Maximal Arc-Length Matching at Multiple Thresholds (MALT) for Variational Shape Recognition

Anjan Goswami
Dept. of Computer Sci. & Elec. Engg.
Univ. of South Florida, Tampa
agoswami@csee.usf.edu

Amitabha Mukerjee
Dept. of Computer Sci. & Engg.
Indian Instt of Technology, Kanpur
amit@iitk.ac.in

Abstract

Widely varying shapes are difficult to recognize especially when present as part of an occluded shape. Handling such situations requires the consolidation of data at different levels of confidence, which is the philosophy motivating the Maximal Arc-Length matching at multi-Thresholds algorithm (MALT). The algorithm matches segments from two contours and rates a pair of matches more highly if the arc-length shift in the two matches is equal.

Unlike traditional matching strategies, dependency on threshold selection is avoided by considering multiple thresholds and ranking the matches. All the models are tested in the polygonal domain, in which variational shape classes at different levels of deformation have been generated using vertex and edge deformation from a nominal model. Although in its current form the algorithm is restricted to polygonal contours, in this domain it exhibits very robust recognition even in the presence high deformations and occlusion.

1 Introduction

We consider model based polygonal or segmented shape recognition where the input shape may contain gross deformations and occlusions. In case of severe model deformation, recognition requires the application of some novel steps beyond the consolidation of individual feature matches. We propose the following steps for this problem:

- Representing shapes based on relative relations of sequence of edges. This is one oldest approach in biological shape analysis [15] and also recently have been tried in vision under names such as relative order indexing or relational indexing [8, 6, 3].

- Consolidate feature matches across multiple error-thresholds (different confidence'levels), and

- Express higher confidence if there are a large number of matching feature pairs with the same arc-length shift between them.

We develop a multi-threshold model of Hausdorff fraction where the decision is based on cumulative error in multiple error zones instead of any specific threshold parameters. We present empirical evidence indicating that the specific choice of differing error zones do not affect the rankings generated by the algorithm. This is further integrated with the dominant arc-length shift heuristic (section 3.2), which says in essence that if a model contains a large number of matching feature pairs which are shifted from the contour by the same arc-length shift, then this model is likely to be actually present in the image and with a similar shift. This model is developed under the belief that relative relation between the sequence of edges will be less sensitive under contour deformation and a shape will no longer be the same shape if these relations change arbitrarily. The combined algorithm, Maximal Arc-Length at multi-Threshold (MALT) is seen to be very robust against shape variation. The model has been tested by a comprehensive procedure for generating Variational Shape Classes, which has been developed by variations on the contours of 2D polygons.

A shape contour constructed from a sequence of control edges (linear or curved) can be represented by a small set of relations between each edge with one or more preceding edges (the supersegment). This representation is scale and displacement invariant and saves considerable auxiliary effort in indexing. Similar representation over any feature space can form quasi-invariants [4] which are of great significance in general object recognition schemes. Test results from the polygonal domain demonstrate successful recognition in some of the worst input shapes.

To obtain the input shapes, we furst deformed a given model to obtain a variational shape class. Deformed shape instances are combined to form occluded images, which are then presented as input shapes for disambiguating the models that may be present in the shape.

malized by perimeter) and the angle from the preceding edge within are varied within a specified tolerance zone (Table 1).

To ensure that all Images generated by this procedure are consistent, any self-intersecting or open polygonal chains arising from this procedure are pruned from the image set to ensure consistency in generated images.

	Shapes
C0101	
C0501	
C0505	
C0510	
C1515	

	Shapes
V05	
V10	
V20	
V30	
V45	

Table 1: Variational Shape Class created by Contour Deformation. a) Relative "chain" variation (C 0105 = length tolerance 01% of average edge length and angle tolerance 05% of 10°. b) Vertex variation (V 05 = 05% of average edge length.

M-0	M-1	M-2	M-3	M-4
62	30	26	15	32
M-5	M-6	M-7	M-8	M-9
18	8	17	40	37
M-10	M-11	M-12	M-13	M-14
6	12	6	56	18
M-15	M-16	M-17	M-18	M-19
37	27	41	43	8
M-20	M-21	M-22	M-23	M-24
24	51	15	10	15
M-25	M-26	M-27	M-28	M-29
18	39	24	13	26

Table 2: Database of Model Shapes (number of edges below each figure).

2   Shape Class

A shape is all geometrical information that remains when translation, scale and rotational effects are filtered out from an object [9]. In computer vision and image processing applications shapes are subjected to variation either due to inexact knowledge of transformations (Rigid body or Affine), natural variation in models (Biological Shapes) or perturbations due to the error in feature extraction etc. The notion of "Shape Class" captures these variations. Given the nominal shapes, finding out deformed samples of it from a occluded background is a very difficult problem. In [7], we addressed the problem of modeling shape variation and have shown different models of object recognition schemes with asymmetric Hausdorff measures.

2.1   Variational Shape Class

Variational shape classes are generated by applying a deformation to a nominal shape contour. (Examples are Shock Graphs [10, 17], Parametric Boundary variation [1], Manufacturing Tolerance models [16], Axis models [12]). We generate variational shape classes in a polygonal model by vertex deformation (perturbing each vertex within a rectangular tolerance zone proportional to the bounding box on the nominal shape or chain-edge perturbation in which the edge length (nor-

2.2   Occluded Variational Shapes

A further challenge in recognizing variational shapes arises in the problem of occlusion where the test image may be constituted by combining partial contours from a number of models, each of them having been deformed to differing degrees. Even in the simple domain of polygon recognition, this is seen as a difficult problem. Whereas considerable work has been done on occluded shapes [2, 11], and on shape class recognition [13, 17], the problem of recognizing a variational shape when it is occluded has not been formally modeled, either from the shape class perspective or from the recognition perspective.

A measure for the degree of overlap of two shapes can be constructed with area occlusion (area$(A \cap B) \div$ area $(A \cup B)$) or perimetric occlusion = (peri$(A \cap B) \div$ peri $(A \cup B)$). While this measure is easily generalized for three or more shapes, it does not reflect the asymmetric occlusion of different shapes - e.g. a small shape may be lost to the extent of 90% by area whereas the occlusion may be shown as 10%. An alternate measure may be to reflect the individual occlusions of each shape, but this involves $\binom{n}{2}$ relations for n objects. The results reported here use perimetric occlusion since the models are contour based.

E-2	E-3	E-4	E-5
80 edges	58 edges	55 edges	63 edges
P.O: .10	P.O: .25	P.O: .33	P.O: .14
exact	V:05 both	V:05 both	V15,V20
E-8	E-10	E-11	E-12
58 edges	54 edges	63 edges	62 edges
M-4, M-9	P.O: .36	P.O: .16	P.O: .18
and M-15	V25 both	V:40 both	V50 both

Table 3: Experimental Images formed by combining various deformations of the models in Table 2. Below each Image: Total Number of Edges, Perimetric Occlusion, models present, and shape variation. E-5, 10, 11, 12 have the same two objects (duck and elephant) with increasing vertex perturbation.

3 Maximal Arc-Length at multi-Threshold (MALT)

A variant of Hausdorff measure can be defined as the maximum number of points in the model such that the distance is below a given error threshold,

$$h_k(M_m, I_n) \leq \psi,$$

where $h_k()$ denotes a distance from m-th Model feature to n-th input feature where the n-th input feature is closest to m-th model feature. Let $k_\psi(M, I)$ denote the maximum k for which the above relation is true, and $\| M \|$ is the number of features in M. The ratio

$$F_\psi(M_j, I) = \frac{k_\psi(M_j, I)}{\| M_j \|}$$

reflects the fraction of the shape M present in I at the confidence indicated by the threshold ψ [14].

3.1 Multi-Threshold Fraction (MT)

Instead of fixing on a specific threshold to identify k_ψ, it may be possible to fuse the conclusions that can be arrived by consolidating the data from different levels of confidence, which is the motivation for the multi-threshold (MT) Hausdorff Fraction model.

In particular, it is unlikely that any model with a single confidence level (single threshold cutoff) would be able to identify severely deformed models, where different features may agree to differing degrees [7].

Given a list of $\mu + 1$ error thresholds ψ_i, in increasing order $\psi_0, \psi_1 \ldots \psi_\mu$, (defining μ error intervals or zones), the multi-threshold fraction $F_{MT}(M_j, I)$ sums the fractions at differing levels of error:

$$F_{MT}(M_j, I) = 1/\mu \left(\sum_{i=0}^{i=\mu} F_{\psi_i}(M_j, I) \right),$$

where, $F_\psi(M_j, I) = \frac{k_\psi(M_j, I)}{\| M_j \|}$ and $k_\psi(M, I)$ denotes the maximum k for which $h_k(M_m, I_n) \leq \psi$.

Since the matches at low values of ψ are also present in the higher values, these edges get reinforced more than features detected for higher ψ which may contain some noise. This process reflects an harmonic assumption that a match found only in the last error zone has $\frac{1}{(\mu - r)}$ less validity than a match at error-zone r. An alternate geometric decay model - that the last error zone has $k^{-(\mu-r)}$ less validity, introduces an additional parameter - the confidence decay rate k - and also did not appear to work very well in preliminary tests.

3.2 Dominant Arc-Length Shift

If the features being detected happen to constitute successive segments in both the image and model shapes, then there is a strong likelihood of a match. More generally, if the j-th and i-th Input shape features are found to be closest to the (j+τ)-th and (i+τ)-th model features then it is possible that shifting the model by τ edges would result in a match between the contour fragments. This belief would be strengthened if a large number of other matching feature pairs also exhibited the same shift, which is the basis of the **Dominant Arc-Shift Heuristic**:

- The dominant arc-length shift (that which is common to the largest set of matched feature-pairs) is

likely indicative of an actual arc-length shift that matches a contour fragment between the image and the deformed model.

- The greater the dominance of the dominant arc-shift, the higher the confidence in the match. (see Figure 1).

This heuristic provides a mechanism for combining multiple feature matches, an idea that has also been proposed recently in [5].

This heuristic is combined with the multi-threshold measure (F_{MT}) to obtain the MALT algorithm. To reflect the higher importance of a greater number of matches at the same arc-length shift, we have chosen the simple measure of counting all feature pairs at each arc-length shift ($p_{\psi i}$), weighted by $\frac{2}{\|M\|}$ to discount the effect of models with many more edges. The sum of this factor over all thresholds is the arc-length dominance measure F_{AL}. The *Maximal Arc-Length at multi-Threshold* (F_{MALT}) measure combines the Multi-Threshold (MT) Hausdorff fraction F_{MT} with this arc-length dominance measure F_{AL}:

$$F_{MALT}(M_j, I) = \frac{1}{\mu}(\sum_{i=0}^{i=\mu}(F_{\psi i}(M_j, I) + 2p_{\psi_i}/\|M\|))$$

Note that p_ψ is the number of matching pairs and not the number of matching edges; thus if f edges from the Model match successively with f edges from the Image (with a constant shift of τ), then $p_\psi = \binom{f}{2}$, which can quickly overwhelm the contribution of MT if there are many sequence matches.

Dominance Fraction

The frequency distribution of the arc-length shift is an important indicator of the confidence one can have in the MALT estimate. The dominant arc-length shift, which has the highest number of matching edge pairs, is strongly indicative of the presence of a long arc-chain from the model in the image.

The Dominance Fraction $\mathcal{D}_\mathcal{M}$ is the maximum number of edges in a single arc-length shift divided by the total number of edges in the model ($\frac{f}{\|M\|}$) (see figure 1). The dominance fraction is a measure of the percentage of the model we are highly confident of finding in the image. It is also directly related to F_{MALT} which is the summation of $\binom{f}{2}$ and is dominated by the highest f.

3.3 Error Threshold Values

The MT model was formulated based on the belief that cumulative data over multiple thresholds would
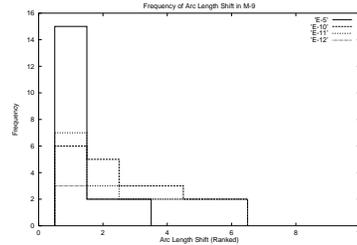
Figure 1: *Effect of Increasing Variation on dominance fraction* As the model variation increases, the arc-length shift is no longer concentrated at a dominant value. Dominance fraction in E-5, E-10, E-11 and E-12 are respectively 0.41, 0.16, 0.19, 0.08.

remain more or less the same if the intervals changed. We have tested this premise empirically over repeated experiments and a sample at different error levels is shown in Figure 2 Furthermore, as the error threshold increases, more and more noisy data will enter and the rate of growth flattens out. This implies that the maximum error threshold ψ_μ is also not very critical. In the experiments that follow, we have used $\mu = 15$ and error thresholds with a maximum threshold $\psi_\mu = 0.6$, in a geometric thresholding scheme (higher errors correspond to a longer error interval). However, as noted in Figure 2 the results are more or less identical for differing choices of these numbers.

4 Algorithm

The basic idea of the algorithm is to find, for each feature in each model, the minimum distance to an image feature and the arc-length shift. The arc-length shift between these two feature is the difference in indices between the two mapped features.

Given a list of $\mu + 1$ error thresholds ψ_i, if the minimum distance is below some threshold ψ_r then the Fraction count F_{MT} is incremented for all thresholds ψ_r and above and also, an index of frequencies is maintained for each occurring arc-length-shift τ indicating the number of feature matches with this arc-length shift. Eventually, if the number of edge-matches at a arc-length shift value is n then there are $\binom{n}{2}$ pairs of matches at the same arc-length distance and the MALT measure F_{MALT} is incremented appropriately.

1. For each Model M (0 to m),
BEGIN
  2. For each feature e_i^M of model M
  BEGIN
    3. min_dist(i)=∞
    4. For each feature e_j^I of image I
      if $dist(e_i^M, e_j^I) < min\_dist(i)$

{ min_dist(i)=dist(i,j); min-posn p=j; }
5. Arc-Length shift $\tau=$ Abs(p-i)
6. Find error-threshold $\psi_r, r \in (1 \dots \mu)$ s.t.
   $\psi_r > min\_dist(i) > \psi_{r-1}$ (Binary Search)
7. F_{MT} += $(\mu - r + 1) \times \frac{1}{\|M\|}$
8. Increment Frequency of Arc length Shift
   at threshold ψ_r : Frequency[r][τ]++
END
9. For each f = Frequency[r][τ], if f $\geq$ 2
   F_{AL} += $\binom{f}{2} \times \frac{2}{\|M\|}$
10. Return ($F_{MALT} = \frac{(F_{MT}+F_{AL})}{\mu}$), $F_{MT} = \frac{F_{MT}}{\mu}$)
END
11. Best matches are k highest F_{MT} or F_{MALT}

Complexity

The complexity of this algorithm is determined by step 4, and is $O(mNn)$ where m is the number of models, N the maximum number of model features (max $\| M \|$), and n the number of features in the input image. Normally, there are more input image features than error-thresholds $(n > \mu)$, so steps 6 and 9 with costs of $O(mN \log \mu)$ and $O(mN\mu)$ are dominated by step 4.
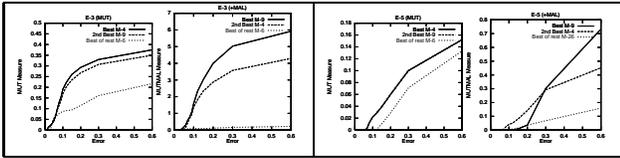
Figure 2: *Effect of Additional Arc Length measure* . The results at different error zones for the Hausdorff Fraction with Multi-Threshold F_{MT} (left), and for with Maximal Arc-Length ("MALT", right). The match fractions remain unaltered by threshold changes.

5 Results

We use a database of 30 hand-drawn model shapes (Table 2). Table 1 shows some samples of generated variational shape classes and table 3 shows 12 experimental shapes. Table 4 shows results of the recognition algorithm with E-2, E-3 and E-4.

5.1 Increasing Shape Variation

If a model appears in a highly deformed state, then it is possible that some other models may contain supersegments that are a better approximation to some of the deformed edges. The "MALT" algorithm is able

E-2		E-3		E-4	
F_{MT}	F_{MALT}	F_{MT}	F_{MALT}	F_{MT}	F_{MALT}
.82	16.94	.35	5.92	.42	9.30
.81	13.10	.37	4.29	.27	1.57
.16	0.16	.21	0.22	.20	0.20

Table 4: *Results for some simple cases*: E-2 (with M-17, M-26, M-8), E-3 (with M-9,M-4, M-0),E-4 (M-9, M-4, M-0). For each case first two models are constituent candidates; the "best of the rest" model is also reported for comparison.

	E-5			E-10	
M	F_{MT}	F_{MALT}	M	F_{MT}	F_{MALT}
9	.13	.72	9	0.11	.27
4	.15	.45	4	0.12	0.19
26	.11	.16	15	0.08	0.17
8	.10	.15	26	0.10	.15
27	.12	.14	0	0.09	0.14
	E-11			E-12	
M	F_{MT}	F_{MALT}	M	F_{MT}	F_{MALT}
9	.09	.20	4	.12	.26
4	.11	.16	23	.13	0.16
8	.08	0.14	8	0.09	0.15
25	.10	.14	1	0.11	0.14
10	.13	.13	10	.13	.11

Table 5: *Results for increased perturbation and occlusion*. All these shapes involve the duck and the elephant (M-4 and M-9). Even in E-11, where the duck shape is hardly recognizable even by humans, the algorithm manages to find it (barely!). In E-12 however, where the vertex perturbation is 50% of the average edge length, it fails to find the elephant (M-9) among the top 5. Data is shown for the 0.6 error level in Figure 3.
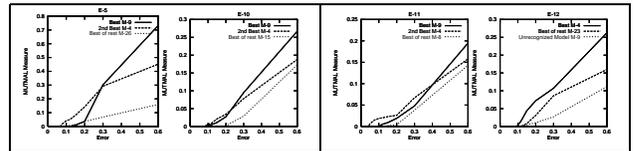
Figure 3: *Effect of high variation*. As the variation from the model shapes increases, results for the "best of the rest" (a model not actually present in the shape) approaches that of a model present due to high noise levels in images E-11 and E-12, which are highly noisy with 40-50% vertex perturbation (see figures in Table 3).

to identify both elephant and duck in E-5, E-10, E-11 where the shapes are highly deformed but it fails to recognize the elephant in E-12 where human beings can also hardly interpret the elephant (Table: 5).

E-9			E-6		
M	F_{MT}	MALT	M	F_{MT}	MALT
9	0.42	9.17	9	0.41	9.61
4	0.19	0.43	4	0.15	0.26
14	0.11	0.23	14	0.11	0.19
3	0.13	0.20	18	0.09	0.19
18	0.10	0.19	6	0.18	0.18

Table 6: As the duck beak gradually disappears into the back of the elephant, the multi-threshold fraction F_{MT} ranks the house (M-6) above the duck in the barely recognizable E-6, but MALT recognizes it.

The shape E-6 is constructed using only five edge segments of the duck in it; of these two are intersecting. Even human beings can hardly recognize it (Table 6).

As the model variation increases, more spurious edges match some of the model features, displacing features at the correct arc-length. Thus the number of feature-pairs at the same arc-length shift reduce (Figure 1) and the dominance fraction decreases, and with it, the confidence of the matching process.

6 Conclusion

The main contribution of this work is in the Arc-Length maximizing search, which is a form of sub-string matching in partial image contours. Two observations are combined in this work − results from multiple thresholds are combined to reduce dependence on any pre-set threshold value, and also higher confidence is allocated for results that match multiple segments with the same "arc-length" shift. The model has been tested on polygonal data but is useful for matching contours that are generated based on the same set of control points as in sketches and artwork, CAD drawings, etc.

The extent of deformation and occlusion permitted are exceptional and in some cases appear to be superior to the human ability to recognize a noisy shape occluded in a contour. This clearly indicates the merit in the basic philosophy of arc-shift matching, though it needs to be developed further for handling arbitrary contour data.

References

[1] B. Aldefeld. Variation of geometries based on a geometric-reasoning method. *Computer Aided Design, April 1988*, 20(3):117–126, 1988.

[2] Nirwan Ansari and Edward J. Delp. Partial shape recognition: A landmark-based approach. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 12(5):470–483, May 1990.

[3] Serge Belongie and Jitendra Malik. Matching with shape contexts. *Proceedings of the IEEE Workshop on Content-based Access of Image and Video Libraries*, 2000.

[4] T. O. Binford, T. S. Levitt, and W. B Mann. Bayesian inference in model-based machine vision. In T. S. Levitt L. N. Kanal and J. F. Lemmer, editors, *Uncertainty in Artificial Intelligence 3*, pages 73–95. Elsevier Science Publishers B. V. (North-Holland), 1989.

[5] Yuri Boykov and D. Huttenlocher. A new bayesian framework for object recognition. In *IEEE Computer Vision and Pattern Recognition*, pages 517–523, 1999.

[6] M. S. Costa and L.G Shapiro. Scene analysis using appearance-based models and relational indexing. *IEEE Computer Society International Symposium on Computer Vision*, pages 103–108, 1995.

[7] Anjan Goswami and Amitabha Mukerjee. Modeling and recognition of planar shape classes. In *Proc. 4th Intl Conf. Advances in Pattern Recognition and Digital Techniques, (ICAPRDT-99)*, pages 170–175, Calcutta, December 1999.

[8] B. Huet and E. Hancock. Relational histograms for shape indexing. *Proc. of ICCV*, pages 563–569, January 1998.

[9] D. G Kendall. The diffusion of shape. *Advances in Applied Probability*, 9:428–430, 1977.

[10] Benjamin B. Kimia, Allen R. Tannenbaum, and Steven W. Zucker. Shapes, shocks, and diffusions i: The components of two-dimensional shape and reaction-diffusion space. Technical report, LEM-Brown, August 1995.

[11] Mark W. Koch and Rangasami L. Kashyap. Using polygons to recognize and locate partially occluded objects. *IEEE Transcation on Pattern Recognition and Machine Intelligence*, 9(4):483–493, July 1987.

[12] O. Kubler and R. L. Ogniewicz. Hierarchic voronoi skeletons. *Pattern Recognition*, 28(3):343–359, 1995.

[13] Konstantin Y. Kupeev and Victor L. Brailovosky. A reinforced random algorithm for a partial contour perceptual similarity problem. *Pattern Recognition Letters*, 19:287–297, 1998.

[14] Clark F. Olson. A probabilistic formulation for hausdorff matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Santa Barbara, CA*, pages 150–156, 1998.

[15] C. R Rao. Test of significance in multivariate analysis. *Biometrika*, 35:58–79, 1948.

[16] A.A.G Requicha and Steve Chan. Representation of geometric features, tolerances, and attributes in solid modelers based on constructive geometry. *IEEE J. Robotics and Automation, September 1986*, RA-2(3):156–166, 1986.

[17] Kaleem Siddiqi, Ali Shokoufandeh, Sven J. Dickinson, and Steven W. Zucker. Shock graphs and shape matching. Technical Report CVC TR-97-002/CS TR-1147/RU DCS-TR-345, CS, Yale University, October 1997.