6. References

[1] C Arcelli, L P Cordella and S Levialdi, 1981. From local maxima to connected skeletons. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 3, pp 134–143.

[2] C Arcelli and G Sanniti di Baja, 1988. Finding local maxima in a pseudo-euclidean distance transform. *Computer Vision, Graphics and Image Processing*, 43, pp 361–367.

[3] C Arcelli and G Sanniti di Baja, 1992. Ridge points in Euclidean distance maps. *Pattern Recognition Letters*, 13, pp 237–243.

[4] E B Goldstein, 1989. *Sensation and perception*, 3rd edition, Wadsworth Publishing Company.

[5] W J Christmas, J Kittler and M Petrou, 1994. Non-iterative contextual correspondance matching. *European Conference on Computer Vision, ECCV'94*.

[6] C R Dyer and A Rosenfeld, 1979. Thinning algorithms for grey scale pictures. PAMI, Vol 1, pp 88–89.

[7] L D Griffin, A C F Colchester and G P Robinson, 1992. Scale and segmentation of grey level images using maximum gradient paths. Image and Vision Computing, Vol 10, pp 389–402.

[8] P L Palmer, J Kittler and M Petrou, 1993. Using focus of attention with the Hough transform for accurate line parameter estimation. *Pattern Recognition*, to appear.

[9] P L Palmer, M Petrou and J Kittler, 1993. Accurate line parameters from an optimising Hough transform for vanishing point detection. *4th International Conference on Computer Vision*, pp 529–539.

[10] K Qian and P Bhattacharya, 1992. A template polynomial approach for image processing and visual recognition. *Pattern Recognition*, Vol 25, pp 1505 1515.

[11] Sha'ashua and Ullman, 1988. *Proceedings of the 2nd International Conference on Computer Vision*, pp 321–327.

[12] S-S Yu and W-H Tsai, 1990. A new thining algorithm for grey scale images by the relaxation technique. *Pattern Recognition*, Vol 23, pp 1067–1076.

A HIERARCHY OF REPRESENTATIONS FOR CURVES

PAUL L. ROSIN

Institute for Remote Sensing Applications, Joint Research Centre

I-21020 Ispra (VA), Italy

paul.rosin@jrc.it

ABSTRACT

A variety of representations are available for describing the shapes of curves. This paper suggests combining a number of alternative curve representations (e.g. lines, superellipses, codons). These are arranged to form a hierarchy of increasing specificity, ranging from qualitative through to quantitative representations. Different levels provide different trade-offs between the various desirable but mutually incompatible properties of shape representations. Depending on the application, the model and image curves may be represented by the complete hierarchies or by the subset of the representations for each sub-sections that is most appropriate in terms of robustness, flexibility, etc. Each representation needs to be applied at the scales which isolate important structures, and so the curve's "natural" scales are selected by an automated process. This multi-scale analysis produces a hierarchy of parts, and thus a hierarchy of hierarchies is formed.

1. Introduction

Many different representations have been suggested or developed for describing the shapes of curves for use in computer vision systems. Examples include straight lines, circular and elliptical arcs, splines, Fourier descriptors, codons, the curvature primal sketch, and scale-space plots of zero-crossings. This paper suggests combining a number of such representations to form a hierarchy of representations.

Hierarchical representations have played an extensive role in Artificial Intelligence. Their advantage is that they provide a more powerful (in terms of flexibility, expressibility, etc.) representation than single flat representations. One of the most common applications of hierarchical representations is for decomposing a model into parts. However, decomposing a model is carried out top-down, with a priori knowledge about the identification and function of its various parts. Since this information is not directly available when analysing an image curve bottom-up it can only be inferred. This is usually performed by analysing the curve at multiple scales. The order of emergence of parts over scale determines their levels in the tree. Of course, this spatial organisation will not always correspond to a functional part decomposition.

Another popular model hierarchy is the specialisation tree. This enables a model to be identified by increasingly precise classifications. These classifications usually require top-down information about the model. It would be difficult to translate this hierarchy into a bottom-up version for describing image curves before they have been assembled into an object and identified from the model database. However, the hierarchy of curve representations proposed here provides a bottom-up equivalent. Whereas the specialisation hierarchy describes the classification of an object in increasing detail, the hierarchy of representations describes the shape of the curve in increasing detail. It should be noted that the increasing classification and shape detail referred to here differ from the increasing spatial detail provided by the decomposition hierarchy.

Despite the popularity of the decomposition (multi-scale) hierarchy for describing curves little attention has been paid to hierarchies of shape representations. Most computer vision systems use a single feature or class of features to represent curves. This paper describes how multiple shape representations can be combined into hierarchies. Since each representation will be applied over multiple scales to form a hierarchy this forms in effect a hierarchy of hierarchies. The benefits that follow from such a scheme are discussed, and its application to image curves is demonstrated.

2. Multiple Shape Representations

A number of important criteria exist for assessing curve representations [11]. These include: expressiveness; completeness; robustness under occlusion and noise; invariance under Euclidean transformations; invariance within a generic viewpoint; the number of parameters and the sensitivity to changes of their values; efficiency of computation; and ease of implementation. These criteria allow curve representations to be compared against each other. None will completely and perfectly satisfy all the criteria. In part, this is because some of the criteria are conflicting, and the various representations are compromises biased towards one set of criteria or another. For instance, it is difficult to be both expressive to subtle variations of shape but yet insensitive to noise. The first criterion is important so that specific shapes can be recognised. The second criterion is essential so that shapes of curves extracted from the real world can be matched reliably. Therefore, rather than search for some ideal, perfect shape representation it is more realistic to use a set of sub optimal representations. These can be combined so that the strengths of one can overcome the weaknesses of another.

Much of the variation between curve representations arises from their amount of reliance on metric descriptions. Some representations are extremely precise while others are relatively vague. Based on this observation we will order representations in terms of increasing specialisation or precision. This can also be considered as a progression from qualitative descriptions through to quantitative descriptions. In other contexts a similar range of shape representations have been recently proposed which are ordered in terms of their invariance to transformations [15] and their robustness to noise and poor resolution [3].

Three benefits of combining the various representations are:

- Robustness - The most appropriate curve representation obviously depends on the shape of the model. For example, for a restricted set of models such as geons, a curve representation consisting of just concave/convex/straight arcs would be entirely adequate [4]. However, the representation's appropriateness is also affected by the context of the scene in which the object is viewed. For instance, a circular boundary of an object may be detected by representing the image curve by ellipses. However, if there is inadequate data either due to noise, occlusion, or because the component appears too small, then ellipse fitting is likely to be unreliable and therefore unsuitable

- Flexibility - Combining different curve representations enables a model to be described by a mixture of feature types. This is particularly interesting if different parts of the model are described by representations at different levels of the qualitative/quantitative spectrum. Certain portions may then be described more rigidly than others. For instance, a section might be constrained to be precisely specified by straight lines and elliptical arcs. Conversely, another section could be allowed some variation of its shape parameters (e.g. size, skewness, compactness) by representing it by codons. This approach has some relation to work done to model industrial parts, allowing tolerances for various types of deviations (e.g. variations in translation, orientation, and area between lines and circles). But rather than a precise metric technique, our method is more qualitative.

- Comparability - An additional benefit of the multi-representation scheme is that it enables comparisons to be easily made between different representations. This will be useful if just a simplified description using a single representation is eventually preferred. Although the design criteria detailed above are one way to compare representations, viewing the results of the fitted representations provides a more experimental approach.

3. Combining Representations

In order to easily combine different curve representations the various feature primitives should share endpoints. Often the determination of breakpoints when segmenting a curve depends on the type of feature being fitted (e.g. straight lines versus circular arcs [16]) or the particular segmentation algorithm (e.g. recursive subdivision [16] versus seed growing [13]). One possibility would be to use one representation to determine the potential breakpoints for other representations. For instance, this is the case when circular arcs are fitted to curves which have already been segmented into straight lines, and subdivision is only considered at line endpoints [16].

However, rather than preselecting some privileged representation, a more general approach will be taken. Segmentation will be based on the singular points of curvature. These have the advantage that they are intrinsic, local properties, and so are invariant under occlusion and Euclidean transformations of the curve within the plane. Under 3D transformations only the zeros of curvature (and not curvature extrema) are invariant [14]. In addition, curvature has been shown to be a perceptually significant feature for human perception of shape [2]. Because of these properties, and because they contain much information about the curve, the singular points of curvature have been a popular tool for analysing curves [5]. Therefore, an important benefit of using singular points is that it enables some of these representations (e.g. codons) to be incorporated.

One of the problems of analysing the curve based on its singularities of curvature is that curvature measurements are extremely prone to noise. However, this can be overcome if the curve is smoothed as part of the multi-scale processing. Since most curve representations operate at a single scale they can then be applied to the curve at each scale.

4. Multiple Model Representations

In addition to representing the image curves in different ways the hierarchical representation approach can be applied to the model boundaries. This would enable a hierarchical matching scheme to be applied to the representation type as well as hierarchical matching of parts (i.e. the multi-scale analysis). Matching between image and model curves would be performed first between the most qualitative representation. The initial potential matches would then be refined and verified by progressing down the tree to more quantitative representations until reliable image feature extraction or adequate image to model feature matching can no longer be performed.

Alternatively, each model part could be described by only a single or a few of the representations from the hierarchy. Several approaches are possible to decide which representation is the most appropriate. First, as mentioned above, is robustness. This suggests that certain representations should be eliminated as unsuitable. For instance, small, short, curved sections of a model will generally not be adequately imaged to provide good ellipse fits. Knowledge of the expected nature of the scene can also be used to influence this decision. Noisy, cluttered outdoor scenes will require more robust representations than well controlled indoor ones. For instance, if occlusion is likely then global features such as Fourier descriptors and moments will perform poorly.

Second, the range of shapes within the object models may eliminate certain representations. This can arise either because the representations are not complete and cannot represent some of the shapes, or because the representation would be unnecessarily cumbersome. For instance, the set of codons defined by Hoffman and Richards [5] are restricted to continuously varying, smooth, closed curves. Curves with straight sections, cusps, or open ends cannot be represented by such codons, and so an alternative representation is required. Another example of an unsuitable representation is the use of straight lines for a model consisting of curved surfaces. Individual curves might have to be represented by many lines, complicating subsequent matching.

The range and similarity of the models to be recognised is also important. If all the models are relatively distinctive then an imprecise, qualitative representation may be sufficient. Otherwise more quantitative representations are necessary to discriminate between more subtle differences in shape. However, only the distinguishing features may need to be represented in this manner. Also discussed above was the issue of flexibility. If some parts of the model are allowed greater variation than others, they can be defined by a qualitative representation rather than a more precise, quantitative representation.

Finally, the requirements of the task of the vision system will indicate the appropriate representations. Certain features can provide more suitable information than others. For instance, concave and convex sections of curves determine the type of adjoining surface curvature (hyperbolic or elliptical), and can be used for reconstructing the 3D shape of the object. Another example is the ellipse which, if it is the projection of a circular feature in the object, determines the pose of the object (with two fold ambiguity).

5. Additional Hierarchies and Axes

There are several other hierarchies and ranges or axes related to the representations. Associated with each curve representation will be a suitable method for assessing correspondences between model and image features. Just as the representations range from qualitative to quantitative, the methods for matching will range from high level symbolic ones to more numeric ones. The more qualitative representations will only be described by labels (e.g. concave/convex/straight) and so matching could simply be a test for identical labels or application of a shape deformation grammar [11]. More quantitative representations provide more information, and comparisons will have to be made between their parameters, resulting in a similarity measure (or equivalently a distance in feature space).

We stated that the curve would be analysed at multiple scales, and the various representations independently applied to the curve at each of these scales. However, there are a range of approaches between standard single and multi-scale techniques, such as determining the natural scales [10]. These methods provide a trade-off between compactness and expressiveness of the representation.

Another issue is the manner in which the multiple scales are generated. The most common technique is to smooth the curve, although alternatives include grouping and pyramids. Smoothing can be performed so that the topology of the curve is preserved (e.g. by smoothing along the curve) or is not preserved (e.g. by smoothing the region the curve encloses). Again there is a range of intermediate approaches [6].

6. An Example Hierarchy

Figure 1 shows an example of a hierarchy of representations. The two most qualitative representations are formed by segmenting at either the zero crossings of curvature to form

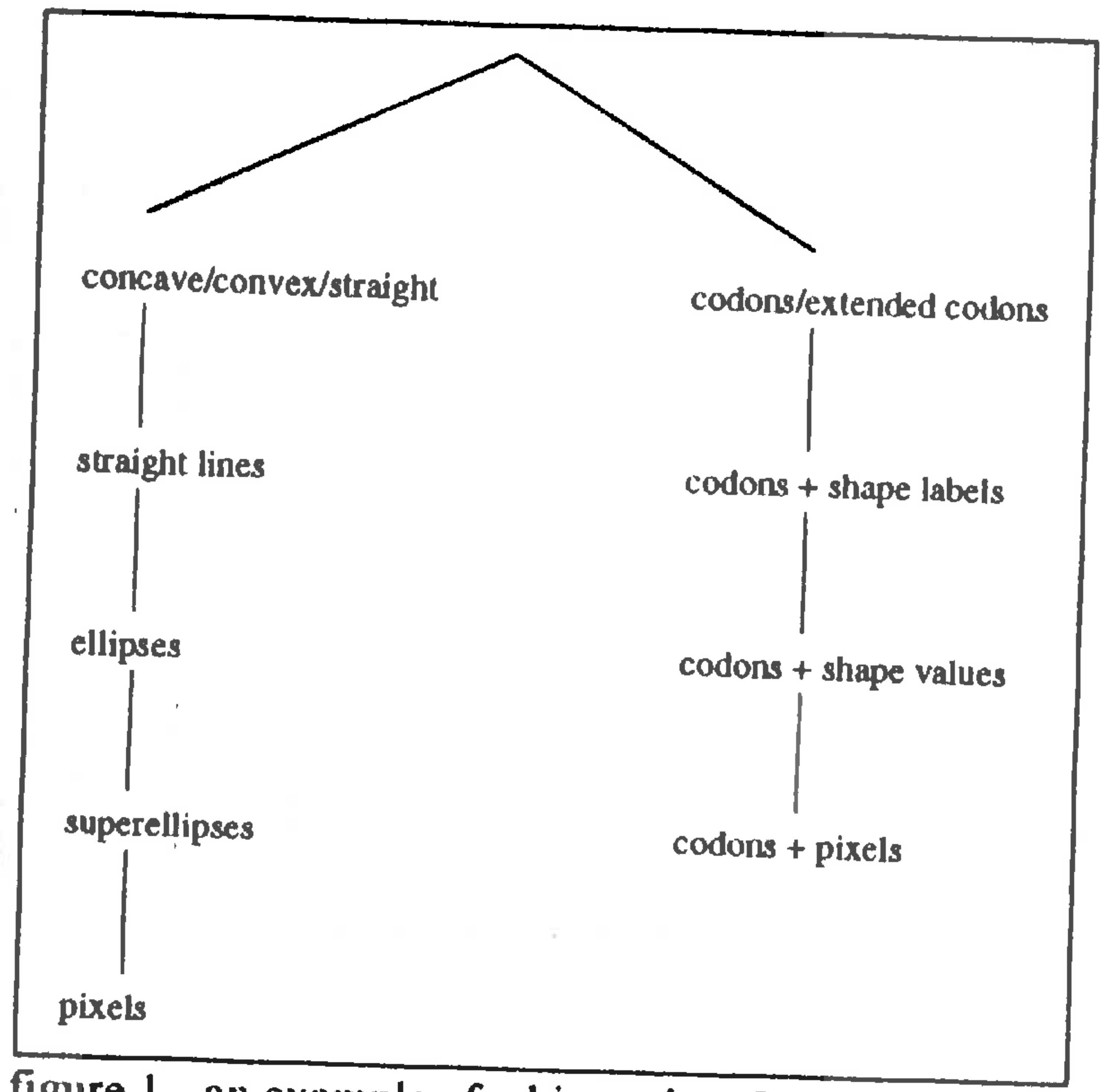


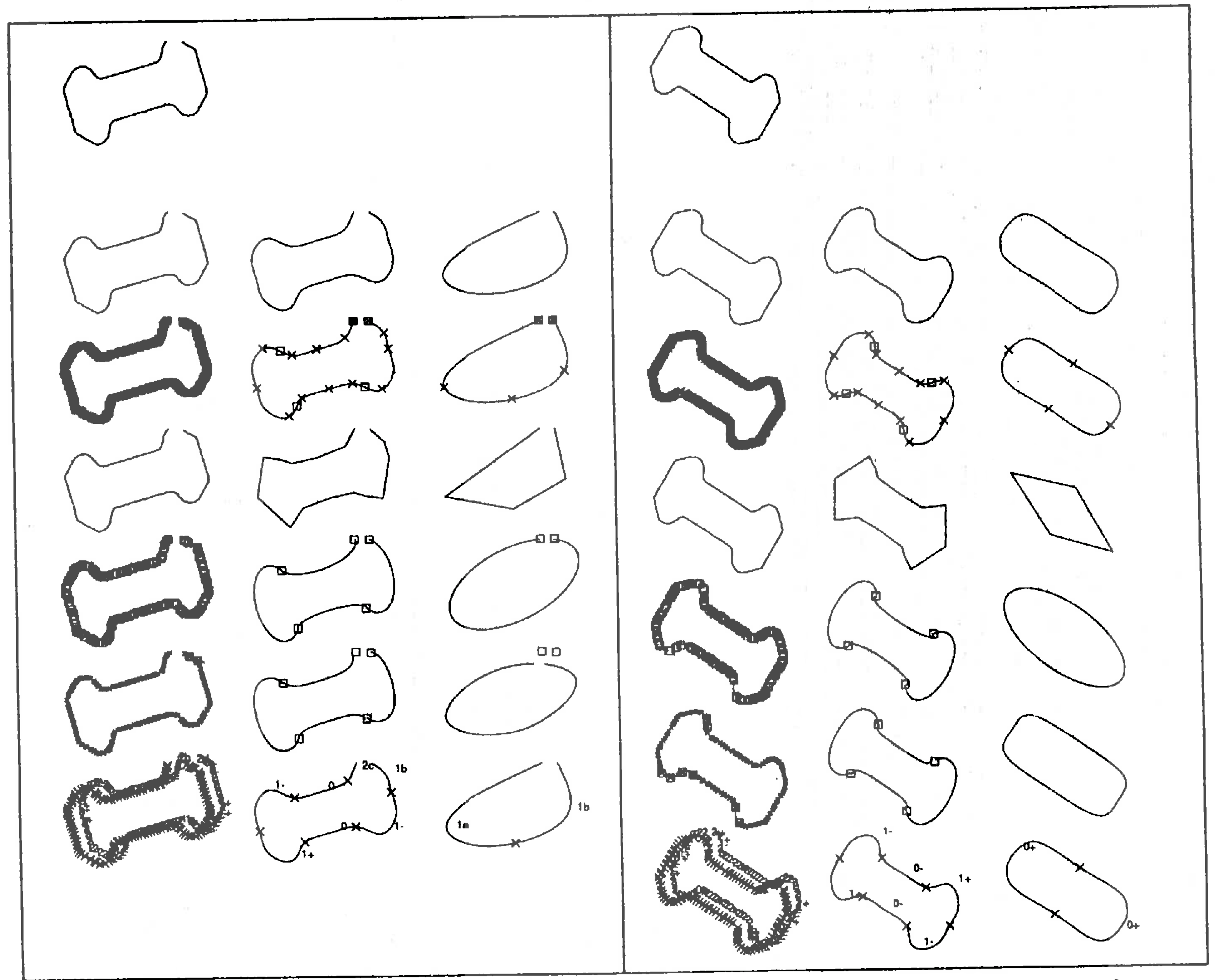
figure 1 - an example of a hierarchy of representations

concave/convex/straight sections, or at the minima of curvature to form codons. Either the restricted set [5] or the extended set [11] of codons can be used. The latter is more complete, and can cope with straight lines, open curves, and cusps, but has the disadvantage that many more codon labels are required. Therefore, whichever is most appropriate to the model data is selected. The concave/convex/straight sections can be described with increasing fidelity by straight lines, ellipses, superellipses, or ultimately, the raw pixel data. Codons can also be described with increasing amounts of detail by incorporating shape labels (e.g. left-skew, very-compact), which can then be quantified (e.g. compactness = 0.68) and augmented by additional metric techniques such as Fourier descriptors and moments. Methods for calculating the codon shape measures of skew, roundness, compactness, length, and orientation are given in [11].

Matching the concave/convex/straight sections or the basic codons would only require checking that they have identical type and shape labels. The codon shape values, and also the lines and ellipses would be compared by calculating the difference in feature space (e.g. the similarity of two straight lines would be some function of their difference in length, orientation, and position). And finally, pixel lists could be compared by correlation techniques.

In this example the curves are only represented at their natural scales [10]. Gaussian smoothing along the curve is performed using Lowe's correction technique [7] to prevent the curve shrinking towards the centre of curvature.

The concave/convex/straight sections are determined as the portions of the curve bounded by the zero crossings of curvature. The codons are terminated by curvature minima, and their labels are identified by examining triples of signed curvature extrema and/or the curvature signs at the ends of the curve [11]. Straight lines are fitted by connecting their



figures 2 & 3.- two views of the gizmo at their natural scales and their segmentation into various features

bounding extremal points.

Ellipses are found by performing a least squares fit of the algebraic distance [1] constrained using Lagrange multipliers to go through the bounding extremal points. The above procedure provides a closed form solution, but has the disadvantage that non-elliptical conics (i.e. hyperbolae and parabolae) may also be fitted. In such situations the conic is discarded and a circular arc is fitted instead.

An iterative technique is necessary to fit the superellipses [12]. This involves minimising the sum of the distances between each data point and the superellipse along the ray from the point to the centre of the superellipse. The procedure is initialised using an ellipse if this can be obtained by fitting a conic, or a circle otherwise. In the latter case the initial orientation is determined using central moments.

In the following figures part (a) shows the original curve, (b) the set of automatically determined natural scales at which to describe the curve, (c) the curve at its natural scales marked with the zero crossings (plotted as boxes) and extrema (crosses) of curvature, (d) straight lines drawn through the singularities of curvature, (e) ellipses fitted to the concave/convex sections, (f) superellipses fitted to the concave/convex sections, and (g) the sections between minima labelled as the extended set of codons.

Figures 2 and 3 show two views of the gizmo - a simple manufactured object. The natural scales capture: (i) the fine noise and quantisation effects, (ii) the basic shape without the noise, and (iii) an overall blob. The open curve is slightly distorted as the highest scale. This is due to difficulties in estimating extensions of the curve which are necessary for the convolution process of Gaussian smoothing. At the finest scales the noise creates numerous spurious singularities of curvature, causing the curve to be segmented into many tiny sections which are of little use when performing object recognition. The coarser scales provide more useful descriptions. Since the descriptions are local they are insensitive to the transformation and occlusion of the gizmo. Comparing the representations at the middle scale the lines are less appropriate since the bulges of the gizmo are fragmented. In contrast, each bulge can be represented by a single convex section. In this example, the superelliptic arcs are only slightly more accurate than the ellipses. The codon representation is similar to the arcs although each bulge is broken into its two halves. This is because even though the ends may appear straight there is a slight curvature. Reliably determining straight sections based on curvature is problematic [11]. At the coarsest scale the two sets of descriptions obtained in figures 2 and 3 are very similar with the exception of the codons. This arises because specific labels are generated for curve sections adjoining open boundaries. While this also occurs at the previous scale the curve sections are relatively short with respect to the complete curve. At the coarsest scale *all* the curve sections terminate at an open end of the curve. The improvement of the superelliptical arcs over the elliptical ones is more evident at the coarsest scale. If each section of curve were to be described by a single representation then at the medium scale the ellipses or codons are most suitable. The superellipses do not provide enough improvement in shape fidelity over the ellipses to be worth the additional processing and potential instabilities of the fitting process. However, at the coarsest scale both codons and lines are unsuitable since the former are sensitive to occlusion and the latter do not describe the shape accurately. The ellipses or superellipses are more appropriate.

7. Conclusions

We have described how various curve representations can be combined to form a hierarchy of representations. This allows the strengths of one representation to be played off against the weaknesses of another. In addition it facilitates the comparison of different representations which is especially useful for complex shapes where the most appropriate representation is not otherwise immediately obvious. The multi-scale analysis (coarse $\rightarrow$ fine spatial scale) and multiple representations (qualitative $\rightarrow$ quantitative shape descriptors) can be thought of as two axes of a two dimensional feature space. The complete feature space is useful for describing shapes rather than just representations at points or lines within the space as is the more common practice.

A feature of the techniques involved in the processing (including the smoothing, scale selection, segmentation, and fitting) is that they are all non-parametric. This is essential if a general purpose, robust system is required that can cope with a variety of data without requiring user intervention for the tweaking of parameters.

The example hierarchy presented here can be extended in many ways. Additional representations could be included, and more degrees of freedom given to the existing representations (e.g. tapering and bending of superellipses). Another approach would be to add representational layers to support higher level groupings of consecutive features. For instance, sequences of straight lines can be combined into corners and arcs [8] and codons can be combined to form various bumps [11]. The current selection mechanism for breakpoints could also be extended to provide greater sensitivity to subtle variations in shape. One approach would be to use singularities of higher order derivatives of curvature.

Finally, it would be of interest to investigate the performance of human visual perception in more detail. Earlier an analogy was made between the top-down specialisation hierarchy and the shape representation hierarchy. Rosch [9] has shown that humans first recognise objects at a specific "basic level" in the specialisation hierarchy, and then continue to categorise them in more detail. Likewise, there may be some support for the hypothesis that object shapes are best initially described at some particular level in the representation hierarchy. Currently we have only informally discussed the suitability of particular curve representations for each model curve section in terms of robustness, discriminability, flexibility, etc. Future work will concentrate on methods for automatically selecting appropriate representations by assessing their suitability.

8. References

[1] A. Albano, "Representation of digitized contours in terms of conic arcs and straight-line segments", *CGIP*, Vol 3, pp. 23-33, 1974.

[2] F. Attneave, "Some informational aspects of visual perception", *Psychological Review*, Vol. 61, pp. 183-193, 1954.

[3] A.F. Bobick and R.C. Bolles, "The representation space paradigm of concurrent evolving object descriptions", *Trans. IEEE PAMI*, Vol. 14, pp. 146-156, 1992.

[4] S.J. Dickinson, A.P. Pentland, and A. Rosenfeld, "3-D shape recovery using distributed aspect matching", *Trans. IEEE PAMI*, Vol. 14, pp. 174-198, 1992.

[5] D.D. Hoffman and W.A. Richards, "Representing smooth plane curves for recognition", in *Proc. AAAI*, pp. 5-8, 1982.

[6] B.B. Kimia, A. Tannenbaum and S.W. Zucker, "Entropy scale-space", *Visual Form*,

pp. 333-344, 1992.

[7] D.G. Lowe, "Organization of smooth image curves at multiple scales" in *Proc. 2nd ICCV*, pp. 558-567, 1988.

[8] T. Pavlidis and F. Ali, "A hierarchical syntactic shape analyzer", *IEEE Trans. PAMI*, Vol. 1, pp. 2-19, 1979.

[9] E. Rosch., C.B. Mervis, W.D. Gray, Johnson D.M., and Boyes-Braem P., "Basic objects in natural categories", *Cognitive Psychology*, Vol 8, pp. 382-439, 1976.

[10] P.L. Rosin, "Representing curves at their natural scales", *Pattern Recognition*, Vol. 25, pp. 1315-1325, 1992.

[11] P.L. Rosin, "Multi-scale representation and matching of curves using codons", *CVGIP: Graphical Models and Image Processing*, Vol. 55, pp. 286-310, 1993.

[12] P.L. Rosin and G.A.W. West, "Curve segmentation and representation by superellipses", *Proc. Australian and New Zealand Conference on Intelligent Information Systems*, Perth, Australia, 1993.

[13] G. Taubin, "Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation", *Trans. IEEE PAMI*, Vol. 13, pp. 1115-1138, 1991.

[14] A. Verri and A. Yuille, "Some perspective projection invariants", *J. Opt. Soc. Am.*, Vol. 5, pp. 426-431, 1988.

[15] D. Weinshall, "A hierarchy of invariant representations of 3D shape", in *Proc. Workshop on Qualitative Vision*, 97-106, 1993.

[16] G.A.W. West and P.L. Rosin, "Techniques for segmenting image curves into meaningful descriptions", *Pattern Recognition*, Vol. 24, pp. 643-652, 1991.

COMPUTING AND COMPARING DISTANCE-DRIVEN SKELETONS

GABRIELLA SANNITI DI BAJA
Istituto di Cibernetica, CNR,
80072 Arco Felice, Naples, Italy

and

EDOUARD THIEL
Equipe TIMC-IMAG, CERMO BP 53
38041 Grenoble Cedex 9, France

ABSTRACT

A unified distance-driven algorithm is presented to extract the skeleton of a digital pattern from its distance map. The algorithm equally runs whichever distance is selected to compute the distance map, among four commonly used path-based distance functions. The resulting skeletons are compared on a large set of 512×512 input pictures, in terms of reversibility and computation time.

1. Introduction

Distance maps, computed according to different distance functions, can be adopted to guide pattern skeletonization. On the discrete plane, path-based distances are commonly used, where the distance between two pixels is defined as the length of a shortest path linking them. The degree of approximation to the Euclidean distance depends on the number of different unit moves permitted along the path, and on the weights used to measure them. City-block distance (one unit move, unitary weight) and chessboard distance (two unit moves, both with unitary weight)) [1] are a natural choice on the square grid, but roughly approximate the Euclidean distance. Better approximations are obtained by using weighted distance functions allowing two (or three) differently weighted unit moves [2-5].

Skeletonization algorithms driven by the city-block distance d_1, the chessboard distance $d_{1,1}$, the two-weight distance $d_{3,4}$ and the three-weight distance $d_{5,7,11}$ can be respectively found in [6-9].

In this paper we provide a unique skeletonization algorithm equally running, whichever among the previous four distance functions is used to build the distance map DM of the pattern to be skeletonized. After computing the DM, the skeletal pixels (i.e. the centres of the maximal discs, the saddle points and the linking pixels) are identified and