Algorithmic Reconstruction of Broken Fragments

Steven ELIUK – Pierre BOULANGER

University of Alberta Computing Science, Edmonton, AB
{eliuk; pierreb}@cs.ualberta.ca

Abstract

Reconstruction of broken fragments is an extremely difficult task, mainly due to the high probability of missing pieces and the likelihood of erosion. This process is tedious and requires many hours to complete even for the experienced archaeologist. This paper describes some of the most successful fragment reconstruction algorithms and presents a new technique for reassembly. The technique relies on a simple reassembly algorithm which uses the Pearson Correlation Coefficient (Pearson 1896) to calculate the linear relation between two curvatures. The Pearson Coefficient calculation does not suffer from some of the apparent pitfalls of other algorithms such as being unaffected by weathering of objects and incomplete fragment sets. Using this new application of the Pearson Coefficient we can help the experienced archaeologist reconstruct broken fragments.

Keywords

Automated reconstruction, curvature correlation, reassembly

1. Introduction

The importance of computers in archeology is becoming more apparent due to the increased demand for 3D digital scans of archaeological sites and reconstruction of fragments. The significance lies in the fact that one can work with a 3D digital scan more easily on a computer. The scale of the site or object can be projected on a monitor providing easy navigation or investigation. Historical artifacts can be extremely heavy and fragile and therefore very susceptible to damage from movement or inspection. Computers provide the ability for the object to be handled in a less intrusive manner,using simulation to create a virtual object. A virtual object is a 3D representation of the original artifact that can be manipulated and modified in a virtual environment. Now investigators can work with a digital representation of the object, with the original being safe from environmental damage. There is even the ability to physically print these 3D virtual objects with 3D printers, in the event that a physical model is needed. Through 3D representation of archaeological sites and objects these techniques have led to the ability to reconstruct damaged artifacts or even the reassembly of artifacts which have been broken for many years. For instance, museums have large collections of pottery fragments. It would be possible to reassemble these fragments to generate a complete or almost complete model. Prior assembly algorithms have relied on knowledge of the object before assembly. This is not always possible, as in many instances the the final shape of the object is unknown. Earlier works, such as Virtual Archaeologist (Papaioannou *et al.* 2001) have been influential in in the development of modern day techniques. They provide a successful reassembly algorithm that requires little user interaction and no knowledge of the object prior to reassembly. Through Virtual Archaeologist, other techniques have been developed and will be discussed in more detail later in this paper.

In this paper, we describe the problem (Section 2), give a rather complete review of present day algorithms (Section 3), then define a new technique (Section 4) for 3D reassembly that relies more on the curvature of the fragment rather then the facet of fracture. We provide an example of the proposed algorithm and provide evidence to demonstrate the algorithm with a real world reassembly of a broken cup. Lastly, we provide information on our current research and considerations for future development.

2. Problem description

Fig. 1 shows 3D scans of pottery fragments (*Fig. 1a*), an individual fragment (*Fig. 1b*), curvature projected on a 2D plane (*Fig. 1c*) and the location of match on an original model (*Fig. 1d*).
3D reassembly is a 3D puzzle which requires,

– scanning of fragments (*Fig. 1a*)
– 3D mesh generation (*Fig. 1b*)
– smoothing and feature extraction (*Fig. 1c*)
– probabilistic pairwise matching
– reassembly of object (*Fig. 1d*)

3D fragment puzzles have a number of difficulties,
- incomplete scans
- number of pieces is unknown
- incomplete data (missing pieces)
- extra data (fragments that belong to other models)
- incomplete surfaces and non-uniform erosion
- orientation unknown/transformations needed

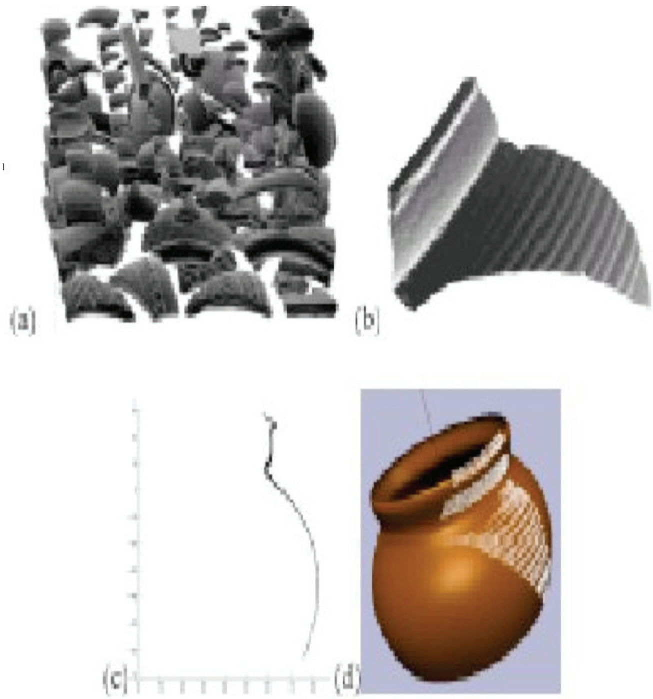


Fig. 1. CARIA: Computer Assisted Reconstruction In Archeology (Aytul et al. *2003).*

The 3D puzzle is a difficult task for an experienced archaeologist working with the actual object. It can take hundreds of hours and is often unsuccessful. Through the aid of computers and a small amount of user interaction its possible to give a good virtual 3D assembly of an object in a reasonable amount of time. This process does not damage the original object and provides a means of careful investigation into reconstruction of the fragment.

3. Previous work

There has been a wealth of knowledge produced in the early 21st century regarding 2D/3D assembly. In this section we describe some of the latest developments in the field and give some of the background methods used in these works. We will briefly describe some the significant pieces in the works. This is by no means a complete literature review on fragment reconstruction, but provides a good look at current day algorithms.

3.1. Virtual Archaeologist: Assembling the past (Papaioannou *et al.* 2001).

Virtual Archaeologist provides a meaningful addition to the literature surrounding 3D reassembly in the 21st century. The paper applies a semi-automatic algorithmic technique to the 3D reconstruction of fragments. The technique relies on a three stage process, the first being mesh segmentation. Papaioannou attempts to restrict the search of possible matches solely to those faces which appear "interesting". They define "interesting" to be those faces which appear coarse, making the assumption that fractured faces are more jagged and rough then their counterparts- human generated "smooth" faces. The second state involves estimating the relative pose for the fragments where matching error is minimized. In the third state, is where full reconstruction is completed. The algorithm selects fragment combinations that minimize global reconstruction error. We will mainly focus on the first and second stage, mesh segmentation and estimating the relative pose for fragment comparison.

The mesh segmentation is completed by a region-growing algorithm (Papaioannou *et al.* 2001). The process begins with a random polygon. Neighboring polygons are classified to the same face if the normal for the neighbor does not deviate from the average face normal by more than a predefined threshold. In the event the normal deviates more than the predefined threshold a new face is formed. As noted in the article it is possible for small faces/regions to be created within larger regions. This is not desirable, so the algorithm merges these regions to eliminate these small artifacts to produce a more reasonable face for comparison. A region is classified as small if it encompasses less than 5% of the entire 3D mesh of the object. The process continues for all objects with each region having an average facet normal, as depicted in *Fig. 2.* center.

These average normals are used to define the transformations needed to compare two 3D fragment faces/regions. This is represented in *Fig. 3.* Two different facets are originally aligned based purely on their respective average normals. The first object is then able to perform a full rotation around the axis of alignment(ρ_1), deviated from the axis(φ_1,θ_1) by up to 10°. The first object is also able to slide along the broken facet (x_1,y_1). Using the average normal the transformation needed to approximately align the two facets is trivial. Comparison and fragment merging is

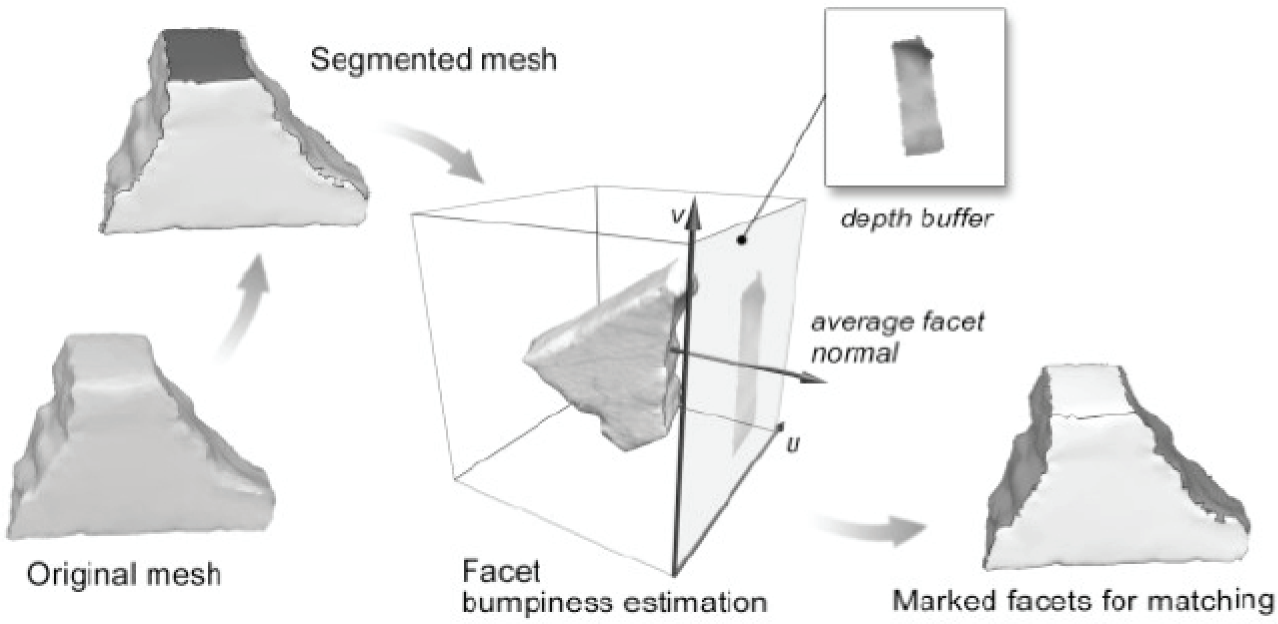


Fig. 2. Virtual Archaeologist: Assembling the past.

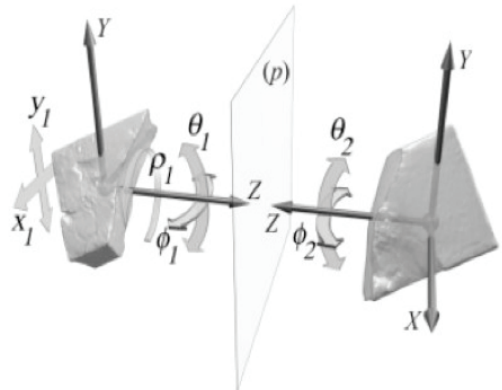


Fig. 3. Virtual Archaeologist: Assembling the past. Alignment of two facets for comparison (Papaioannou et al. *2001).*

then completed in the final stage where the algorithm minimizes the global reconstruction error.

The strategy described by Papaioannou provides a good technique for mesh segmentation and reconstruction. The technique has difficulties when attempting to reconstruct similar shapes, such as a broken tablet where many fragments have similar flat faces. The underlying assumption of this method is that the fracture faces are nearly planar and they match completely as seen in (Huang *et al.* 2006). This technique does not take into account any of the image based textures that could be used for matching. However, the paper is influential in the development of other modern techniques and provides techniques for mesh segmentation and alignment of fragments for comparison.

3.2. Reassembling Fractured Objects by Geometric Matching (Huang *et al.* 2006)

The paper provides improvements over previous techniques by being able to handle arbitrary shaped fractures and partial matches. These abilities make the algorithm less dependent on the segmentation step. The algorithm uses mesh segmentation, alignment, pairwise matching and global multi-piece matching. The pairwise matching provides a quality rating for each match. From the information obtained from the pairwise matching process, the global multi-piece matching is performed. Huang uses integral invariants, found originally in (Manay *et al.* 2006), to help with the segmentation process. The paper defines integral invariants as integrating spatial functions over moving domains centered at surface points (Huang *et al.* 2006). Huang introduces integral invariants (Huang *et al.* 2006), more details and further references can be found in Manay *et al.* 2006. *Fig. 4* represents one of the experiments used to test the algorithm. All of the test objects have rather large solid pieces with large fracture faces. As shown in *Fig. 4*, the results are impressive, however tend not to focus on objects where the fractured face is relatively small like that of pottery or non-solid objects. This would require more emphasis on the curvature of the fracture because the unique geometric features in the

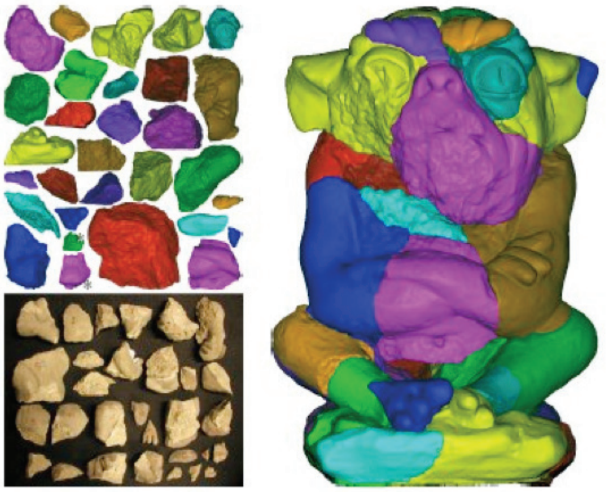


Fig. 4. Reassembling Fractured Object by Geometric a Matching. Reassembling a gargoyle statue. Photo of fragments (bottom left), 3D model (top left), final assembly (right). (Huang et al. *2006.)*

fractured face would not be as abundant. The use of integral invariants seem very successful. As with the previous technique, the process does not take into account any image based textures. Nonetheless, the technique is very successful for reassembly of objects with rather large fracture faces.

3.3. Evaluation of obstetric gestures: an approach based on the curvature of 3d positions (Moreau *et al.* 2006)

Moreau evaluates curvatures correlation using the Pearson correlation coefficient (Pearson 1896), to calculate the linear relation between two obstetric curvatures(1).

$$r_{pr} = \frac{\sum_{i=1}^{i=n} (A_i - \overline{A_m})(B_i - \overline{B_m})}{\sqrt{\sum_{i=1}^{i=n} (A_i - \overline{A_m})^2 \sum_{i=1}^{i=n} (B_i - \overline{B_m})^2}} \tag{1}$$

Equation (1) represents the Pearson Correlation Coefficient, denoted r_{pr}. The A_i is the ith component of the first curvature vector and likewise, B_i represents the ith component of the second curvature. A_m and B_m represent the average of the components of the first and second curvatures, respectively. The technique provides a single value that represents the linear relation between two curves. If the curves are similar they will have a greater correlation coefficient.

The overview of prior techniques all provide important insight into the development of the method presented in this paper.

4. Tools and methods

The curvature of a break contains much of the information needed to make good matches in limited facet breakage. The previous methods rely on unique features in the facet of breakage to make good matches. As noted this face is sometimes small and possibly weathered, which in turn creates difficulties for current day algorithms. The goal of this paper is to provide a good matching algorithm for pottery type matches where the facet of breakage is small. The algorithm relies on four stages. These stages are outlined in *Figs. 5–7*. The first is the segmentation of the artifact into interesting curves. We define "interesting" to be curves where the relative change between points along the curve is coarse. Most human generated curves are smooth and consistent, with few abrupt changes. The second stage relies on alignment of the curves, mainly based upon Papaioannou's alignment algorithm (Papaioannou *et al.* 2001), where the average normal is used for initial alignment. After the initial alignment, a series of small euclidean transformations are performed in both rotation and translation in order to to minimize the Pearson Correlation Coefficient between the two curves in question. The third stage involves Moreau's

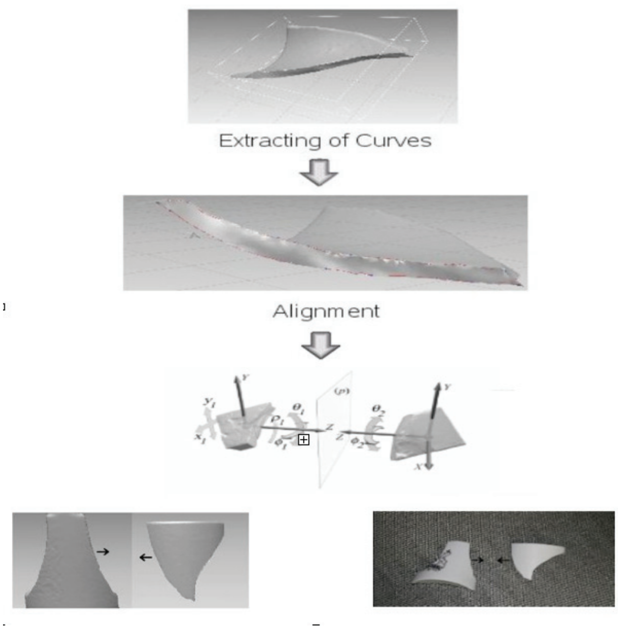


Fig. 5. Stage 1 - Extraction of curves Stage 2 – Alignment (Papaioannou et al. *2001).*

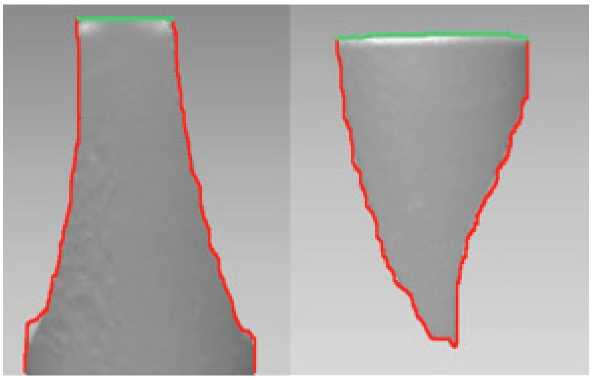


Fig. 6. Green curvatures represent "man made" classification Red curvatures represent "fracture point" classification.

technique for curvature comparison (Moreau *et al.* 2006) in order to evaluate the similarity of two curves. The fourth stage involves pairwise assembly of fragments. The new joined piece is added to the pool of fragments. The process then begins again, extracting curves, alignment, etc until all pieces are matched to a relative certainty. Pieces that can not be matched to predefined certainty can then be investigated further by an experienced archaeologist.

I. Stage 1: Segmentation to find interesting curves

The segmentation of an object into interesting curves involves analysis of the fragment and its edges. Once a perimeter is determined for the object, we segment this perimeter based on abrupt changes between the previous perimeter point and the current point. We do this incrementally, starting initially at an arbitrary location on the perimeter. If the angle between the current point and the previous point is greater then a predefined threshold, a new curve is then produced. Once the entire perimeter has been segmented into curves, interesting curves are determined by the change in coarseness of the curve. If the coarseness is greater than a predefined threshold, which may be domain specific, the curve will be classified as "man made" or a "fracture point". There is a number of ways this can be accomplished. One can count the

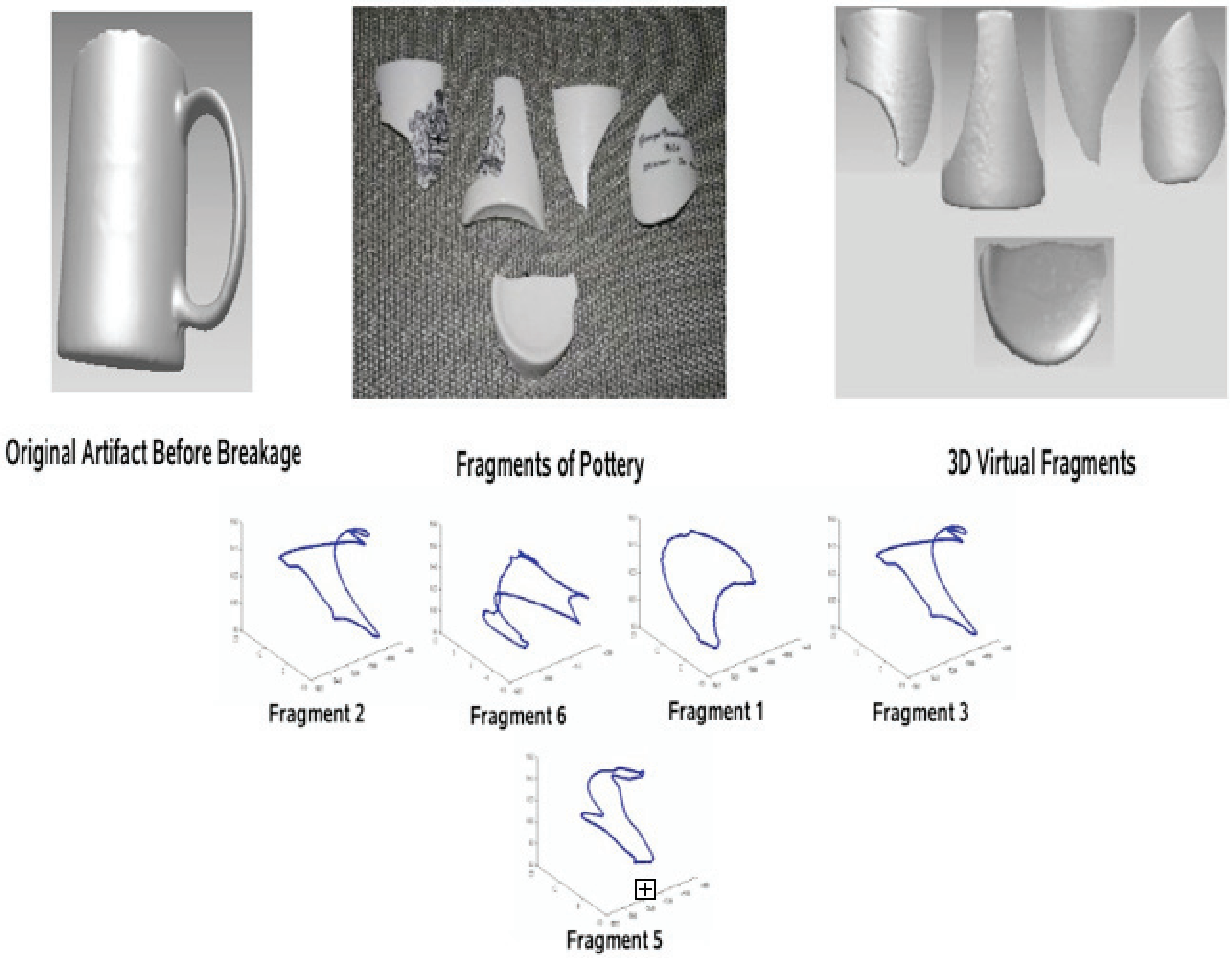


Fig. 7. Original 3d object (left), fragments after breakage (center), 3D virtual fragments (right), curves generated from 3D virtual fragments (bottom) note: the curvatures (bottom) could represent a transformation of the objects original pose when curvature was generated.

number of abrupt changes that represents coarseness. If the count exceeds a predetermined threshold the curve is classified as a fracture point. If the count does not exceed the predetermined threshold the curve is classified as man made. Another technique would be to fit a polynomial of a predetermined degree to the curve, if the squared error between the fitted polynomial and the actual curve is greater than a predetermined threshold the curve would be classified as a fracture point. If the squared error is below the threshold, the curve is relatively smooth and most likely man made. *Fig. 6* represents these classifications, red being the fracture points and green being the man made curves. Each fragment contains a list of interesting curves. These curves are then used for comparison purposes in the later stages.

II. Stage 2: Alignment of curves

Fig. 5 represents the alignment of curves based upon (Papaioannou *et al.* 2001). Initially an average normal is computed for each curve. The alignment algorithm then does a trivial alignment based solely on the average normal. A quick refresher of the alignment scheme is needed. As noted earlier referring to

Figs. 3 and *5*, the first object is able to perform a full rotation around the axis of alignment(ρ_1), deviated from the axis(φ_1,θ_1) by up to 10˚. The first object is also able to slide along the broken facet (x_1,y_1). This is accomplished by a transformation matrix acting on all the points of a curve. As the number of transformations increase so does the chance of aligning two curves. This means an incremental approach can be adopted. Initially starting with a low number of transformations. If there are many pieces that can not be classified the algorithm then increases the bounds on available rotations and translations. One could also increase the precision of these transformations. For instance, if we initially used rotations with 1° increments we could use 0.5° increments attempting to get a better alignment. The less transformations that are considered the quicker the algorithm performs and the better the incremental approach seems to work.

This entire process can almost be eliminated because curvature comparison, stage 3, does not require alignment to compare two curves. Simple translation along a curve is needed to obtain the best correlation coefficient. This reduces the complexity of the algorithm greatly and in turn reduces the runtime.

III. Stage 3: Curvature comparison

The virtual representation of the objects provides points in Euclidean space. We rely on Moreau's technique to express the data in respect to arc length and not simply point data. A review of Moreau's technique to express data in respect to Arc length is below, see Moreau *et al.* 2006 for a complete description.

The arc length *Si*, seen Equations 3 and 4, is defined as the Euclidean distance between two consecutive points where *i* is from 1 to *n* – 1, where *n* is the number of sampled data points defining the original curve.

The distance in each direction is defined by:

$$\begin{aligned}\Delta(x_i) &= (x_{i+1} - x_i);\\ \Delta(y_i) &= (y_{i+1} - y_i);\\ \Delta(z_i) &= (z_{i+1} - z_i);\end{aligned} \qquad (2)$$

Where xi, yi, zi are the *i*th components of the fracture curve.

$$s_i = \sqrt{\Delta(x_i)^2 + \Delta(y_i)^2 + \Delta(z_i)^2} \qquad (3)$$

The cumulated arc length *l* is defined as:

$$l = \begin{bmatrix} 0 & s_1 & s_1 + s_2 & \dots & \sum_{i=2}^{i=n} s_{i-1} \end{bmatrix}^T \qquad (4)$$

The curvature *k* is now extracted by taking the norm of the second derivative according to the accumulated arc length *l* seen in Equation (4).

$$\kappa(l) = \left\| f''(l) \right\| \qquad (5)$$

This technique to extract the curvature produces peaks in *k* where there are significant changes in the curvature. This essentially highlights the features in the curve where strong correlation between curves can be generated.

There is also the unique problem of non-uniform sampling of the 3D fragment while scanning. When comparing curvatures in the later stages of the algorithm, we assume the sampling is uniform in the sense that there are equal number of points for any of the curves over the same distance. This method guarantees a one-to-one comparison over a given distance. We accomplish this uniform distribution of points along a curve by making sure the Euclidean distance between any two points along a curve is at a predefined value. If the distance is below, a point is inserted into the curve. This involves a small amount of preprocessing.

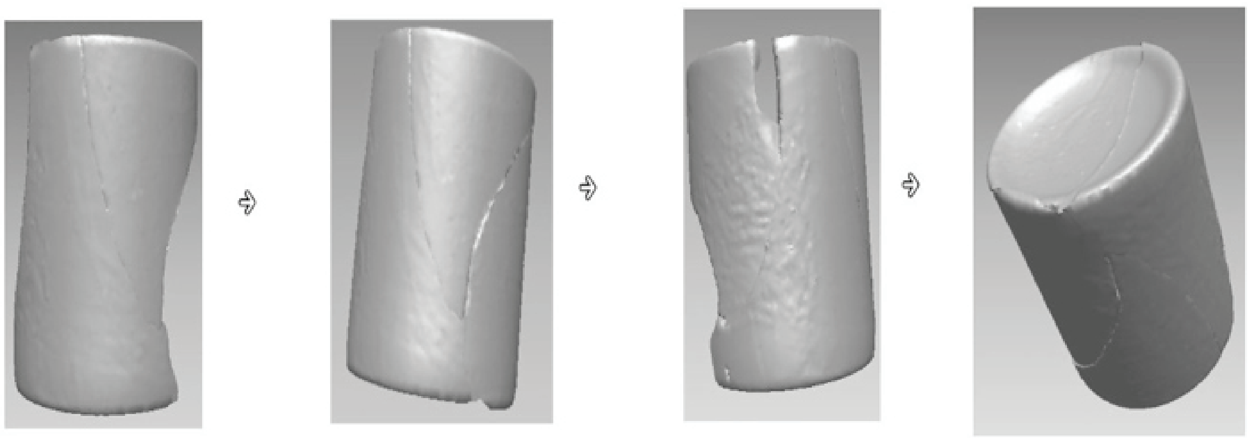
Using the Pearson Coefficient, Equation (1) as a guideline for determining the level of correlation between two curves, the algorithm chooses the transformation that results in the highest correlation curvature. As mentioned earlier, this is accomplished by normalizing the curves. This is needed because each curve must have the same number of points over a given distance, a one-to-one correspondence is needed in order to use the Pearson Coefficient. The best correlation coefficient between the two curvatures being compared is then stored. This value will be the determining factor when deciding if the curvature in question has a match among the available curves. If the fragment cannot be matched at this stage there is still the potential that a suitable match pair could be produced by the joining of two pieces. When two pieces are combined a new curve can be produced which may provide a suitable match pair in a future iteration of the algorithm.

IV: Stage 4: Pairwise comparison, choosing the best pair

There were a couple of techniques attempted. Initially we chose a random piece and compared its “interesting” curves with all other fragments and their respective “interesting” curvatures. This technique can be improved. Analyze all the pieces, instead of a one-to-many comparison looking for the best match for a given fragment, or do a many-many comparison. Compare all the pieces and only choose the best match, where the best match is determined by the Pearson Correlation Coefficient from stage 3. This provides an improved matching algorithm because it does not always rely on the piece in question making a match. This technique has the ability to assemble small subsections of the original fragment with higher certainty in the initial stages. This high certainty in the initial stages provides the framework to successfully reassemble of the object. A good analogy for the one-to-many comparison would be building a bridge with a foundation out of wood. It could work, but it’s probable that the steel support structure would crush the foundation in the later stages of building. With the many-to-many comparison we start the bridge building process with an extremely solid concrete foundation. This solid foundation provides a very good initial start to the building process and ensures a greater certainty there will be less errors in the beginning of the reassembly process.

5. Experimental results-Section 4

The algorithm is tested on a pottery like object, (*Fig. 7.* left). The cup was broken into larger pieces and these pieces were scanned, (*Fig. 7.* center). The point clouds produced by the laser scan were used to produce 3D models of each piece as represented by (*Fig. 7.* right). Later the perimeter curves were extracted *Fig. 7.* lower. The choice of object is domain specific, the small facet of fracture does not provide many face features and makes feature extraction difficult. By focusing on the curvature of the fracture we are able to reproduce the original object as shown in *Fig. 8.*



Fragment 1 and 6 Merged

Fragment 1,6,3 Merged

Fragment 1,6,3,2 Merged

Fragment 1,6,3,2,5 Merged

Fig. 8. Reconstruction of 3D virtual fragments.

6. Future work

We would like to determine how the algorithm handles objects with larger faces such as Huang's gargoyle or the map fragments from the Forma Urbis Romae project (Koller *et al*. 2005). Domains where there are many like curves, such as the Forma Urbis Romae project are predicted to be difficult using only the curvature of the break point. Another interesting strategy would be to use the texture information of an object to help the alignment of pieces in stage 2. Incorporating the curvature techniques outlined in this paper, with facet feature based techniques Huang *et al*. 2006 one should obtain better matches in all domains.

References

Ercil, Aytul, Yucel Yemez, Lale Akarun and Ali Vahit Sahiner (2003). *Caria: Computer Assisted Reconstruction in Archaeology*. Sabanci University, Lecture.

Huang, Qi-Xing, Simon Flory, Natasha Gelfand, Michael Hofer and Helmut Pottmann (2006). Reassembling Fractured Objects by Geometric Matching. *ACM Transactions on Graphics* 25(3), 569–578.

Koller, David, Jennifer Trimble, Tina Najbjerg, Natasha Gelfand and Marc Levoy (2005). Fragments of the City: Stanford's Digital Forma Urbis Romae Project. *This Williams Symposium on Classical Architecture*. *Journal of Roman Archeology*.

Manay, Siddharth, Daniel Cremers, Byung-Woo Hong, Anthony Yezzi and Stefano Soatto (2006). Integral Invariants for Shape Matching. *Transactions on Pattern Analysis and Machine Intelligence, IEEE* 28(10), 1602–1618.

Moreau, R., Ochoa V., Pham M. T., Boulanger Pierre., Redarce T., Duuis O., (2006). *Evaluation of Obstetric Gestures: An Approach based on the curvature of a 3D positions*.

Papaioannou, Georgios, Evaggelia Karabassi and Theoharis Theoharis *(2000). Segmentation and Surface Characterization of Arbitrary 3D Meshes for Object Reconstruction and Recognition. Proceedings of the International Conference on Pattern Recognition '2000, IEEE. 734–737*.

Papaioannou, Georgios, Evaggelia Karabassi and Theoharis Theoharis *(*2001). *Virtual Archaeologist: Assembling the past. Computer Graphics and Applications. IEEE* 21(2) 53–59.

Pearson, Karl (1896). Mathematical contributions to the theory of evolution III.Rregression, heredity and panmixia. *Philosophical Transactions of the Royal Society* 187, 253–318.