Reconstruction of Fragmented Objects

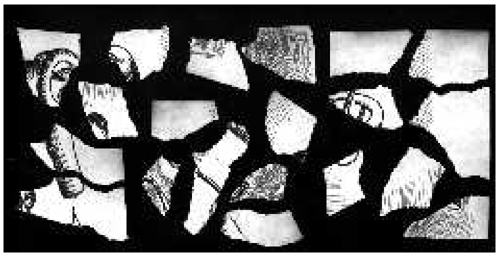
Helena Cristina da Gama Leitão[1]
Jorge Stolfi[2]

Institute of Computing
Universidade Estadual de Campinas
Caixa Postal 6065 - 13081-970 Campinas, SP, Brasil
{hcgl|stolfi}@dcc.unicamp.br

Abstract. We consider the problem of automatic reconstruction of fragmented objects, such as ancient vessels and documents, mural paintings, fossils, etc.. Our approach combines an original filtering technique with multiscale curvature analysis and biosequence indexing algorithms.

1. Introduction and motivation

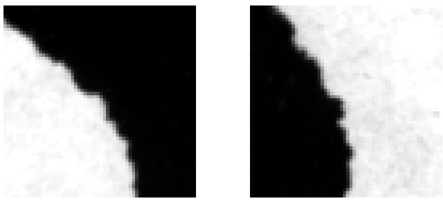
The topic of this work is the reconstruction of unknown objects that have been broken or torn into a large number of irregular fragments, such as the ones in (1).



(1)

This problem occurs in several contexts, including the reconstruction of ancient manuscripts, ceramic articles, clay tablets, mural paintings, fossils, and so on [1].

Once the fragment images have been acquired, the most time-consuming part of the reconstruction is the identification of pairs of fragments that were adjacent in the original objects. Fortunately, in many of the applications mentioned above, such pairs can be identified, with high confidence, by the congruence of their outlines at sub-millimeter scale. See figure 2.



(2)

This problem has aspects in common with the recognition of objects from their outlines, a classical problem in computer vision [3]. However, most works in this area assume that the the outlines are to be matched against a small set of fixed templates, so that their goal is to reduce the cost of comparing an outline against a template [5].

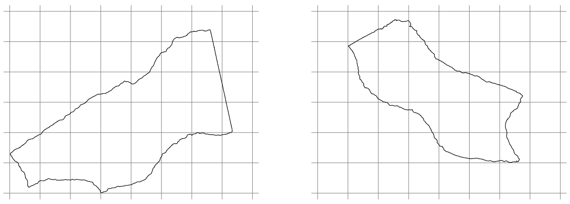
In the applications that interest us, however, the "templates" are the fragment outlines themselves. Since the number of fragments n may be quite large—tens of thousands—it is not feasible to compare every outline against every other. Therfore, our main goal is to to decrease the *asymptotic* cost of of locating congruent fragments, from $\Omega(n^2)$ to something closer to $O(n)$.

2. Details of our approach

Acquisition of fragment outlines: In the case of flat fragments, such as paper or mural paintings, we can obtain the outlines by digitizing each fragment with a flatbed document scanner (either directly or through photographs), and then applying a suitable edge detection algorithm. In the case of ceramic or fossil fragments, the fracture lines are three-dimensional curves, and their extraction will require stereoscopic vision techniques.

In any case, it is important that the contours be determined with sub-millimeter accuracy, since the reliable identification of adjacent fragments depends on details of that magnitude. In particular, each outline must follow the fracture line on a specific side of the object, in spite of color variations, turned edges, etc..

Figure 3 shows the outlines of two fragments from figure 1. Each grid square mesures 50×50 pixels. One can see two stretches of about ≈ 200 pixels, one on each contour, that are congruent within a couple of pixels.



(3)

Outline filtering: The extracted outlines are usually contaminated by "noise" due to meterial erosion and digitization errors. Therfore, before comparing the outlines,

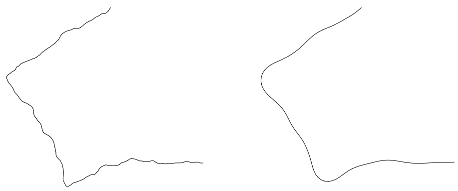
[2] CNPq Research Grant 301016/92-5.
[1] Capes PICD(UFF) Grant

it is necessary to filter away their high-frequency components, which are the ones most affected by such noise.

Filtering a curve is harder than filtering a time series, since the natural parameter for describing the curve—its length—may shrink substantially and non-uniformly as result of filtering. This shrinking changes the wavelength of the curve components, which invalidates the filtering. Ideally, the curve should be parametrized by its length *after* filtering, and the filtering should be based on this parametrization.

In order to break this vicious circle, we use the following iterative process. Let p be the extracted outline, presented as a closed polygonal line with vertices $p_0, \ldots, p_{n-1}$. Initially, we assign to each point p_i a parameter value t_i, equal to the length of the polygonal p from p_0 to p_i. then, at each iteration, we compute a continuous curve $u(t)$, that approximates the points p_i at times t_i. A low-pass filter is applied to $u(t)$, resulting in a smoothed curve $v(t)$. Each t_i is then redefined as the length of $v(t)$ from $t = t_0$ to $t = t_i$. In the next iteration, the curve $u(t)$ is recomputed, using the original points p_i with the revised parameters t_i; and so on.

This process usually converges after a few iterations. Figure (4) shows a detail of the first outline from figure (3), before and after filtering.



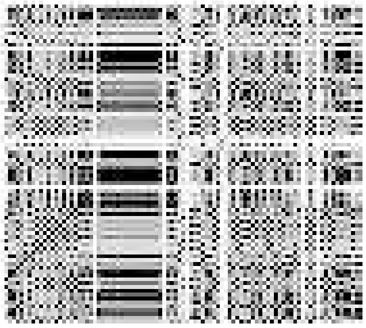
(4)

Finding congruent stretches: The next step is to identify pairs (C, C') of outlines that have long and approximately congruent stretches.

In order to make this task independent of fragment rotation and translation, we use the standard technique of representing each outline by the graph of its local curvature (and torsion, in the case of three-dimensional curves) as a function of arc length [5].

When we discretize the curvature into a finite set of "symbols" (curvature values), two curves with similar shapes will be mapped into similar strings of symbols. Thus we are left with the problem of finding all pairs of similar sub-chains (modulo inversion and negation) in a given set of strings. This is a standard problem in computational geometry, for which many efficient algorithms are known [2].

Image (5), inspired by the FAST biosequence matching technique, shows the comparison of curvature graphs for the two outlines in figure (3). Each column i in the image corresponds to an extremal curvature value (maximum or minimum) a_i in the first outline, and each line j ocorresponds to an extremum b_j in the other. The lightness of pixel (i, j) is proportional to the difference $|a_i - (-b_j)|$.



(5)

The dark diagonal line, in the lower left part of the image, reveals a long sequence of curvature maxima and minima, in one of the outlines, that occurs—reversed and negated— on the other outline. Thsrevealsi coincidence is a strong indication that the two fragments were originally adjacent.

Multiscale techniques: The many dark spots in figure (5) make it difficult to distingush algorithmically the true shape congruences from random coincidences.

Part of the difficulty comes from the fact that two curves with very different shapes may have quite similar curvature graphs. To get around this problem, we plan to perform the identification at *multiple scales* of resolution [4]. Instead of comparing only the curvature graphs of the original outlines, filtered with cutoff frequency w, we compare also the outlines filtered with cutoffs $w/2$, $w/4$, ... The justification being that two curves with very different shapes will have very different curvature graphs in at least one of these scales.

References

[1] A. D. Kalvin et al., *Using visualization in the archaeological excavations of a pre-Inca temple in Peru.* IBM Res. Rep. RC 20518 (1996).

[2] J. Meidanis and J. Setubal, *An Introduction to Computational Molecular Biology.* PWS Publishing (1997).

[3] A. R. Pope, *Model-based object recognition: A survey of recent research.* Tech. Rep. 94-04, CS Dept., Univ. of British Columbia (1994).

[4] F. Mokhtarian and A. K. Mackworth, *A theory of multiscale, curvature-based shape representation for planar curves.* IEEE Trans. on Pattern Analysis and Mach. Intell. **14** (8), 789–805 (1992).

[5] H. J. Wolfson, *On curve matching.* IEEE Trans. on Pattern Analysis and Mach. Intell. **12** (5), 483–489 (1990).

[6] W. R. Pearson and D. J. Lipman, *Improved tools for biological sequence comparation.* Biochemistry - Proc. Natl. Acad. Sci. **85**, 2444–2448 (1988).