., and A. Elfes (eds.) *Autonomous Mobile Robots: Perception, Mapping, and Navigation*, IEEE Computer Society Press, 1991, pp. 60-71. 448
5. Ficet-Cauchard, V., Porquet, C. and Revenu, M.: An Interactive Case-Based Reasoning System for the Development of Image Processing Applications. In: B. Smyth and P. Cunningham (eds.) *Advances in Case-Based Reasoning, Proc of EWCBR-98*, Lecture Notes In Artificial Intelligence, **1488**, Springer, 1998, pp. 437-447. 443

6. Grimnes, M. and Aamodt, A.: A Two Layer Case-Based Reasoning Architecture for Medical Image Understanding. In: I. Smith and B. Faltings (eds.) *Advances in Case-Based Reasoning, Proc of EWCBR-96*, Lecture Notes In Artificial Intelligence, **1168**, Springer, 1996, pp. 164-178. 443
7. Henkind, S. J. and Harrison, M. C.: An Analysis of Four Uncertainty Calculi. IEEE Transactions on Systems, Man, and Cybernetics, **18**(5), 1988, pp.700-714. 448
8. Hsu, Y. N. and Arsenault, H. H.: Optical Pattern Recognition Using Circular Hermonic Expansion. *Applied Optics*, **21**, November 1982, pp. 4016-4019. 445
9. Jacovitti, G. and Neri, A.: Multiscale Image Features Analysis with Circular Harmonic wavelets. In: *Wavelets Applications In Signal and Image Processing III*, Proc. of SPIE 2569, July 1995, pp. 363-372. 445
10. Jacovitti, G., Manca, A. and Neri, A.: Hypercomplete Circular Harmonic Pyramids. In: *Wavelets Applications In Signal and Image Processing IV*, Proc. of SPIE 2825, August 1996, pp. 352-363. 445
11. Jacovitti, G. and Neri, A.: Content based Image Classification with Circular Harmonic Wavelets. In: *Hybrid Image and Signal Processing VI*, Proc. of SPIE 3389, 1998. 445
12. Jacovitti, G. and Neri, A.: Multiscale Circular Harmonic wavelets: a Tool for Optimum Scale-Orientation Independent Pattern Recognition. In: *Wavelets Applications V*, Proc. of SPIE 3391, 1998. 445, 446, 454
13. Klir, G. J., and T. A. Folger.: *Fuzzy Sets, Uncertainty and Information*, Prentice Hall, 1988. 448
14. Leonard, J. J. and Durrant-White, H. F.: *Directed Sonar Sensing for Mobile Robot Navigation*, Kluwer Academic Publishers, 1992. 448
15. Martens, J. B.: Local Orientation Analysis in Images by Means of the Hermite Transform. *IEEE Transactions On Image Proc.*, **6**(8), July 1997, pp. 1103-1116. 445, 446
16. Micarelli, A. and Sciarrone, F.: A Case-Based System for Adaptive Hypermedia Navigation. In: I. Smith and B. Faltings (eds.) *Advances in Case-Based Reasoning, Proc. of EWCBR-96*, Lecture Notes in Artificial Intelligence, **1168**, Springer-Verlag, Berlin, 1996, pp. 266-279.
17. Micarelli, A., Sciarrone, F., Ambrosini, L. and Cirillo, V.: "A Case-Based Approach to User Modeling". In: B. Smyth and P. Cunningham (eds.) *Advances in Case-Based Reasoning, Proc. of EWCBR-98*, Lecture Notes in Artificial Intelligence, **1488**, Springer-Verlag, Berlin, 1998, pp. 310-321.
18. Michel, S., Karoubi, B., Bigun, J. and Corsini, S.: Orientation Radiograms for Indexing and Identification in Image Databases. *Signal Processing VIII*, Proc. of EUSIPCO, Trieste, Italy, 10-13 Sept. 1996, pp. 1693-1696. 445
19. Oriolo, G., Ulivi, G. and Vendittelli, M.: Motion Planning with Uncertainty: Navigation on Fuzzy Maps. *Proc. 4th IFAC Symposium on Robot Control*, (SY.RO.CO.94), **1**, Capri, 1994, pp. 71-78. 447, 448, 449
20. Oriolo, G., Ulivi, G. and Vendittelli, M.: On-Line Map Building and Navigation for Autonomous Mobile Robots. *Proc. 1995 IEEE Int. Conference on Robotics and Automation*, Nagoya, Japan, 1995, pp. 2900-2906. 447, 448, 449
21. Polaroid Corporation *Ultrasonic Ranging System*, 1987. 447
22. Ravichandran, G., and Trivedi, M. M.: Circular-Mellin Features for Texture Segmentation. *IEEE Transactions on Image Processing*, **4**(12), December 1995, pp. 1629-1640. 445

23. Sheng, Y. and Arsenault, H. H.: Object Detection from a Real Scene using the Correlation Peak Coordinates of Multiple Circular Harmonic Filters. *Applied Optics*, **28**(2), January 1989, p. 245. 445
24. Simoncelli, E. P.: A Rotation Invariant Pattern Signature. *Proc. of the IEEE Int. Conf. On Image Proc. ICIP 96*, Lausanne, Switzerland, September 16-19, 1996. 445
25. Zimmermann, H.-J.: *Fuzzy Set Theory and Its Applications*, Kluwer Academic Publishers, 1991. 448

Appendix

Given an image $\mathbf{g}^{(i)}(\mathbf{x})$ and a lattice

$$\mathbf{M} = \{\mathbf{m} = \left[\, m_1\delta \; m_2\delta \,\right]^T, m_1 = 1, \cdots, N_1, m_2 = 1, \cdots, N_2\}$$

defined on the image support, we can locally expand $\mathbf{g}^{(i)}(\mathbf{x})$ around any lattice node, so that

$$\mathbf{g}^{(i)}(\mathbf{x}) = \frac{1}{\sum_{\mathbf{m}\in\mathbf{M}} V\left(\frac{|\mathbf{x}-\mathbf{m}|}{\sigma}\right)} \sum_{m\in M} \sum_{n} \sum_{k} \mathbf{C}^{(i)}_{n,k}(\mathbf{m}) \frac{1}{\sigma} \mathcal{L}^{(n)}_k \left(\frac{|\mathbf{x}-\mathbf{m}|}{\sigma}, \gamma(\mathbf{x}-\mathbf{m}) \right)$$

where

$$\mathcal{L}^{(n)}_k(r,\gamma) = (-1)^k 2^{\frac{(|n|+1)}{2}} \pi^{\frac{|n|}{2}} \left[\frac{\mathrm{k!}}{(|\mathrm{n}|+\mathrm{k})!} \right]^{\frac{1}{2}} r^{|n|} L^{|n|}_k(2\pi r^2) e^{-\pi r^2} e^{jn\gamma}$$

are the Laguerre-Gauss (LG) circular harmonic functions, and $L^{(n)}_k(t)$ are the generalized Laguerre polynomials defined by the Rodrigues formula

$$L^{(n)}_k(t) = \frac{\mathrm{t^{-n} e^t d^k}}{\mathrm{k!}} \frac{\mathrm{d^k}}{\mathrm{d_{t^k}}} \left[t^{k+n} e^{-t} \right] = \sum_{h=0}^{k} (-1)^h \binom{n+k}{k-h} \frac{\mathrm{t^h}}{\mathrm{h!}}$$

The expansion coefficients are given by the inner products between the current pattern and elements of the representation basis as follows:

$$\mathbf{C}^{(i)}_{n,k}(\mathbf{m}) = \langle \mathbf{g}^{(i)}(\mathbf{x}) V\left(\frac{\mathbf{x}-\mathbf{m}}{\sigma}\right), \frac{1}{\sigma} \mathcal{L}^{(n)}_k \left(\frac{|\mathbf{x}-\mathbf{m}|}{\sigma}, \gamma(\mathbf{x}-\mathbf{m}) \right) \rangle \qquad (2)$$

where $V(\mathbf{x})$ is a Gaussian weight function, i.e.: $V(\mathbf{x}) = e^{-\pi|x|^2}$. Let R_φ be the rotation operator. Then, as detailed in [16], the distance between the actual image $\mathbf{f}(\mathbf{x})$ with LGT coefficients $\mathbf{D}_{n,k}(\mathbf{m})$ and the i-th pattern with LGT coefficients $\mathbf{C}^{(i)}_{n,k}(\mathbf{x}_0)$, associated with the L^2-norm is

$$\min_{\varphi_i, \mathbf{b}_i} \{ [\mathbf{f}(\mathbf{x}) - \mathbf{g}^{(i)}(R_{\varphi_i}(\mathbf{x}-\mathbf{b_i}))]^T [\mathbf{f}(\mathbf{x}) - \mathbf{g}^{(i)}(R_{\varphi_i}(\mathbf{x}-\mathbf{b_i}))] \}$$

On the other hand,

$$[\mathbf{f}(\mathbf{x}) - \mathbf{g}^{(i)}(R_{\varphi_i}(\mathbf{x} - \mathbf{b_i}))]^T[\mathbf{f}(\mathbf{x}) - \mathbf{g}^{(i)}(R_{\varphi_i}(\mathbf{x} - \mathbf{b_i}))] = \|\mathbf{f}(\mathbf{x})\|^2 + \|\mathbf{g}^{(i)}(\mathbf{x})\|^2 - 2Re\left[\langle \mathbf{f}(\mathbf{x}), \mathbf{g}^{(i)}(R_{\varphi_i}(\mathbf{x} - \mathbf{b_i}))\rangle\right]$$

due to the properties of the LG functions, since when $\delta = \sigma/\sqrt{\pi}$,

$$\sum_{m \in M} V\left(\frac{|\mathbf{x} - \mathbf{m}|}{\sigma}\right) \cong 1$$

denoting with the over bar the complex conjugate of a complex number, the inner product in the previous formula can be approximated as follows:

$$\langle \mathbf{f}(\mathbf{x}), \mathbf{g}^{(i)}(R_{\varphi_i}(\mathbf{x} - \mathbf{b_i}))\rangle \cong \sum_{\mathbf{m} \in \mathbf{M}} \sum_{k} \overline{\mathbf{C}}^{(i)}_{n,k}(\tilde{\mathbf{x}}_0 + \mathbf{m})\mathbf{D}_{n,k}(\tilde{\mathbf{x}}_0 + \mathbf{b_i} + R_{-\hat{\varphi}_i}\mathbf{m})$$

where the pattern orientation is estimated by means of the recursive formula (see [12]),

$$\hat{\varphi}_i^{(m)} = \hat{\varphi}_i^{(m-1)}(\mathbf{b_i}) - \frac{\sum_{\mathrm{n}=-\infty}^{\infty} \mathrm{nM}_{\mathrm{n}}^{(\mathrm{i})}(\mathbf{b_i}) \sin\left[\mathrm{n}\hat{\varphi}^{(\mathrm{m}-1)}(\mathbf{b_i}) - \eta_{\mathrm{n}}^{(\mathrm{i})}(\mathbf{b_i})\right]}{\sum_{\mathrm{n}=-\infty}^{\infty} \mathrm{n}^2\mathrm{M}_{\mathrm{n}}^{(\mathrm{i})}(\mathbf{b_i}) \cos\left[\mathrm{n}\hat{\varphi}^{(\mathrm{m}-1)}(\mathbf{b_i}) - \eta_{\mathrm{n}}^{(\mathrm{i})}(\mathbf{b_i})\right]}$$

with

$$M_n^{(i)}(\mathbf{b}_i)e^{j\eta_n^{(i)}(\mathbf{b}_i)} \cong \sum_{\mathbf{m} \in \mathbf{M}} \sum_{k} \overline{\mathbf{C}}^{(i)}_{n,k}(\mathbf{m})D_{n,k}\mathbf{b}_i$$