

Binary Shape Normalization Using the Radon Transform

Salvatore Tabbone and Laurent Wendling

LORIA

Campus scientifique

BP 239, 54 506 Vandœuvre-les-Nancy Cedex, France

{`tabbone,wendling`}@loria.fr

Abstract. This paper presents a novel approach to normalize binary shapes which is based on the Radon transform. The key idea of the paper is an original adaptation of the Radon transform. The binary shape is projected in Radon space for different levels of the (3-4) distance transform. This decomposition gives rise to a representation which has a nice behavior with respect to common geometrical transformations. The accuracy and the efficiency of the proposed algorithm in the presence of a variety of transformations is demonstrated within a shape recognition process.

1 Introduction

By definition The Radon transform of an image is determined by a set of projections of the image along lines taken at different angles. For discrete binary image data, each non-zero image point is projected into a Radon matrix. Earlier works [6,10,11] on the 2D Radon transform were dedicated to find high-valued coefficients in the transformed domain, in order to detect specific shape primitives like straight lines or arcs of conics. In all these approaches the information encoded is contour-based allowing the characterization of simple shapes. Furthermore, this kind of representation is not suited to a recognition task because it needs to be normalized with respect to geometric parameters (translation, rotation and scaling). Indeed, it is difficult to recover all the geometric parameters of the transformation between two objects using directly the Radon transform. To overcome this problem we propose an original adaptation of the Radon transform. We define a new representation which has a low time complexity and a nice behavior with respect to common geometrical transformations. The key idea is to project each binary shape in the Radon space for different levels of the (3-4) distance transform. In this manner we take into account the link between the internal structure and the boundaries of the shape. Thus, we provide a global description of any binary shape whatever its type and form are.

The remainder of the paper is organized as follows. A brief review of shape representation methods is presented in Section 2. The definition of the Radon transform is recalled in Section 3. The current method is described in Section 4 and experimental results are given in Section 5. Finally, Section 6 presents conclusions and the future work.

2 Related Works

The normalization of binary shape [4] is a mandatory step in many computer vision system especially when the main goal is the discrimination of objects depending on their forms. Shape representation for object recognition has been the subject of much research, and extensive surveys of shape analysis can be found in [7,12]. The choice of a particular representation scheme is usually driven by the need to cope with requirements such as robustness against noise, stability with respect to small distortions, invariance to common geometrical transformations, or tolerance to occlusions.

Many approaches have been proposed to describe the boundary contours from a small set of features. Fourier descriptors [8,15] have been widely used and modified versions [16] have been proposed to compute the affine transformation between one shape to another. In most cases the centroid of the shape is required to define the geometric transform and it is well known that the position of the centroid is sensitive to noise. Moreover, the Fourier descriptors represent the global appearance of a shape from their most important components. The number of required coefficients relies on the shape of the given object and is usually rather large.

In curvature approaches [13,23] a shape is described in a scale space by the maximum of the curvature. The similarity of two shapes is determined by measuring the distance between their corresponding scale space representations. In some approaches the similarity is computed at a high scale or at all scales. These methods yields interesting results. However the number of scales is set manually because it is difficult to compute it automatically. For the same reason the first and the last scales are set manually too.

To solve the correspondence problem of contour points between two shapes, S. Belongie [1] links a shape context to each contour point. The shape context at a contour point captures the distribution around it and enables to solve the correspondences as an optimal assignment problem. This method has the advantage that it does not require ordered boundary points over the previous one. Nevertheless, matching local contexts does not necessarily preserve the coherence of shapes.

A recognition system which recognizes objects from their silhouettes has been proposed in [24]. Each instance of an object is represented by a graph built on the medial axis of the shape silhouette. The nodes are skeleton junctions and edges are the primitive points between them. An improved version is defined by the shock structure notion [9] which is obtained by viewing the medial axis as being the locus of singularities. This structure is represented in a shock graph which describes the shape more accurately than the medial axis graph.

Several approaches have been proposed to compare shock graphs. K. Siddiqi [21] modifies the shock graph into a shock tree and matching is performed by subgraph isomorphism or by finding the maximal cliques [14]. An alternative approach, also based on the singularities of a curve evolution process, has been given in [20]. The underlying graph is hierarchical, but more complex. Novel recognition frameworks have been proposed for matching shock graphs of 2D

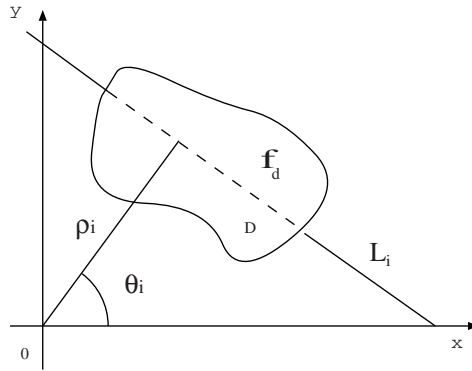


Fig. 1. Definition of the Radon transform.

shapes outlines [18]. In these approaches the edit distance is used to measure the similarity between shapes. It is performed by searching the optimal sequence in the space of all possible transition sequences. These methods are highly effective since they rely on global optimizations. Partial and occluded shapes are well matched. The main disadvantage is that they are computationally expensive, although heuristics can be used to reduce the complexity [19] in practice. A further problem is that they do not take into account the internal structure of general objects and the medial axis is difficult to extract from real images due to its noise-sensitivity.

3 The Radon Transform

Let $f(x, y)$ be an image. Its Radon transform is defined in [5]:

$$T_{Rf}(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos(\theta) + y \sin(\theta) - \rho) dx dy \tag{1}$$

where $\delta(\cdot)$ is the Dirac delta-function ($\delta(x) = 1$ if $x = 0$ and 0 elsewhere), $\theta \in [0, \pi[$ and $\rho \in [-\infty, \infty]$. In other words, T_{Rf} is the integral of f over the line $L_{(\rho, \theta)}$ defined by $\rho = x \cos(\theta) + y \sin(\theta)$.

In the context of shape recognition it is of particular interest to consider the case where the general function f is replaced by (see Figure 1):

$$f_D(x, y) = \begin{cases} 1 & \text{if } (x, y) \in D \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Let L_i be in normal form (ρ_i, θ_i) in the plane (see Figure 1). The Radon transform $T_{Rf}(\rho, \theta)$ describes the length intersection of all the line L_i with the function f_D for all $\theta_i \in [0, \pi[$ and $-\rho_{min} < \rho \leq \rho_{max}$. When an implementation is considered, ρ_{min} and ρ_{max} are finite and rely on the image size.



Fig. 2. A 2D shape and its Radon transform.

Since the Radon transform is linear by definition, geometric properties like straight lines or curves can be made explicit by the Radon transform which concentrates energies (loci of intersection of several sinusoidal curves) from the image in few high-valued coefficients in the transformed domain. These remarks are illustrated in Figure 2 where white pixels are the loci of high energies.

The Radon transform has several useful properties. Some of them are nice for shape representation [5]:

- Periodicity: $T_{Rf}(\rho, \theta) = T_{Rf}(\rho, \theta + 2k\pi)$, k integer. The period is therefore 2π .
- Shift of translation vector $\mathbf{u} = (x_0, y_0)$: $T_{Rf}(\rho - x_0 \cos(\theta) - y_0 \sin(\theta), \theta)$. A translation of f results in the shift of its transform in the variable ρ by a distance equal to the projection of the translation vector on the line $\rho = x \cos(\theta) + y \sin(\theta)$.
- Rotation of θ_0 : $T_{Rf}(\rho, \theta + \theta_0)$. A rotation of the image by an angle θ_0 implies a translation of the Radon transform in the variable θ .
- Scaling of α : $\frac{1}{|\alpha|} T_{Rf}(\alpha \times \rho, \theta)$. A scaling of f results in a scaling of both the ρ coordinate and the amplitude of the transform.

To be useful, a shape recognition framework should allow explicit invariance through the operations of translation, rotation and scaling. To measure the similarity between the Radon matrix of two shapes it is necessary to know the underlying geometric transformations from one shape into the other. However, we can see from the previous properties that if a given shape is translated, rotated and scaled, it will be difficult to recover all the parameters of the geometric transformations from the Radon transform. To overcome this problem we propose in the next section an original adaptation of the Radon transform.

4 \mathcal{R} -Transform

Let the following transform, called \mathcal{R} -transform, be:

$$\mathcal{R}_f(\theta) = \int_{-\infty}^{\infty} T_{Rf}^2(\rho, \theta) d\rho \quad (3)$$

where T_{Rf} is the Radon transform of f . We can show the following properties:

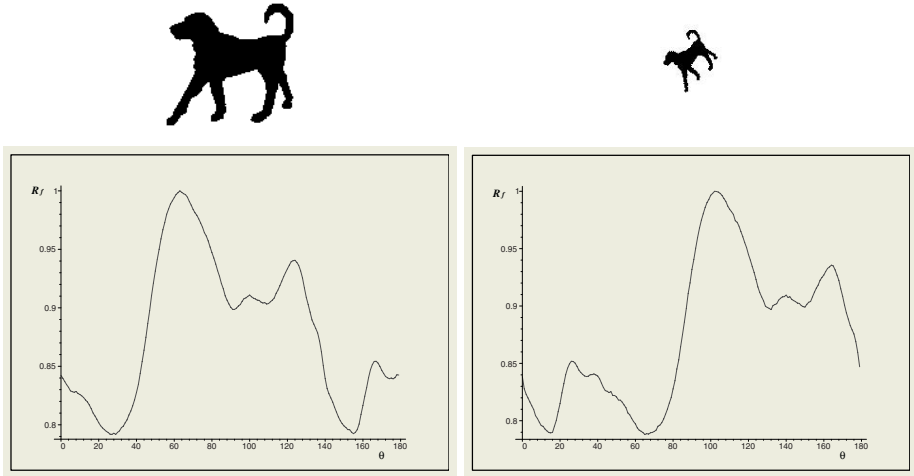


Fig. 3. \mathcal{R} -transforms of the same shape which has been rotated, scaled and translated. Only the rotation provides a shift of the \mathcal{R} -transform.

- **Periodicity:** $\mathcal{R}_f(\theta \pm \pi) = \mathcal{R}_f(\theta)$. The period is therefore set to π .
- **Rotation:** $\mathcal{R}_f(\theta + \theta_o) = \int_{-\infty}^{\infty} T_{Rf}^2(\rho, \theta + \theta_o) d\rho$. A rotation of the image by an angle θ_o implies a translation of the \mathcal{R} -transform of θ_o .
- **Shift:** $\int_{-\infty}^{\infty} T_{Rf}^2(\rho - x_o \cos(\theta) - y_o \sin(\theta)) d\rho = \mathcal{R}_f(\theta)$. The \mathcal{R} -transform is invariant under a translation of f by a vector $\mathbf{u} = (x_o, y_o)$.
- **Scaling:** $\frac{1}{\alpha} \int_{-\infty}^{\infty} T_{Rf}^2(\alpha\rho, \theta) d\rho = \frac{1}{\alpha^2} \mathcal{R}_f(\theta)$ ($\alpha > 0$). A scaling of f causes in a scaling of only the amplitude of the \mathcal{R} -transform.

To summarize, the \mathcal{R} -transform is invariant under translation and scaling if the transform is normalized. A rotation of the shape implies a translation of the transform modulo π . Figure 3 shows two \mathcal{R} -transforms of the same object which has been rotated, scaled and translated. We note that only the rotation provides a modification of the function.

Given a large collection of shapes, unraveling its redundancies with only one \mathcal{R} -transform per shape is not efficient because the \mathcal{R} -transform provides a highly compact shape representation. In this perspective, to improve the description, each shape is projected in the Radon space for different segmentation levels of the (3 – 4) distance transform. A distance transformation is an operation that converts a binary image to an image where each element is valued to the distance of the nearest boundary contour [2,17]. There are different families of distance transformation (see [2,17] for more details). The Chamfer distance, called the (3–4) distance transform, is fast and simple to implement and provides a good approximation of the Euclidean distance. To compute the (3–4) distance transform, the two additive masks of Figure 4 are applied in two passes on the image. In the forward pass the first mask starts in the upper left corner of the

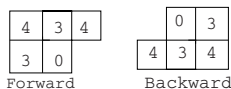


Fig. 4. The masks for computing the (3 – 4) distance transform.

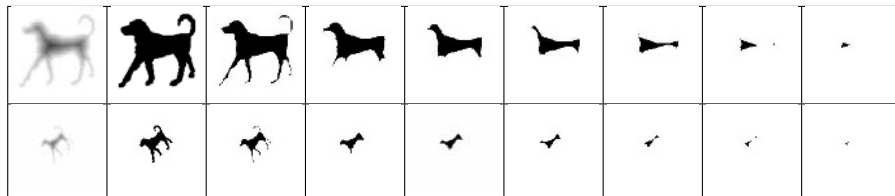


Fig. 5. First column: (3-4) distance transform. Other columns: segmented images for 8 different levels of distance transform.

image moving from left to right and from top to bottom. The opposite operations are performed for the backward mask.

Given the distance transform of a shape, the distance image is segmented in n equidistant levels in order to keep the segmentation isotropic. For each distance level, pixels having a distance value superior to that level are selected and at each level of segmentation, an \mathcal{R} -transform is computed. In this manner, we capture both the internal structure and the boundaries of the shape.

Since the Radon transform is linear, all the \mathcal{R} -transforms are computed in only one step. That is, each non-zero point (x_i, y_i) is projected simultaneously into different Radon matrices. The number of projections relies on the number of segmentation levels and on the value of the distance transform at that point. So, each binary shape is composed of a set of \mathcal{R} -transforms describing a 2D surface and verifying the previous geometric properties. Figure 5 shows the distance transform of the two dog images of Figure 3 and their corresponding distance images for 8 segmentation levels. We note that the isotropy of the segmentation is kept. The corresponding 2D surfaces are presented in Figure 6. We can see that one surface is very close to a circular permutation of the other. In the next section, we give some indications on the way of computing the discrete \mathcal{R} -transform.

5 Digital Considerations

Since the aim of this paper is not to provide a new version of the discrete Radon transform but to emphasize its application, we only give some tips about its implementation. A great number of algorithms have been proposed which are more or less fast. We have adapted the approach proposed in [3] which is well-suited to deal with binary images even though the complexity is high. Let a single point image with coordinates (x_0, y_0) be:

$$I(x, y) = \delta(x - x_0)\delta(y - y_0). \tag{4}$$

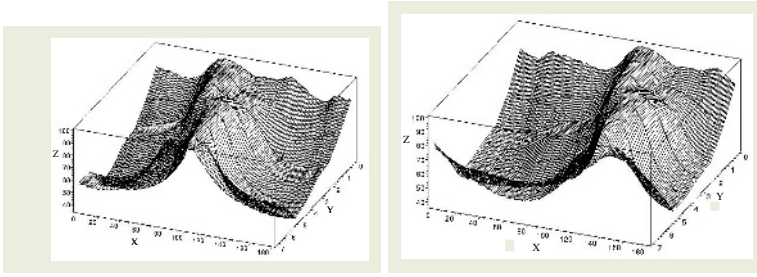


Fig. 6. Surface visualization of the corresponding 2D \mathcal{R} -transforms of the Figure 5. On the X-axis is the number of orientations in the Radon transform. On the Y-axis is the number of level cuts in the distance transform.

Its Radon transform is by definition:

$$T_{R^I}(\rho, \theta) = \delta(\rho - x_0 \cos(\theta) - y_0 \sin(\theta)). \tag{5}$$

That is, a single point has a Radon transform which is non-zero along a sinusoidal curve of equation $\rho = x_0 \cos(\theta) + y_0 \sin(\theta)$. Therefore, since the Radon transform is by definition linear, a way to compute the transform of a binary image is to map every non-zero image point, using the normal parameterization $\rho_i = x_i \cos(\theta_i) + y_i \sin(\theta_i)$, into a Radon matrix. That is, for each point (x_i, y_i) of the image, i is fixed and the value ρ_i is calculated using stepwise increments of θ_i from 0 to π . The increment is defined to avoid aliasing following the Shannon theory (see [22] for more details). Here we set $\Delta\theta = \Delta\rho = 1$, $\Delta x = \Delta y = \frac{1}{2}$ and the sampled values of ρ_i are defined by a linear interpolation. This algorithm requires time $O(N^2M)$ for an image of size $N \times N$ and M different angles (here $M = 180$).

Two major optimizations are made to reduce the complexity. The cosine and the sine of all the possible values of θ_i are computed one time only. The values of ρ_i are also defined recursively. That is, since the step increment Δx is set to $\frac{1}{2}$, we have: $\rho_{i+\frac{1}{2}} = \rho_i + \frac{\cos(\theta_i)}{2}$. Therefore, when we move in the x -direction, ρ is incremented by $\frac{\cos(\theta)}{2}$. Similarly, in the y -direction ρ is increased by $\frac{\sin(\theta)}{2}$.

Hence, the discrete Radon transform is represented by a digital image and the discrete \mathcal{R} -transform is defined by:

$$\hat{\mathcal{R}} = \sum_{n=0}^{\pi} \hat{T}_{R^f}^2(\hat{\rho}, \hat{\theta}). \tag{6}$$

We are aware that the properties of the continuous Radon transform carry over to the discrete Radon transform only approximately, due to errors of discretization. However, we will see in the next section that the provided results have a tendency to show that the errors are small.

6 Experimental Results

To show the efficiency of the proposed algorithm we provide experimental results within a shape recognition process. The method was tested on a database of D. Sharvit [20] who made it kindly available to us on his Web site. This database consists of nine categories with 11 shapes in each category. Figure 7 shows an example of matching one shape of each category against all other shapes of the database. Since there were ten possible similar shapes excluding the shape itself, we provide as results the ten nearest neighbors.

The similarity of two surfaces is defined by the χ^2 distance [1]. Such a distance can be efficiently computed and it is well-adapted to our context:

$$\chi^2 = C(h_q, h_m) = \sum_{k=1}^{\#levels} \sum_{\theta=0}^{\pi} \frac{(h_q(k, \theta) - h_m(k, \theta))^2}{h_q(k, \theta) + h_m(k, \theta)}, \quad (7)$$

where h_q and h_m are respectively the 2D surfaces of query and model shapes belonging to the database. Given the previous cost, we look for the best permutation Π which minimizes: $P(\Pi) = \min\{C(h_{\Pi(q)}, h_m)\}$. Due to the (3-4) transform level cuts, a 2D surface may be composed of non uniformly sampled data. In this perspective we interpolate the values at uniformly spaced points before computing the χ^2 distance.

We can observe from Figure 7 that most of the shapes are well classified. For example, the wrench image provides a good demonstration of the behavior of the proposed approach with respect to geometric transformations and small deformations. Furthermore, images with different sizes and types of occlusion have been added in the database. For example, five occluded images of the fish class have been inserted in the database. We can remark from Figure 8 that our approach have retrieved all the occluded shapes concerning the fish query.

Each image of the database is represented by about 128×128 pixels. All the 2D \mathcal{R} -transforms are computed off-line and the running time is about 0.5s per shape. These results are obtained on a Pentium III, 866 MHz running under Linux.

7 Conclusion

We have presented a new approach to shape recognition which is simple and easy to apply. The key characteristic of our approach is the definition of a new transform, based on both the Radon and the (3-4) distance transform, allowing to capture the shape at different levels of distance. In our experiments we have shown the invariance of our approach under several common image transformations, including the ability to handle partial and occluded shapes. Currently the number of segmentation levels is defined manually. It is obvious that a small-sized shape does not require the same levels of description than a larger one. In this case we need also more precision on the distance decomposition. To achieve more accuracy, further works will be devoted to get a better approximation of the Euclidean distance and to define automatically the number of level cuts.

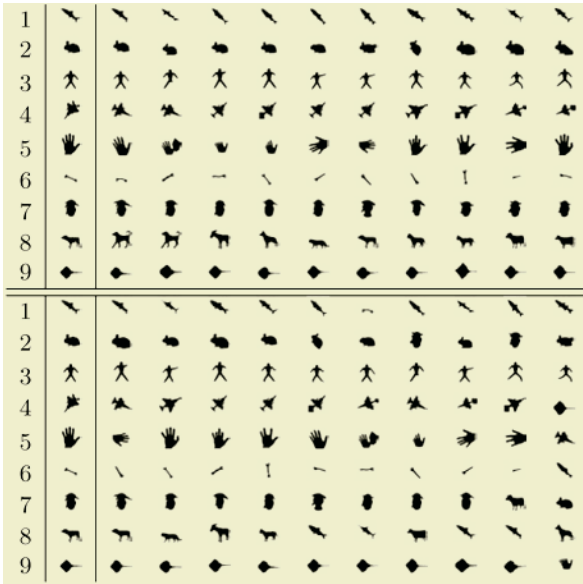


Fig. 7. Top: A database of 99 shapes. Each shape of the first column is matched against every other shapes in the database. Bottom: the 10 nearest neighbors. The self-matching that is a perfect match is excluded from the results.



Fig. 8. Robustness to occlusion. The occluded images are represented in gray. Each shape of the first column is matched against every other shapes in the database (self-matching is excluded from the results). Right: the 14 nearest neighbors.

References

1. S. Belongie, J. Malik and J. Puzicha. Shape Matching and Object Recognition Using Shape Contexts. *IEEE Transactions on PAMI*, 24(4):509–522, 2002.
2. G. Borgefors. Distance Transformations in Arbitrary Dimensions. *CVGIP*, 27:321–345, 1984.
3. R. N. Bracewell. Two-Dimensional Imaging. *Englewood Cliffs, NJ: Prentice Hall*, 1995. pp. 505–537.
4. J. Cortadellas, J. Amat and M. Frigola. Robust Normalization of Shapes. *DGCI 2002*, Bordeaux, France, 2002.
5. S. R. Deans. Applications of the Radon Transform. *New York: Wiley Interscience Publications*, 1983.
6. P. Fränti, A. Mednongov, V. Kyrki and H. Kälviäinen. Content-based Matching of Line-drawing Images Using the Hough Transform. *International Journal on Document Analysis and Recognition*, 3(2):117–124, 2000.

7. A. K. Jain, R. P. W. Duin and J. Mao. Statistical Pattern Recognition: A Review. *IEEE Transactions on PAMI*, 22(1):4–37, 2000.
8. H. Kauppinen, T. Seppänen and M. Pietikäinen. An Experimental Comparison of Autoregressive and Fourier-Based Descriptors in 2D Shape Classification. *IEEE Transactions on PAMI*, 17(2):201–207, 1995.
9. B. B. Kimia, A. R. Tannenbaum and S. W. Zucker. Shapes, Shocks, and Deformations I: The Components of Two-Dimensional Shape and the Reaction-Diffusion Space. *IJCV*, 15:189–224, 1995.
10. V. F. Leavers. Use of the Radon Transform as a Method of Extracting Information about Shape in two Dimensions. *Image Vision and Computing*, 10(2):99–107, 1992.
11. V. F. Leavers. Use of the Two-Dimensional Radon Transform to Generate a Taxonomy of Shape for the Characterization of Abrasive Powder Particles. *IEEE Transactions on PAMI*, 22(12):1411–1423, 2000.
12. S. Loncaric. A Survey of Shape Analysis Techniques. *Pattern Recognition*, 31(8):983–1001, 1998.
13. F. Mokhtarian and S. Abbasi. Shape Similarity Retrieval under Affine Transforms. *Pattern Recognition*, 10(2):31–41, 2002.
14. M. Pelillo, K. Siddiqi and S. Zucker. Matching Hierarchical Structures Using Association Graphs. *IEEE Transactions on PAMI*, 21(11):1105–1119, 1999.
15. E. Persoon and K. Fu. Shape Discrimination using Fourier Descriptors. *IEEE Transactions on SMC*, 7(3):170–179, 1977.
16. Y. Rui, A. She and T. S. Huang. A Modified Fourier Descriptor for Shape Matching in MARS. *Image Databases and Multimedia Search*, 8:165–180, 1998.
17. G. Sanniti di Baja and E. Thiel. Skeletonization algorithm running on path-based distance maps. *Image and Vision Computing*, 14:47–57, 1996.
18. T. B. Sebastian, P. N. Klein and B. Kimia. Recognition of Shapes by Editing Shock Graphs. *ICCV 2001*, 755–762, 2001.
19. T. B. Sebastian, P. N. Klein and B. Kimia. Shock-based Indexing into Large Shape Databases. *ECCV 2002*, 731–746, Denmark, 2002.
20. D. Sharvit, J. Chan, H. Tek and B. Kimia. Symmetry-based Indexing of Image Databases. *Journal of Visual Communication and Image Representation*, 1998.
21. K. Siddiqi, A. Shokoufandeh, S. J. Dickinson and S. W. Zucker. Shock Graphs and Shape Matching. *IJCV*, 35(1):13–39, 1999.
22. P. Toft. The Radon Transform - Theory and Implementation Ph.D. thesis. Department of Mathematical Modelling. Technical University of Denmark, June 1996.
23. C. Urdiales, A. Bandera and F. Sandoval. Non-parametric Planar Shape Representation Based on Adaptive Curvature Functions. *Pattern Recognition*, 35:43–53, 2002.
24. S. C. Zhu and A. L. Yuille. FORM: A Flexible Object Recognition and Modelling System. *IJCV*, 20(3):187–212, 1996.