CMU-CS-85-122

Alignment and Connection of Fragmented Linear Features in Aerial Imagery

David M. McKeown, Jr.
John F. Pane

April 23, 1985

Abstract

Computer vision systems that attempt to extract cultural features from aerial imagery are often forced to interpret segmentations where the actual features are broken into numerous segments or fragments. For example, roads and road-like features are difficult to completely segment due to occlusions, poor contrast with their surroundings, and changes in surface material. Often the nature of the segmentation process is designed to err toward oversegmentation of the image, since the joining of feature descriptions is believed to be simpler than their decomposition. No matter what the cause, it is necessary to aggregate these incomplete segmentations, filling in missing information, in order to reason about the overall scene interpretation. This paper describes a method to select sets of such fragments as candidates for alignment into a single region, as well as a procedure to generate new linear regions that are linked composites of the original sets of fragments. Portions of the composite region that lie between pairs of the original fragments are approximated with a spline. The resulting composite region can be used to predict the areas in which to search for missing components of the cultural feature.

This research partially sponsored by the Defense Advanced Research Projects Agency (DOD), ARPA Order No. 3597, monitored by the Air Force Avionics Laboratory Under Contract F33615-81-K-1539 and by the Defense Mapping Agency Under Contract DMA 800-85-C-0009. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency, of the Defense Mapping Agency, or the US Government.

To be presented at the IEEE Conference on Computer Vision and Pattern Recognition, June 9-13, 1985 in San Francisco, California.

Table of Contents

List of Figures

Alignment and Connection of Fragmented Linear Features in Aerial Imagery*

David M. McKeown, Jr.
John F. Pane
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pa. 15213

Abstract

Computer vision systems that attempt to extract cultural features from aerial imagery are often forced to interpret segmentations where the actual features are broken into numerous segments or fragments. For example, roads and road-like features are difficult to completely segment due to occlusions, poor contrast with their surroundings, and changes in surface material. Often the nature of the segmentation process is designed to err toward oversegmentation of the image, since the joining of feature descriptions is believed to be simpler than their decomposition. No matter what the cause, it is necessary to aggregate these incomplete segmentations, filling in missing information, in order to reason about the overall scene interpretation. This paper describes a method to select sets of such fragments as candidates for alignment into a single region, as well as a procedure to generate new linear regions that are linked composites of the original sets of fragments. Portions of the composite region that lie between pairs of the original fragments are approximated with a spline. Therefore, the resulting composite region can be used to predict the areas in which to search for missing components of the cultural feature.

---

*This research partially sponsored by the Defense Advanced Research Projects Agency (DOD), ARPA Order No. 3597, monitored by the Air Force Avionics Laboratory Under Contract F33615-81-K-1539 and by the Defense Mapping Agency Under Contract DMA 800-85-C-0009. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency, of the Defense Mapping Agency, or the US Government.

1. Introduction

Computer vision systems that attempt to extract cultural features from aerial imagery are often forced to interpret segmentations where the actual features are broken into numerous segments or fragments. For example, roads and road-like features are difficult to completely segment due to occlusions, poor contrast with their surroundings, and changes in surface material. Often the nature of the segmentation process is designed to err toward oversegmentation of the image, since the joining of feature descriptions is believed to be simpler than their decomposition. No matter what the cause, it is necessary to aggregate these incomplete segmentations, filling in missing information, in order to reason about the overall scene interpretation.

This paper describes ALIGN, a program which hypothesizes and aligns linear regions in aerial imagery. This program is implemented as a component of SPAM, a rule-based system for airport scene interpretation[1, 2]. The image segmentation fragments are derived from a low-level image processing program that generates image segmentations using a region-growing technique[3]. The region-growing process merges regions based on spectral similarity, size, and shape criteria such as compactness and linearity. The segmentations are projected through an image camera model into a geodetic <latitude,longitude> coordinate system. SPAM then attempts to use the image segmentation to develop an interpretation of the airport scene, and invokes ALIGN on those regions that are determined to be candidates for interpretation as roads, taxiways, or runways. Photo 1-1 is an area of National Airport in Washington, D.C. containing roads near the terminal building. Figure 1-2 shows linear regions that were segmented in this portion of the image.

Feature alignment and connection proceeds in two phases:

1. The selection of candidate regions for possible alignment within an area of the image.
2. The evaluation of candidate regions to select the best connection.

ALIGN uses a medial axis to represent the shape of each region being aligned. It prunes candidates for alignment based on geometric evaluation of spatial constraints. Once pairs of regions are selected, a spline interpolation is used to smoothly join the aligned regions. A composite output region is generated whose medial axis is the constructed spline. Areas of the image that are not contained within the original candidates may be contained within the composite output region. These areas are used by SPAM to predict where to look for missing segments or to invoke other methods for feature extraction[2].

Section 2 discusses the extraction of the medial axis and other shape descriptions from the segmentation fragments. In Section 3 we describe a method to select sets of such fragments as candidates for alignment into a single linear region. This initial analysis greatly prunes the number of regions in a subarea of the image that are actually evaluated for subsequent alignment into a composite region.

Section 4 describes various geometric constraints used to further evaluate and prune region candidates for alignment. Section 5 discusses linking of the candidates and generation of a composite region based on the spline approximation. Section 6 gives some examples of the alignment procedure using machine generated segmentations from an airport scene. Section 7 describes future work and suggestions for improvement.

2. Region Representation

ALIGN is invoked by SPAM with a a set of linear image region candidates. Each region is represented as a polygon whose vertices are points in a geodetic <latitude/longitude> coordinate system. Since regions are represented by their ground coordinates, ALIGN is not constrained to work with features from a single image, and is capable of working with fragments from different images of the same ground area.

Regions are chosen by SPAM based on their linearity, which makes them good candidates as fragments of roads, runways and taxiways. The perimeters of these polygons are generally very jagged as a result of the region growing process. The shapes of the regions are composed of both straight and curved portions. Initial attempts to use techniques based on Hough transforms[4], which are often used to aggregate edges and/or line

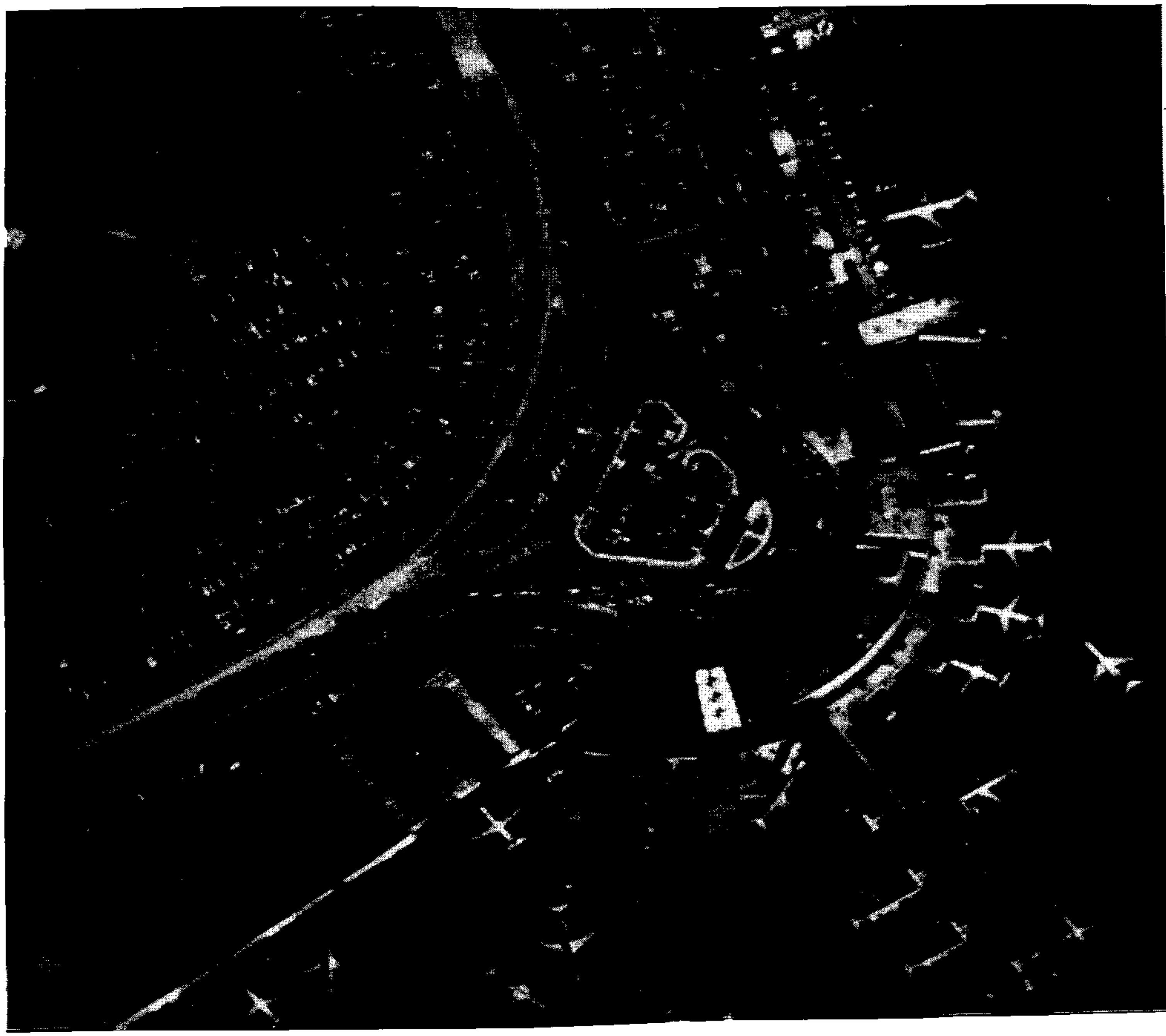


Figure 1-1: Roads Near Terminal Building of National Airport

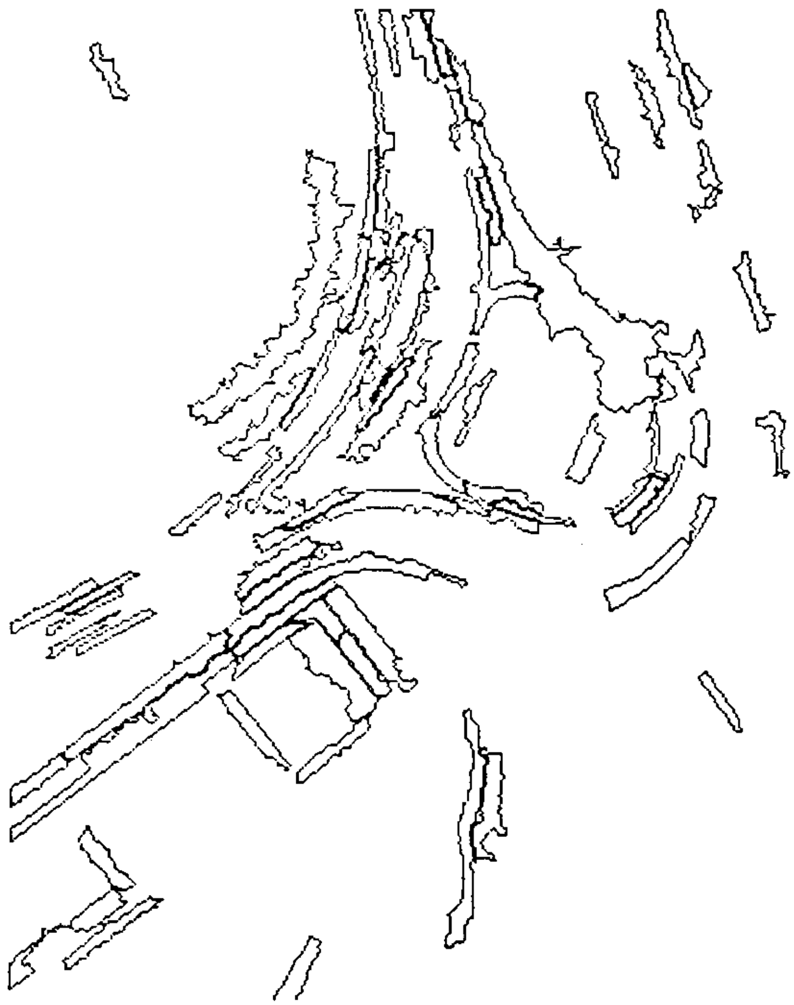


Figure 1-2: Region Segmentation Near Terminal Building

segments, did not provide acceptable results in our application.

During the candidate selection process a simple description of the regions is desired, one which contains position, direction or orientation, and a representation of the extent of the feature. Attempts to fit a simple linear approximation to the set of points were not satisfactory for regions which have significant curvature. The use of a piecewise linear approximation using the medial axis of the region proved to be a useful representation.

A first attempt used a direct medial axis transformation[5, 6] of the fragments. Because of the jagged nature of the boundary of many regions, this method often resulted in a linear approximation which did not provide a complete description of the region. For instance, the medial axis transformation usually did not traverse the entire length of the region, and often did not provide a good indication of the true direction of the region. This is because the medial axis algorithm actually found a series of disjoint spines, and our implementation simply chose the longest of these.

To solve this problem, a Fourier approximation of the region is done to smooth some of the roughness of the original region.[6] Our Fourier description uses nine orders for the approximation. The result is a description which provides a very good representation of the fragment, with the local jaggedness filtered. Running a medial axis transform on these Fourier descriptions results in a much better linear approximation, which is used to represent the feature throughout the rest of the program. Figure 2-2 shows the Fourier shape approximation of the region in Figure 2-1, and the resulting medial axis is shown in Figure 2-3. The Fourier description also generates values for the major and minor axes of a bounding ellipse. The minor axis is used as a representation of the width of the feature.

3. Selection of Candidates for Alignment

Original attempts to align the regions before pruning the search area of unsuitable regions did not yield acceptable results. The combinatorics of performing the geometric analysis on the entire domain of region fragments proved inefficient, and led to results which were confusing or inappropriate for the road alignment application at hand. Doing some pre-alignment selection over the fragment domain prunes the search area, and helps to strengthen the alignment heuristics used later in the algorithm.

The segments are first sorted by length, because the longer fragments have a higher reliability as hypothesized road fragments. Then, beginning with the longest fragment supplied to the program, a locus in latitude/longitude space is created, which is simply an expansion of the minimum bounding rectangle (MBR) of the fragment. Any other fragment whose MBR intersects this locus, is selected as a candidate, and a new locus is formed by iteratively expanding a bounding rectangle which contains both the original

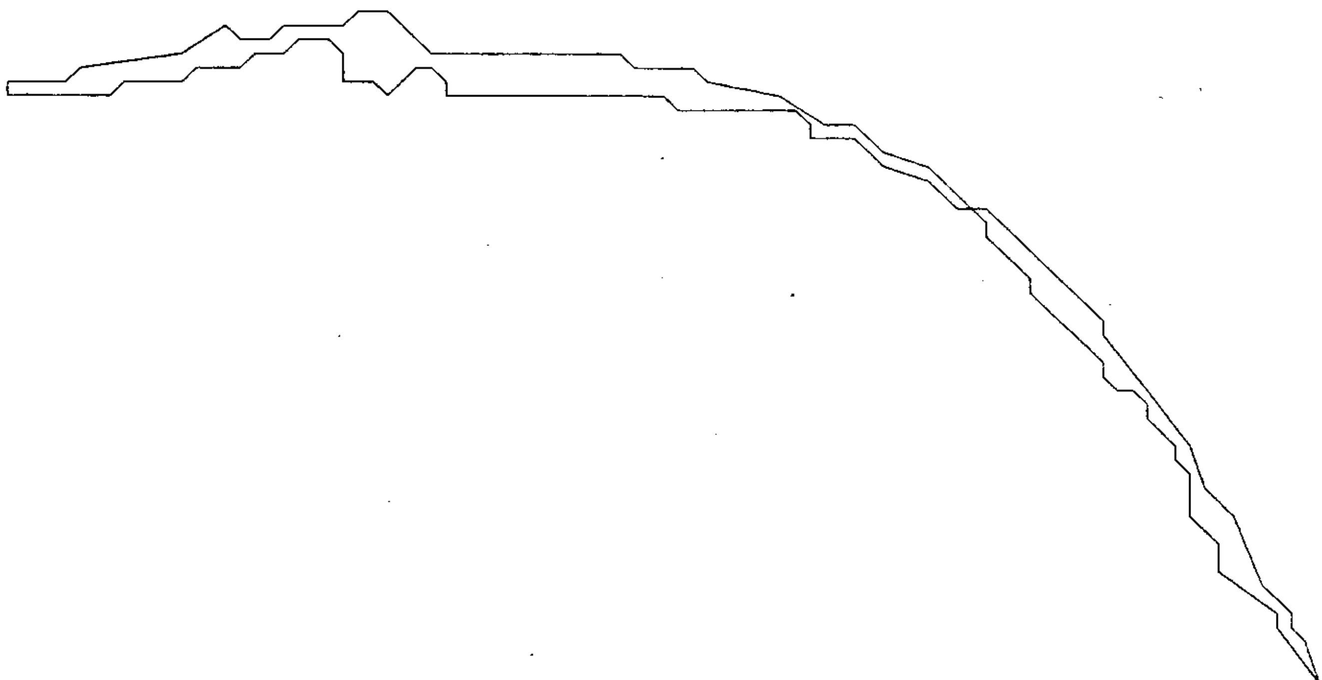


Figure 2-1: The fragments as extracted from the image are ragged.

and all newly selected fragments. This eliminates fragments which are far apart (relative to their lengths) from consideration for future criteria for alignment. The distance tolerance can be adjusted by altering the factors by which the MBR is expanded. Note that this relative selection criteria allows the algorithm to be used on sets of fragments without regard to their actual sizes.

The length (longer side) of the MBR is expanded by a different (greater) factor than the width (shorter side), because roads which align with a given feature will lie somewhere off the end of the feature, and not somewhere along its side. The expansion of the width cannot be ignored, however, because this metric is dependent on orientation, a well known weakness in such a simple representation. For instance, a straight line which has a

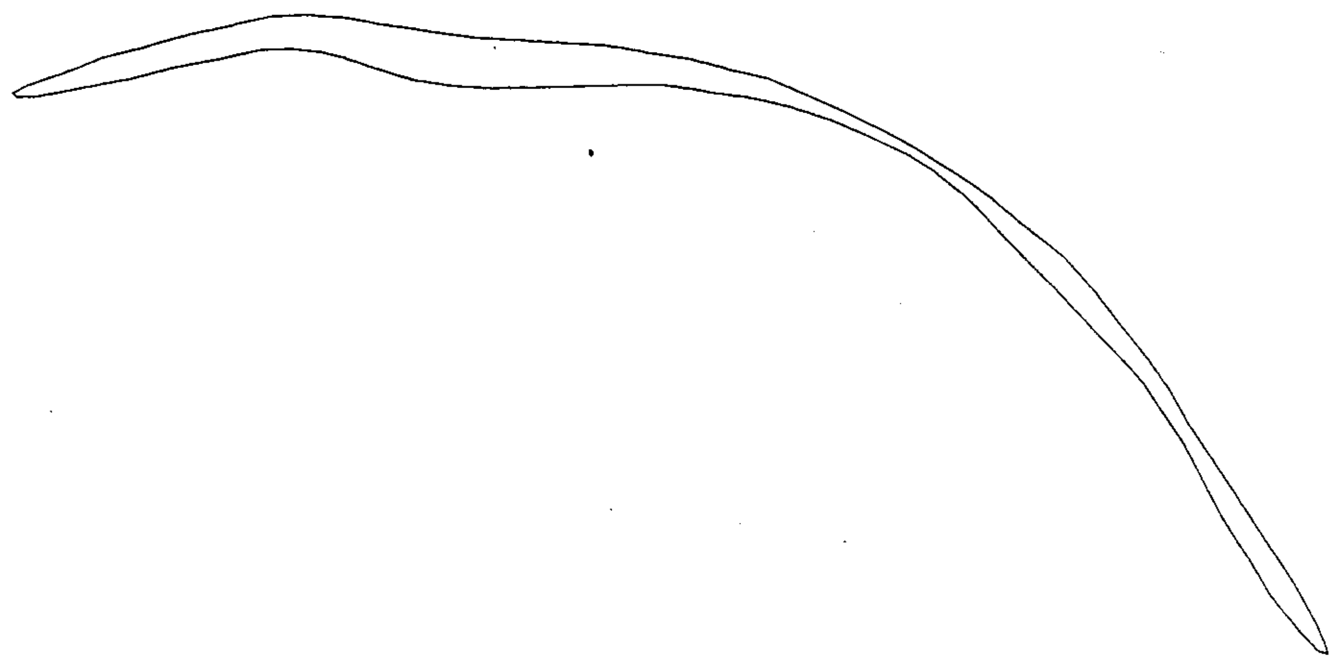


Figure 2-2: Fourier approximation of the fragment in Figure 2-1.

diagonal orientation with respect to the coordinate system will have equal widths and lengths, even though a perfect model would assign a width of zero to a straight line. Figure 3-1 illustrates this problem.

Another weakness in this method of determining the candidate segments is that the repeated expansion of the MBR can cause it to expand to include all of the regions supplied to the program. This is especially likely when there is a very dense set of regions supplied, because of the iterative expansion of the locus. These problems can be solved by using a more complex representation of the locus. This paper does not explore such possibilities, because it was found that the current selection method is adequate for this application. The problems inherent in this selection method are handled satisfactorily later in the

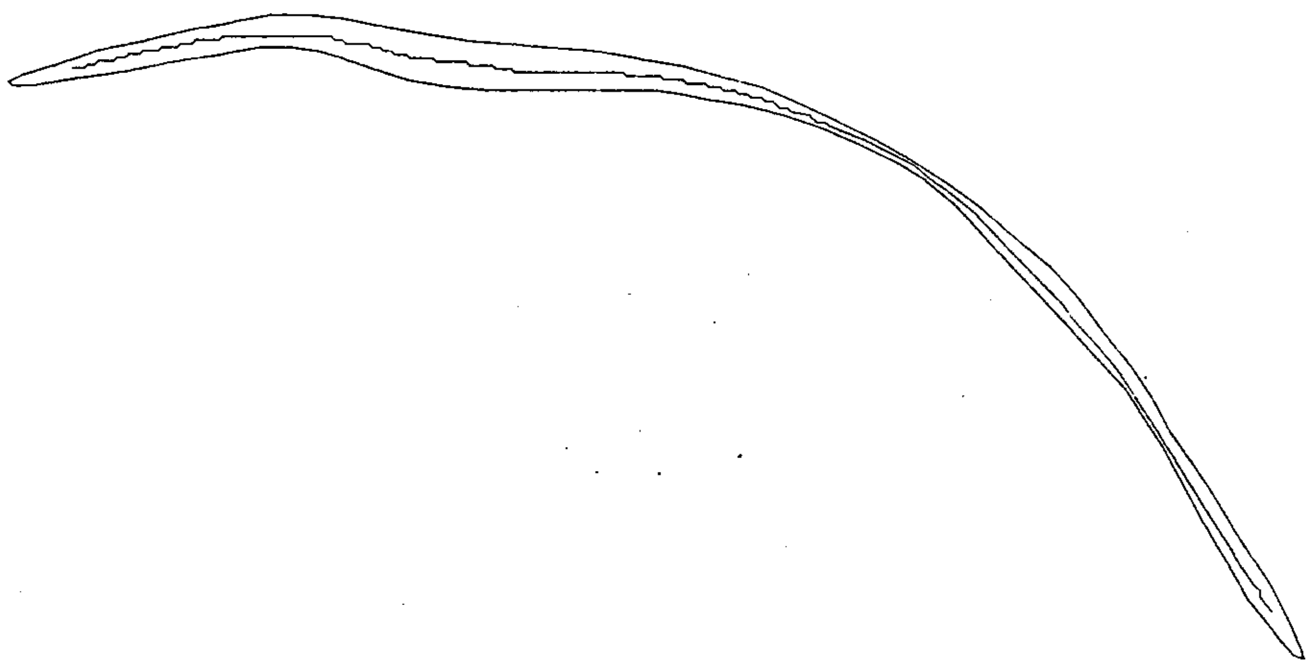


Figure 2-3: The Fourier approximation of the fragment with its medial axis.

program.

The values of the expansion coefficients were empirically determined to be 50% in the length direction and 20% in the width direction. A possible improvement may be obtained by basing these coefficients on the length to width ratio of the feature in question. Once again, this possibly was not explored, because of the satisfactory results which are obtained using this rough metric.

In the following section we discuss how each group of candidate regions is processed by the alignment algorithm, which determines which (if any) of the fragments can be joined.

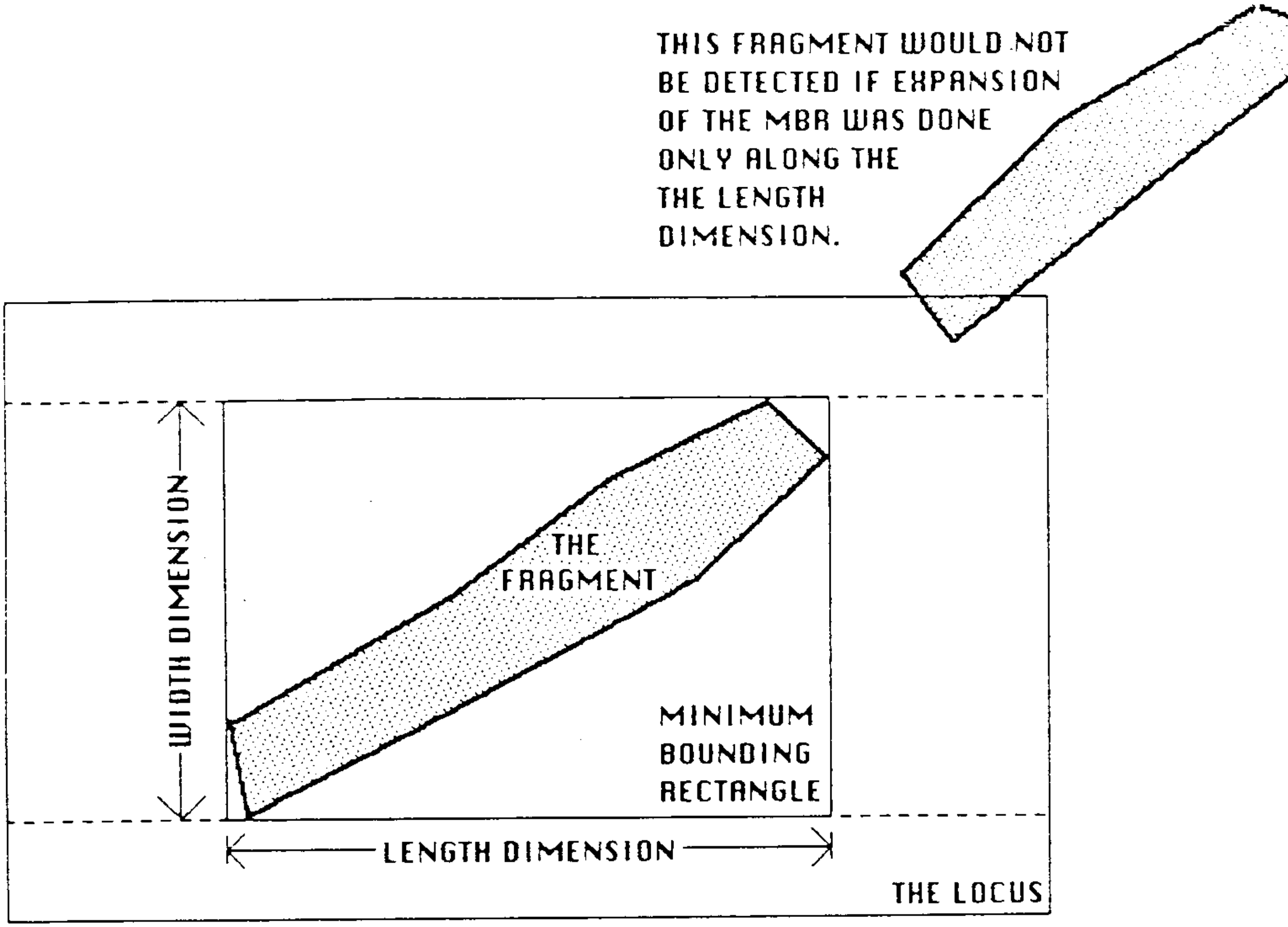


Figure 3-1: Using the MBR to Define a Search Area For Neighboring Regions

4. Criteria for Alignment

In order for a candidate set of fragments to satisfy our alignment tests, they must meet the following criteria:

1. The fragments must be in close physical proximity.
2. They must not overlap.
3. They must have similar orientations.
4. They must connect smoothly, without a bump during the transition from one fragment to the next.

4.1. Proximity

The fragments must be in close physical proximity. The fragments are required to be within a certain distance of one another. In our implementation this distance is equal to the length of the longer of the two fragments in question. This provides tolerance of fairly large gaps in the fragments, without allowing unreasonably long portions of the fragment to be composed of hypothesized (filler) regions.

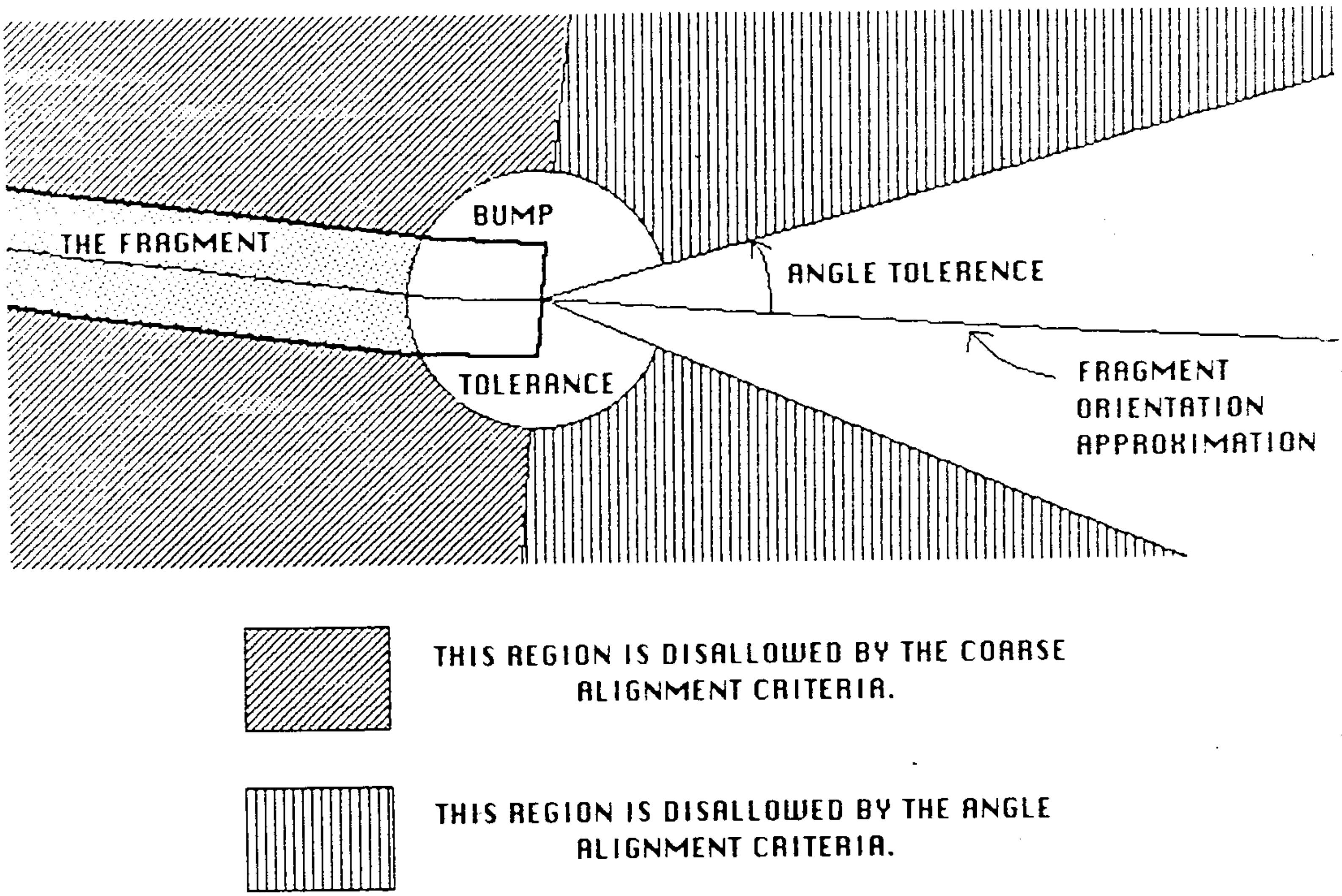
4.2. Overlap

The fragments must coarsely align. In order for fragments to satisfy the criteria for potential alignment, they must not overlap, and must be arranged such that no portions are broadside to one another. In other words, although the fragments are not constrained to be collinear, they must not require the curve that joins them to double back on itself. Figure 4-1 depicts the zones relative to a given fragment that are forbidden by this coarse alignment constraint. Other fragments having one or more points in the forbidden zone are eliminated as a candidates for alignment with this fragment.

This geometric information is determined by relative distances between endpoints and next-to-endpoints of the two fragments. For instance in the model shown in Figure 4-2, there is a series of geometric constraints that can be applied to the *distances* between labeled points. In all, there are four sets of constraints to be applied depending on the relative order of the segments being compared, and on the internal order of representation of the points which describe these segments.

4.3. Orientation

Thirdly, in order for fragments to be aligned, their orientations must be within some tolerance of the orientation of the line which connects them. Fragments that satisfy this criteria must be oriented in a similar direction. To allow for some tolerance of curvature, a line is fit to only the last third of the points on the medial axis representing the segment. This line represents local orientation, rather than an overall orientation of the region. This is important in features which curve significantly since we are most interested in

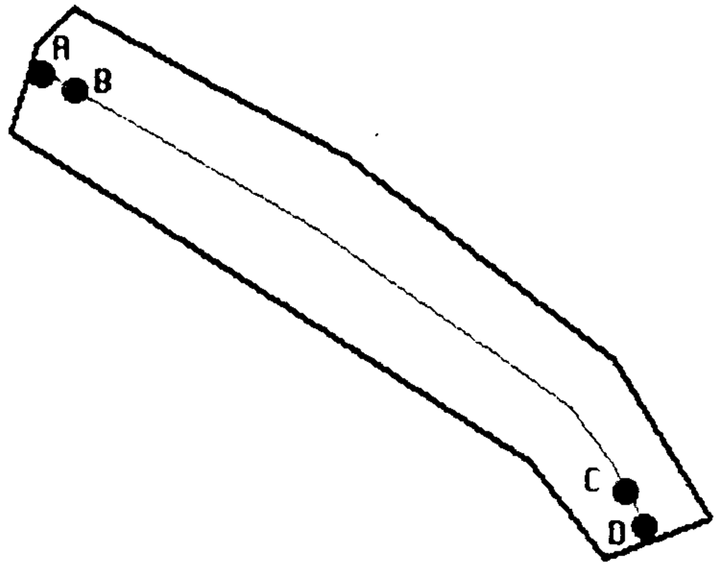


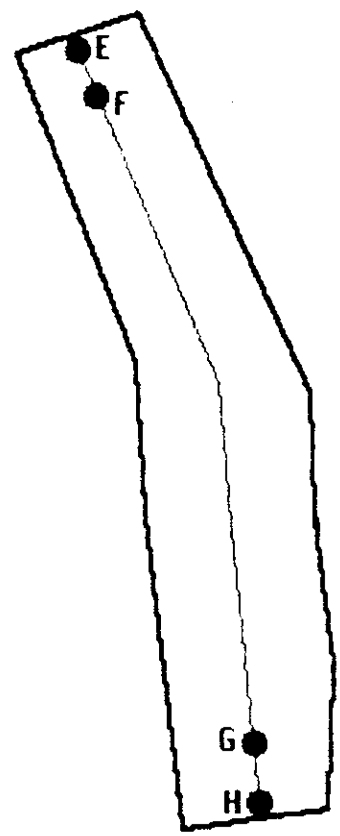
This Figure illustrates the effects of the alignment criteria. The unshaded regions indicate space where an acceptable fragment for alignment can terminate.

Figure 4-1: Tolerance Zones for Alignment Criteria

obtaining a good approximation of the direction of the ends of the region.

Using a fixed percentage of points poses some problems, however. If the segment contains very few points, the last third might seriously misrepresent the actual direction of the segment. This is especially problematic due to the above mentioned jaggedness of the features. An empirical solution to this problem was to force a minimum of 4 points to be





This model demonstrates the points on the fragment approximations which are used for the coarse alignment tests. In this particular instance, the series of tests that apply are labeled #1 in the accompanying chart.

If the features line up, one of the following columns of distance inequality conditions will be satisfied:

#1	#2	#3	#4
ae > de	ae > de	ae < de	ae < de
ah > dh	ah > dh	ah < dh	ah < dh
be < ae	be < ae	be > ae	be > ae
ce > de	ce > de	ce < de	ce < de
bh < ah	bh < ah	bh > ah	bh > ah
ch > dh	ch > dh	ch < dh	ch < dh
df > de	df < de	df > de	df < de
dg < dh	dg > dh	dg < dh	dg > dh
af > ae	af < ae	af > ae	af < ae
ag < ah	ag > ah	ag < ah	ag > ah

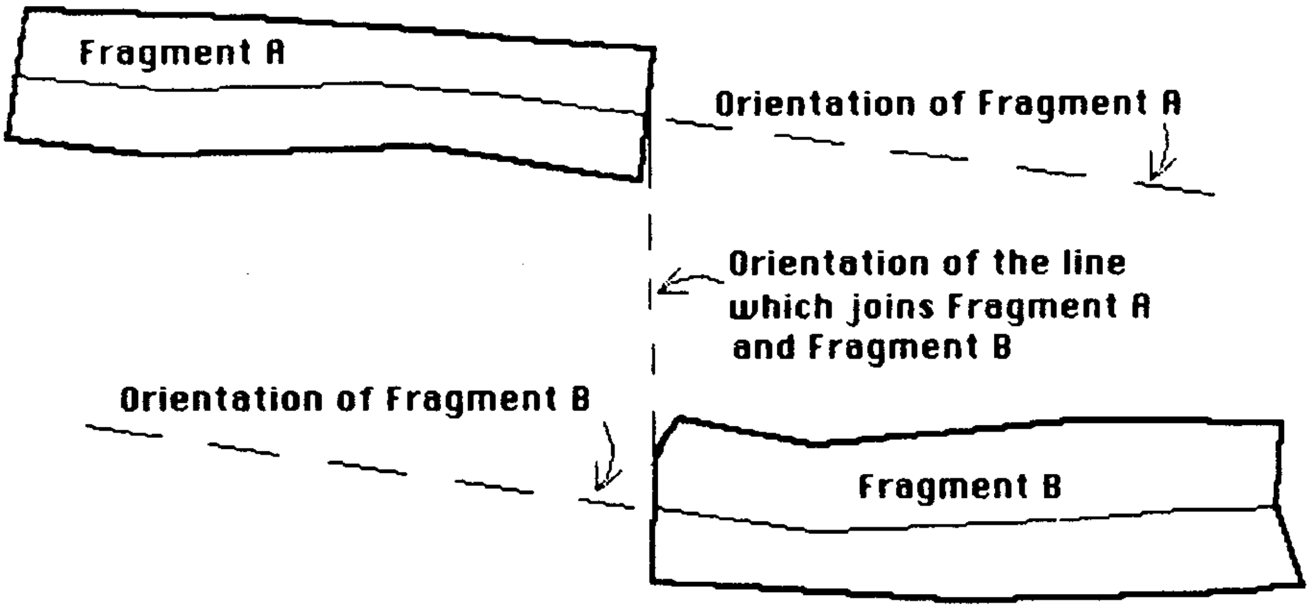
Figure 4-2: Matching Endpoints for Coarse Alignment

used for the linear approximation of the orientation.

Next, a line segment is created which joins the endpoints of the two fragments. Taking the arctangent of these three lines gives a set of angular orientations for the fragments, and the line which connects them. These angles are compared, and both of the fragments are required to be within some tolerance of the connecting line. A satisfactory experimental value for this angular tolerance was determined to be 15 degrees. The region approximations are not directly compared to each other, because while they may be oriented in the same direction, they may be offset from each other in such a manner that a line which connects them would be perpendicular to the regions. This would allow the generation of a composite feature which had a "bump" in it (see Figure 4-2), which is not characteristic of the linear features we are attempting to align. The following section discusses the problem and techniques to accommodate these local discontinuities.

4.4. Tolerance of bumps

Eliminating the possibility of a "bump" in the composite feature presents another problem when the two fragments in question are very close together. The Fourier approximations that are used to smooth the boundaries of the initial regions sometimes overshoots causing the medial axis to extend outside of the actual initial region. This can cause the medial axes to end in nearly the same location, but slightly offset from each other. If this offset is perpendicular to the orientations of the fragments, the above angle test disallows the connection of the fragments. To solve this problem, an exception is made to the angle criteria test if the distance of the jump is smaller than the average width of the two regions (see Figure 4-4). The width of the region is represented by the minor axes of their Fourier approximations.

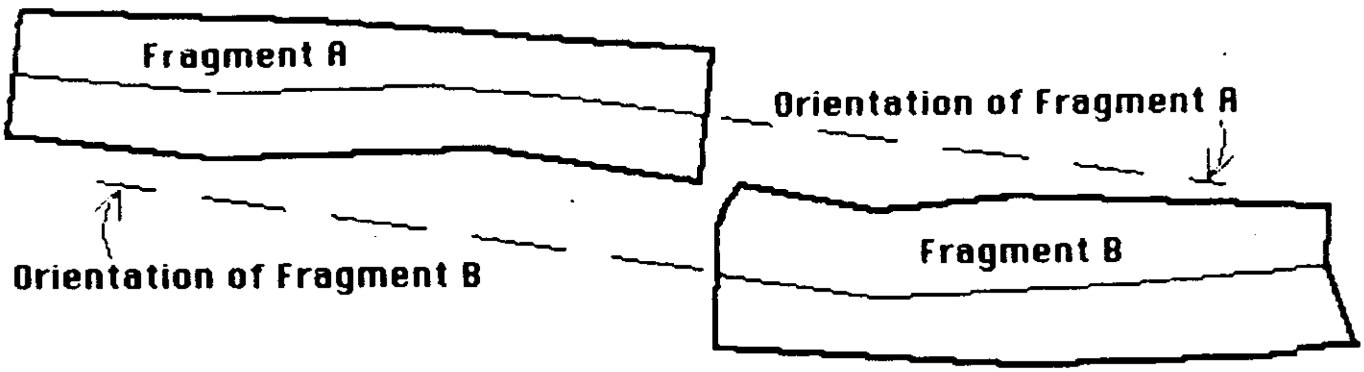


This Figure demonstrates the angle criterion. Pictured are two fragments which have similar orientations, but are offset from one another in such a manner that they do not align properly. The angle criterion requires that the fragments be oriented within some tolerance of the line which connects them. In this example that criterion is obviously not satisfied.

Figure 4-3: An Unacceptable Bump

5. Determining Connection Order of Candidates

If the alignment criteria discussed in Section 4 are met, the fragments are sorted for possible alignment. For each segment in a set of candidates, a link is made to its two neighbors: a neighbor to the first point in the segment, and a neighbor to the last point in the segment. If two or more candidates satisfy the criteria as the same neighbor of the same segment, the candidate is chosen which is closest to the segment in question. Furthermore, a segment may not claim a candidate as one of its neighbors, unless that candidate also selects that segment as (one of) its neighbors.



This Figure Illustrates the case where a "bump" in the alignment is tolerated because the two fragments are very close together. In this case the two fragments must be oriented similarly, but it is not necessary for the line which connects them to also be oriented in a similar direction.

Figure 4-4: A Tolerable Bump

For example, suppose A has two candidates, B and C, which satisfy the alignment criteria, but B is closer. Suppose, also, that both B and C consider A to be a candidate neighbor. Since A will choose B, and B will choose A, a neighbor link is established in both directions. However, when C attempts to choose A as its neighbor, it will fail, because they would not be mutual neighbors (Figure 5-1).

When the neighbor links are established, ALIGN assembles the <latitude,longitude> coordinates of the individual fragments into a vector for the spline algorithm. The points are sorted by the following method: One of the end fragments is chosen first. These can be detected by searching along the neighbor links for a feature which is missing one of its neighbors. This feature's list of points is inserted into the vector beginning with the endpoint which is farthest from the fragment's neighbor. The link to the next neighbor is then followed, and this feature's points are inserted beginning, once again, with the endpoint farthest from the next neighbor. The list is traversed along the neighbor links in

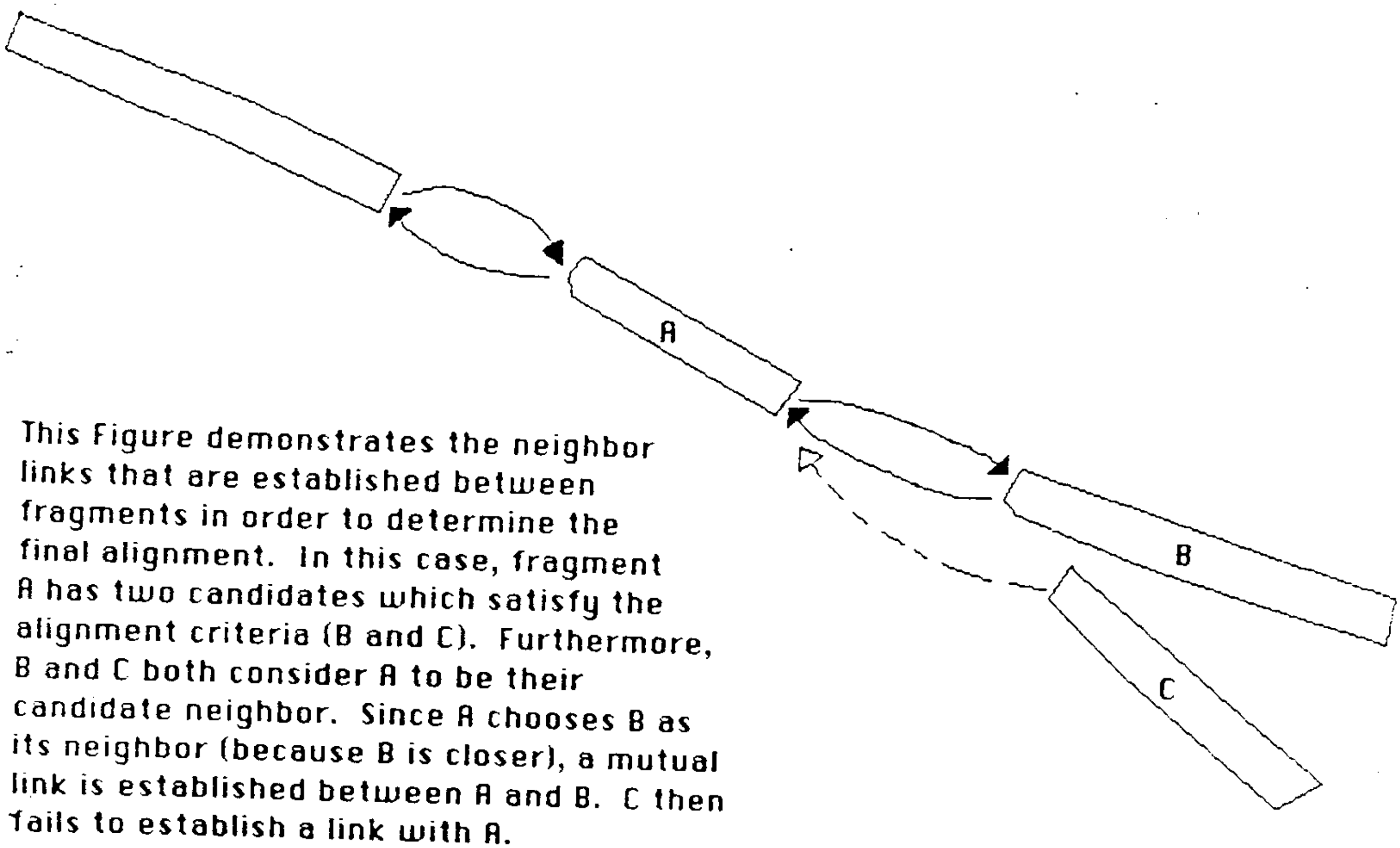


Figure 5-1: Determination of Neighbors

this fashion until the last fragment is encountered. Since this fragment has no next neighbor on which to base the insertion order, the leading endpoint is chosen by determining which is closest to the last point that has been inserted into the vector.

As just described, ALIGN chooses its best hypothesis for the connection of a given feature, and does not even explore the alternative hypotheses. A possible improvement could be incorporated by allowing the generation of alternate results for further exploration. Also, if the chosen hypothesis is later discarded as unsatisfactory, these alternatives could be re-evaluated by the algorithm.

5.1. Region Interpolation Using Spline Approximation

Because there are usually gaps between segments, some interpolation must be performed between aligned segments. A smooth interpolation is desirable, because it would maintain any curvature which the feature possesses. A spline was chosen because it provides a smooth curve while maintaining the constraint that all supplied points are passed through by the spline.

When the vector of points is filled, it is sent to the spline algorithm, which interpolates a smooth (third-order) curve through all of the supplied points. Inter-fragment regions are filled smoothly at the same scale as the original medial axis data. The pointlist which is returned represents the axis of the linear composite feature.

A minor problem was encountered with the spline, due to the jaggedness of the input features. The discontinuities caused by these jumps in the data cause the spline to occasionally oscillate very widely about the axis of the feature. This problem was overcome by smoothing the data points before attempting to generate the spline. The latitude/longitude coordinates of the points are rounded to the nearest tenth of a second. Also, before insertion into the input vector, the points are filtered to remove those which do not represent a shift in both latitude and longitude. This removes redundant points and eliminates the "step" effect which appears after rounding.

5.2. Generation of Composite Features

Each composite feature is reconstructed into a polygon whose medial axis is the spline generated by the previous procedure. The polygon is constructed about the new medial axis using a width taken from the average value of the fragments' minor axes. If necessary, the composite feature could be constructed from the original fragments, rather than from the connected medial axes of the Fourier descriptions. This latter method wasn't chosen in our case, because the smoothing inherent in the former technique is a desirable feature for further processing.

These regions are inserted as standard entries into the SPAM database. A status file is created which reports which fragments were connected, and where the resulting polygon can be found. This allows SPAM to access these features in exactly the same manner as the initial regions created by the region-growing segmentation process.

6. Some Results

ALIGN produces results that, in many situations, allows SPAM to hypothesize the existence of roads, taxiways, and runways--cases where the geometric properties of the individual fragments are not sufficiently valid to warrant such a hypothesis. A typical example of a set of unaligned fragments can be seen in Figure 6-2. This scene contains fifteen individual fragments although they are not all discernible, due to the scale of the drawing.

Figure 6-3 shows the results generated by the program in this example. The four resultant regions shown are composed of eleven of the original fifteen fragments.

Figure 6-4 shows a close up of the alignment of three fragments which are in close physical proximity to each other, and Figure 6-5 shows an alignment which includes a large hypothesized region between the two original fragments that were aligned.

Within the SPAM system, ALIGN is often invoked to attempt to link existing specific runway, road, and taxiway hypotheses, or to generate new hypotheses based on regions whose linear class interpretations were not reliable enough to generate a specific hypothesis. Table 6-1 gives some performance information.

```
Linear Interpretations:
  Runway 6   Taxiway 30   Road 234   Unclassified 44

Alignments:
  Attempted 285
  Adjoined  129
  Average Number Regions Aligned  2.3
  Minimum Number Regions Aligned  2
  Maximum Number Regions Aligned  5
```

Figure 6-1: SPAM Statistics for ALIGN

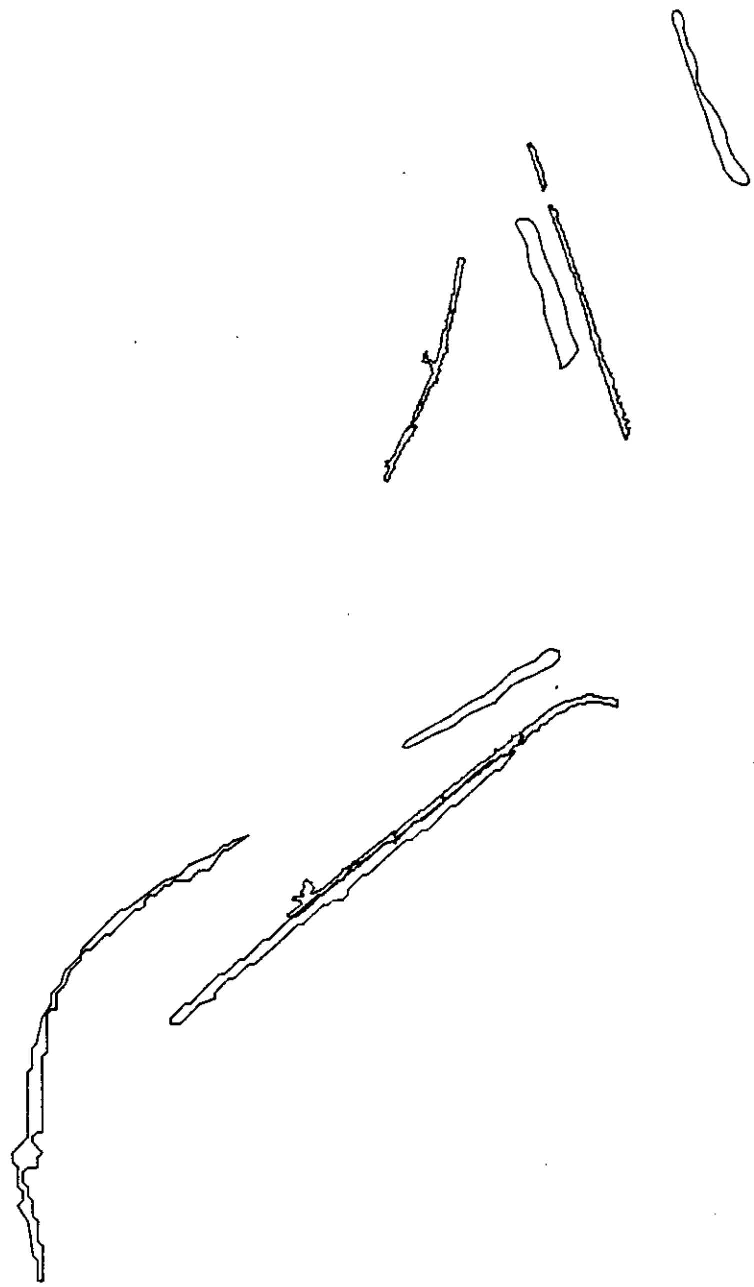


Figure 6-2: A typical scene before alignment.

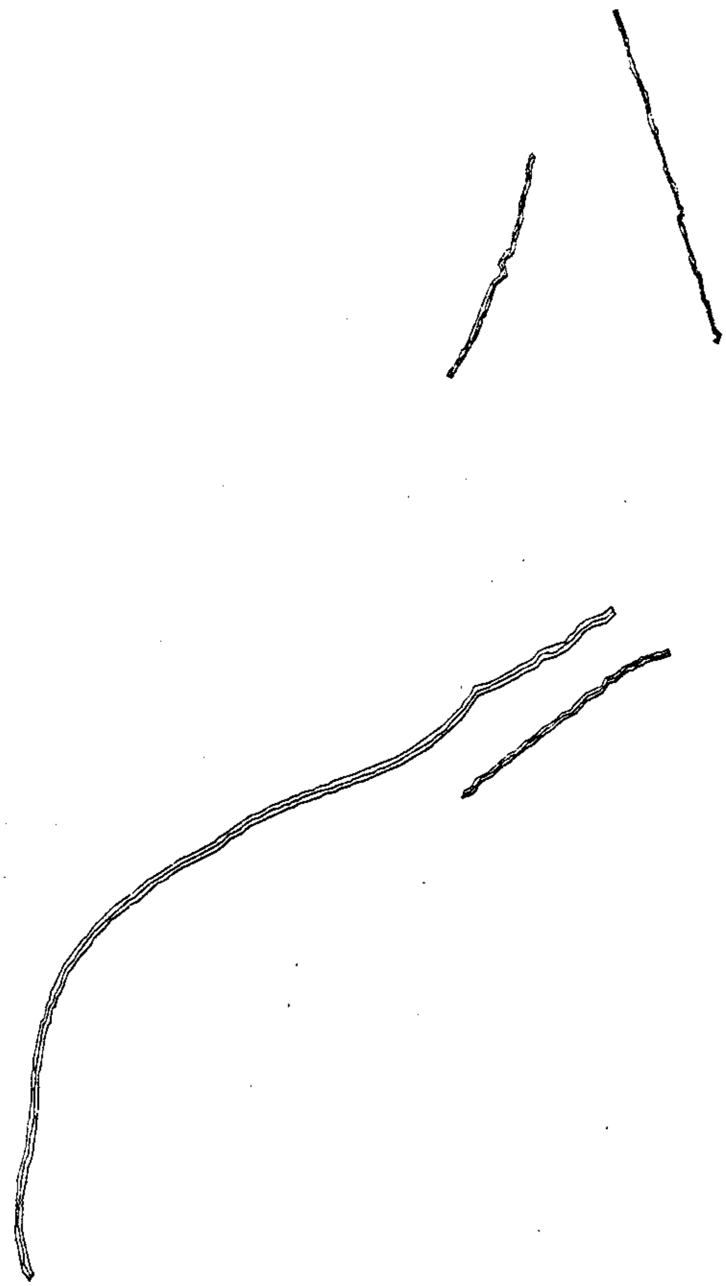


Figure 6-3: The results generated by ALIGN for the scene in Figure 6-2.

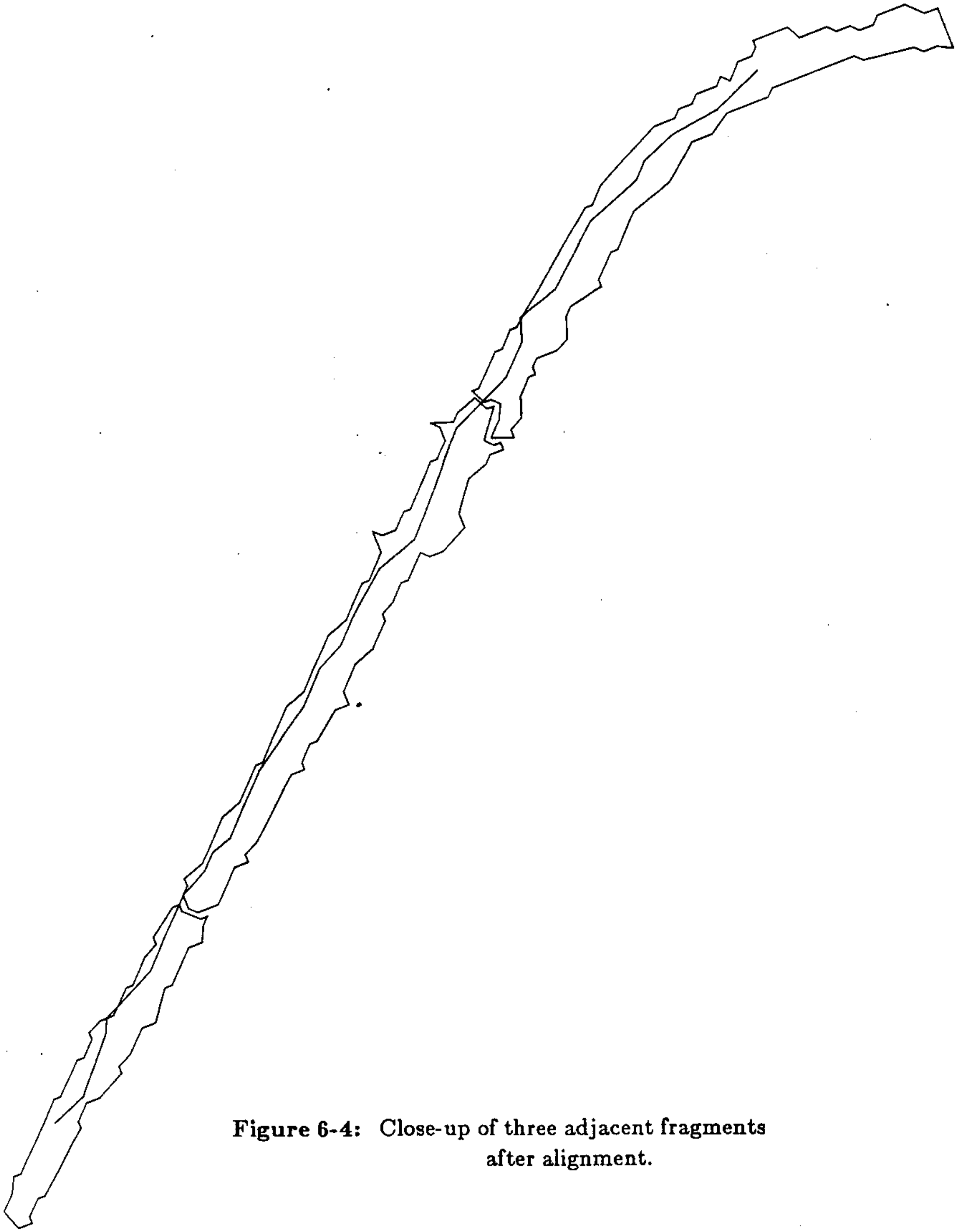


Figure 6-4: Close-up of three adjacent fragments after alignment.

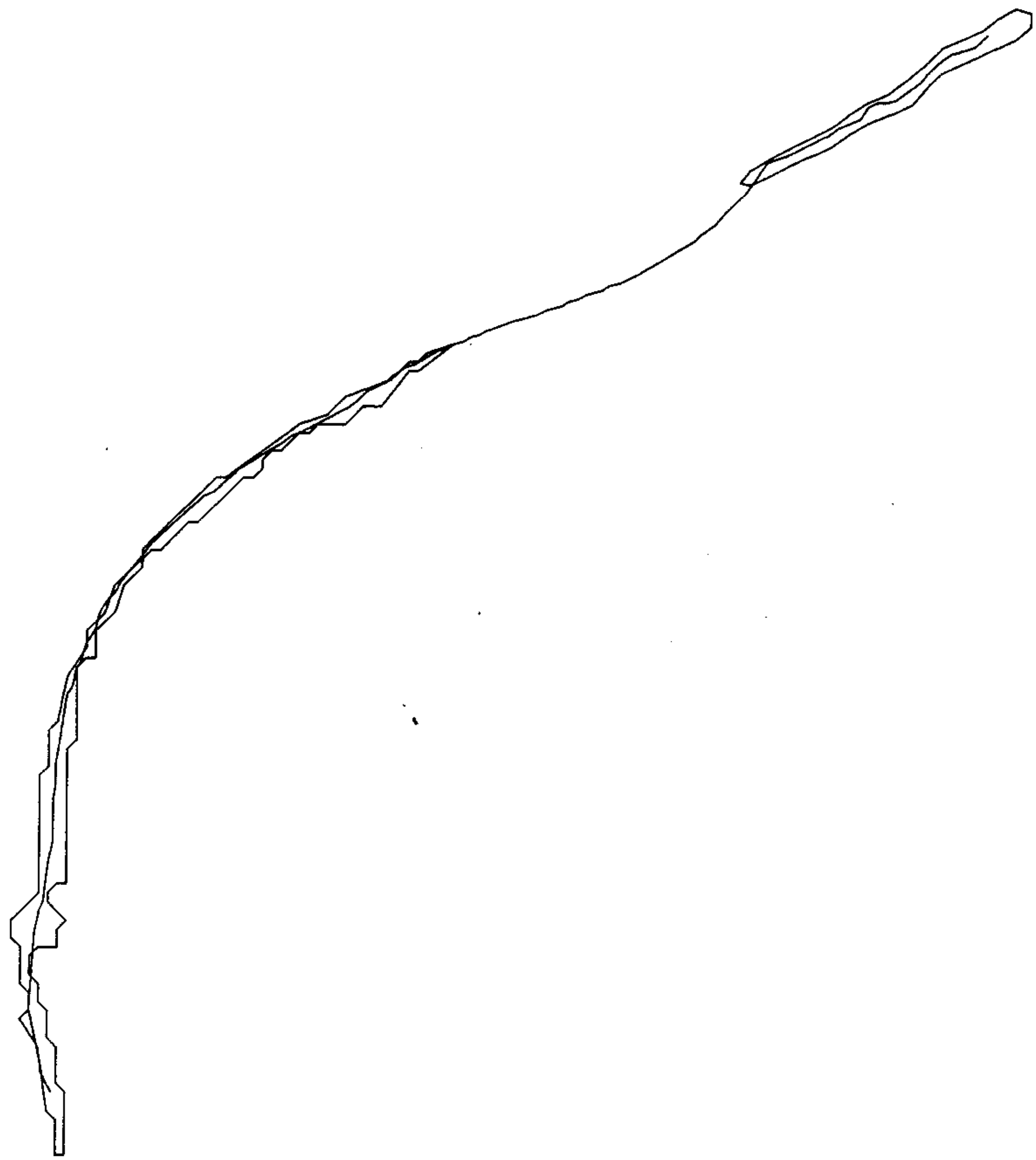


Figure 6-5: Two aligned fragments with an inter-fragment zone.

7. Future Work

The algorithms used by ALIGN can be tailored to the application by adjusting the various tolerance factors. ALIGN shows some weakness when it operates on features that exhibit little linearity because the length of the medial axis decreases, and its orientation becomes less certain. This same effect occurs on very small features, which may lose significant shape information when the smoothing algorithm is performed.

The parameters which determine our criteria of tolerance are based on heuristics of the geometric and structural constraints of our specific problem. They were empirically selected by adjusting them and monitoring how they affected performance of the alignment on many sets of examples. They offer flexibility to operationalize a wide range of domain constraints or search strategies. For instance, an application of this method might begin with a set of very strict selection and alignment constraints, iteratively adjusting the regime until satisfactory results are produced. Further, by altering the types of alignments, this method could be specialized to look for corners or other cultural features based on topological relationships between fragments.

As well, there are some possible improvements or extensions to the algorithm which have not been completely explored:

- A quality-of-fit factor might be generated by the spline algorithm which could be used to assess the value of the hypothesis made. If this quality is judged unsatisfactory, the program might backtrack and explore alternate hypotheses.
- If there are multiple alternatives which satisfy the alignment criteria, the program might hypothesize more high-level cultural feature such as intersections or forks in roads.
- Similarly, a comparison might be made of the actual distance versus the hypothesized distance (along the spline) between each pair of fragmented features. A significant difference between these two values might indicate an unsatisfactory alignment attempt.
- The method of locus expansion in the candidate selection algorithm might be improved to provide a better representation of the length and width of the features, and might be more selective about expansion of the locus.

- Smoothing of the feature might be eliminated on those features which are very small. If they already have a limited number of points to represent them, it is unwise to remove those points.

8. Conclusion

This paper describes a method to select and align segmentation fragments into a single region so that high level interpretation can proceed using more reliable and meaningful segmentation primatives. In its current state ALIGN produces useful results that, in many situations, allows SPAM to hypothesize the existence of roads, taxiways, and runways-- cases where the geometric properties of the individual fragments are not sufficiently valid to warrant such a hypothesis. While local linking algorithms may provide adequate analysis at a very low level, they leave much to be desired when the analysis moves toward a more global assessment of a scene. As we attempt to build image understanding systems that are capable of generating descriptions of cultural features in aerial imagery, programs such as ALIGN are necessary in the intermediate level to provide cues that may not be directly available from low-level segmentation algorithms.

9. Acknowledgment

The authors wish to acknowledge comments and helpful suggestions on an earlier draft of this paper by Steve Shafer, and by an anonymous referee for CVPR85.

10. Bibliography

1. McKeown, D.M., and McDermott, J., "Toward Expert Systems for Photo Interpretation," *IEEE Trends and Applications '83*, May 1983, pp. 33-39.

2. McKeown, D.M., Harvey, W.A. and McDermott, J., "Rule Based Interpretation of Aerial Imagery," *Proceedings of IEEE Workshop on Principles of Knowledge-Based Systems,* Dec 1984, pp. 145-157.

3. McKeown, D.M., Denlinger, J.L., "Map-Guided Feature Extraction from Aerial Imagery," *Proceedings of Second IEEE Computer Society Workshop on Computer Vision: Representation and Control,* May 1984, Also available as Technical Report CMU-CS-84-117

4. Sloan, K., and Ballard, D., "Experience with the Generalized Hough Transform," *Proceedings: DARPA Image Understanding Workshop,* April 1980, pp. 150-156.

5. Bookstein, F.L., "The Line-Skeleton," *Computer Graphics and Image Processing,* Vol. 11, 1979, pp. 123-127.

6. Ballard, D.H., and Brown, C.M., *Computer Vision,* Prentice Hall, Englewood Cliffs, NJ, 1982.