EUROGRAPHICS 2013 / I. Navazo, P. Poulin
(Guest Editors)

Volume 32 (2013), Number 2

Interactive Exploration and Flattening of Deformed Historical Documents

Kazim Pal Melissa Terras Tim Weyrich

University College London

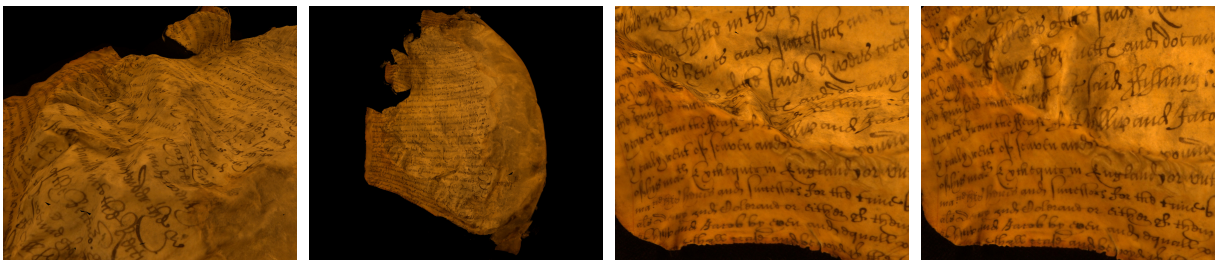


Figure 1: From left to right: A 3D reconstruction of a damaged parchment, a global flattening of the parchment, a locally-affine undistortion of a section of the text, and a local flattening of that same section.

Abstract

We present an interactive application for browsing severely damaged documents and other cultural artefacts. Such documents often contain strong geometric distortions such as wrinkling, buckling, and shrinking and cannot be flattened physically due to the high risk of causing further damage. Previous methods for virtual restoration involve globally flattening a 3D reconstruction of the document to produce a static image. We show how this global approach can fail in cases of severe geometric distortion, and instead propose an interactive viewer which allows a user to browse a document while dynamically flattening only the local region under inspection. Our application also records the provenance of the reconstruction by displaying the reconstruction side by side with the original image data.

Categories and Subject Descriptors (according to ACM CCS): I.3.8 [Computer Graphics]: Applications—I.7.5 [Document and Text Processing]: Document Capture—J.5 [Computer Applications]: Arts and Humanities—

1. Introduction

In this paper we present an interactive system for navigating the surface of a 3D reconstruction of a document which at all times undistorts the local region of text that the user is currently focussed on. Our system aims to improve the accessibility and legibility of text in highly distorted documents such as those shown in Figure 2. The work is particularly relevant to libraries and archives who often possess highly valuable historical documents whose contents are inaccessible and which cannot be restored by conventional means, or read in person due to the high levels of damage and the fragile nature of the material. Such damaged documents are surprisingly common in archives across the world.

Previous work addressing this problem typically deals with documents exhibiting small geometric distortions; however, many historical documents contain more severe types of damage. A common medium for medieval documents is parchment, which is made from limed animal hide and therefore consists of an irregular structure of organic fibres. These fibres are sensitive to their environment and can shrink, swell, and buckle if exposed to heat or humidity, creating dramatic and irregular geometric distortions. We

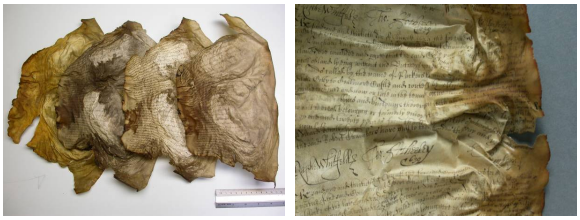


Figure 2: Folios with strong geometric distortions caused by fire and moisture damage. (Reproduced with the permission of The Honourable The Irish Society and the City of London Corporation, London Metropolitan Archives)

demonstrate our approach on a set of centuries-old parchment called the Great Parchment Book, a 17th century property survey of the Ulster estates managed by the City of London commissioned by Charles I whose contents are of great interest to historians studying the history of this region. The book fell victim to both fire and moisture damage, causing such strong distortions which conservators tell us cannot be restored by current conservation methods.

Our system also addresses the issue of *provenance*. For historians studying the text through a digital representation, it is important to be able to judge whether a feature present in the digital representation was also present in the original text or whether is it an artefact of the reconstruction pipeline. Terras [Ter11] discusses this issue at length, focussing mainly on imaging artefacts. However in our case the most likely source of error is the 3D reconstruction process. We therefore document the provenance of the reconstruction by providing the user with smart access to the original image collection. By comparing the 3D reconstruction with the original images the user can better assess the content of the text in areas of the 3D reconstruction which seem to contain errors.

The viewer can also explore arbitrary 3D models, to inspect interesting surface details on objects for the purpose of, for example, archaeological or forensic examination.

2. Related Work

The majority of the previous work on document flattening deals with the problem of flattening images of printed text captured by a flat-bed scanner or a camera, and makes limiting assumptions about the content of the images. For example, Zhang & Tan [ZT05] assume the physical deformations to be caused by the spine of a book, and both Tian & Narasimham [TN11] and Schneider et.al. [SBR07] rely on the text being printed on a light background in a regular font so that individual letters and strokes can be detected.

Brown & Seales [BS01] and Sun et.al [SYY*05] approach the problem of flattening more general documents, with fewer assumptions about their shape or content, by capturing a 3D model of the document using a structured-light scanner and a single image, and then flattening the resulting mesh with a mass-spring model. The mesh is allowed to fall into a planar configuration under a gravity force while spring forces maintain its structure. We observe that this mimics the physical conservation approach of softening the parchments and then stretching them out. Brown et.al. [BP05,BSY*07] later replace the mass-spring approach with the Least-Squares Conformal Mapping (LSCM) algorithm [LPRM02], which maps the mesh onto a 2D plane.

Pietroni et.al. [PCCS11] proposed a system in which a user is able to interactively cut out subsections of 3D scan of a sculpture which can then be flattened and inspected in image space.

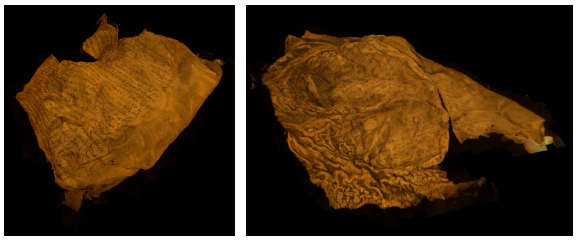
The approach of exploring a 3D model of a page and switching to original camera images is somewhat similar to Snavely et.al.'s Photo Tourism [SSS06] system where a user explores a scene by interpolating a collection of calibrated images and point geometry. The system allows the user to move from image to image and morphs between the images based on the related camera and point geometry.

3. Problem Analysis

Mass-spring based systems are not suitable for flattening documents which contain high levels of physical distortion as they can introduce self-intersections in the flattened mesh. Global parameterisation algorithms such as LSCM or Angle-Based Flattening [SdS01] [SLMB05] do not have this issue since they parameterise the underlying 2D manifold, but they can suffer from other problems.

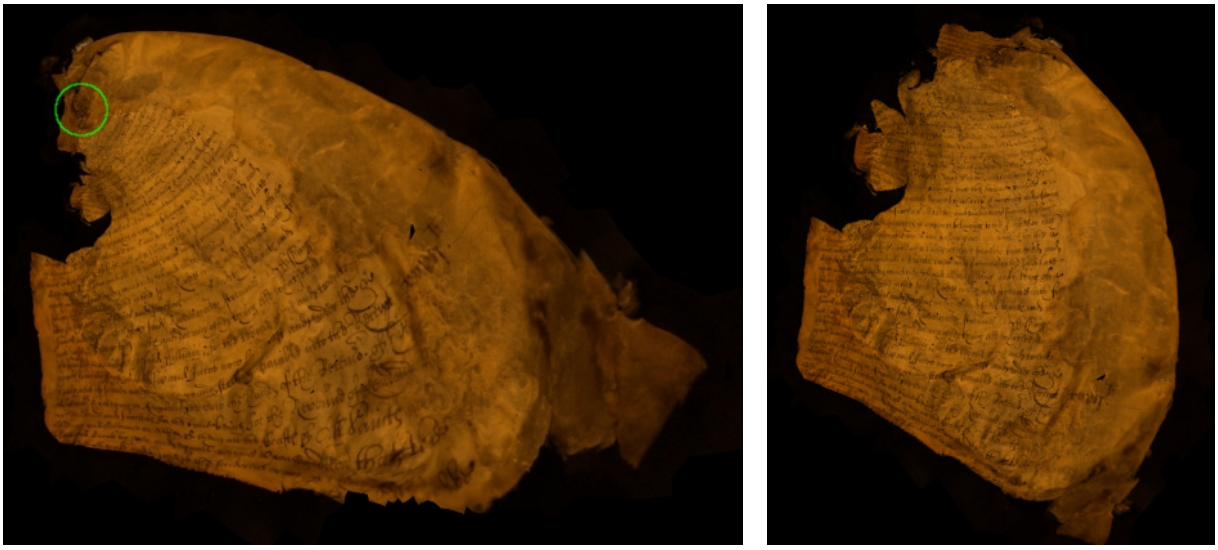
Conformal mappings introduce large stretch deformations in meshes with a high surface-area-to-perimeter ratio (as can be the case in severely damaged documents). Also, in meshes containing small bridges and other artefacts that are often introduced when reconstructing complex and intricate geometry, they can produce extremely sharp deformations around the bridge which affect the entire mesh. Extra mesh cleaning could potentially help with this problem, but removing these artefacts automatically is very difficult.

Finally, as is common with damaged parchment, the documents may contain non-isometric distortions such as shrinking and swelling which cannot be undone by flattening alone. Such distortions typically do not vary smoothly over the parchment due to its irregular fibre structure and the fact



(a) (b)

Figure 3: 3D Reconstructions of Parchments A and B



(a) (b)

Figure 4: Conformal global flattening of Parchment A. In (a) the mesh contains a single small bridge (in the area marked in green) produced as an artefact of the 3D reconstruction pipeline. In (b) this artefact has been manually removed prior to flattening.

that certain areas of a page can receive different types and amounts of damage. This means that estimating the true distortion field which would return the document to its original state is infeasible.

The result of globally flattening Parchment A (Figure 3a) conformally can be seen in Figure 4. Figure 4a shows the extreme distortions which occur due to a small bridge in the mesh geometry, but even when the bridge is removed manually the flattened text (Figure 4b) still contains shrinking, shearing, and bending distortions which result in non-rectified text.

Parchment B exhibits very severe wrinkling around its edges (seen in the bottom of Figure 3b) as well as a large crater-like deformation in its centre, giving it a high surface-area-to-perimeter ratio. Flattening this parchment globally results in drastic stretch deformations as shown in Figure 5.

These distortions are caused in part by the fact that conformal mappings preserve angles, not lengths. Alternatives include Tenenbaum et.al.'s Isomap algorithm [TSL00] which we can use to compute a mapping of the mesh into 2D which preserves the geodesic distances between its vertices. This type of mapping removes the problem of stretch distortion but since it is not angle-preserving, it can introduce shear distortions (Figure 6) which can make the text illegible.

4. Approach

We instead adopt the approach of undistorting local subsets of the mesh. Local regions will have a lower surface-area to perimeter ratio and will thus undergo very little stretch distortion, and since they are flattened independently they will not be affected by reconstruction artefacts elsewhere in the mesh. Lastly, we can more easily rectify these local patches so that the text always runs from left to right without having to compute a consistent global distortion field.

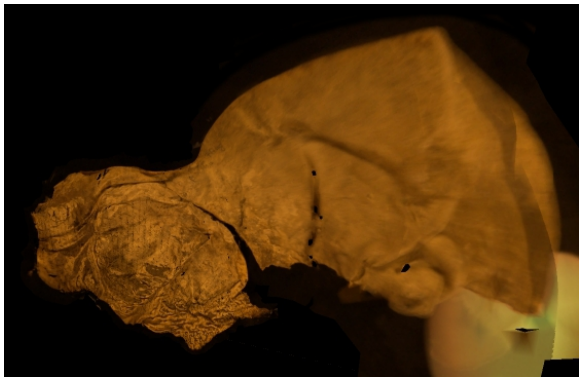


Figure 5: A global flattening of Parchment B creates large stretch distortions due to its large surface-area-to-perimeter ratio.

We can make an analogy to the problem of map making. The global-flattening approach is equivalent to unwrapping the entire globe onto a single map of the earth which will contain the distortions typical of large map projections. Our approach is equivalent to navigating over the 3D globe while creating numerous local maps as we go, each of which contains low levels of distortion.

The goal of our viewer is that, as the user looks at a particular region (the target region) of the mesh, it should be

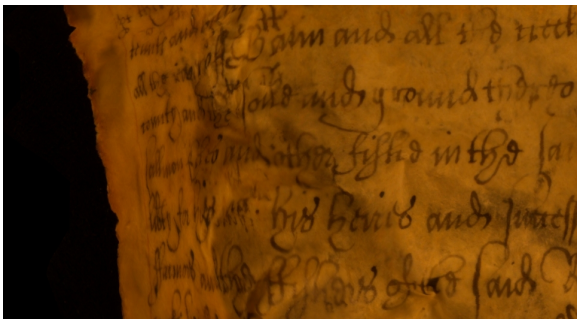


Figure 6: Isometric flattening of Parchment A. Anisotropic distortions, especially at the edge of the page, make the text unreadable.

displayed in a way that is optimal in terms of its readability. This optimality can be defined in terms of three sub-goals:

1. **Visibility**: The text in the region should all be visible so that the reader can properly take account of the context.
2. **Clarity**: The text should not appear distorted.
3. **Screen-Alignment**: Lines of text should be rectified with the screen so that it runs horizontally from left to right, and navigation over the parchment should follow the lines.

We propose two modes of undistortion to solve goals (1) and (2):

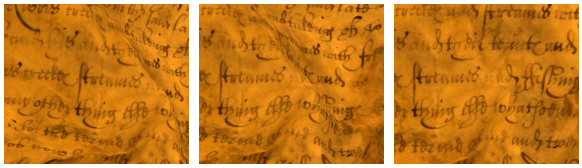
1. **Local-affine**: the mesh is rendered in 3D, transformed so that the target region is oriented to face the camera.
2. **Local-flattening**: the target region is flattened into 2D (independently of the rest of the mesh).

In order to obtain an estimate of the physical warp due to damage, we employ a pre-processing step in which we compute a vector field V defined on the mesh surface which flows parallel to the lines of text. We use V to solve goal (3). We can rectify the text since V encodes the local orientation of the text at each point, and can navigate along lines of text by moving through the flow lines of V. Moving horizontally on screen corresponds to moving parallel to V, and moving vertically on screen to moving perpendicular to V.

To allow interactive speeds, we use the local-affine mode as the user pans over the mesh (essentially a preview of the local-flattening), and the local-flattening mode when the user stops panning. This interaction sequence is illustrated in Figure 7 , and in the supplementary video.

4.1. Interactive Viewer

Our application at all times keeps a view target $\mathbf{p}$ which is a point on $\mathcal{M}$ with associated normal $\mathbf{n}$, and flow-vector $\mathbf{v}_{right}$ found by evaluating V at $\mathbf{p}$. The user is able to click and drag to move $\mathbf{p}$ over the mesh surface. In any given step, $\mathbf{p}$ is updated by transforming the on screen drag motion into motion through the flow lines of V.



(a) (b) (c)

Figure 7: The *local-affine view* and *local-flattening* modes. Between (a) and (b) the user pans slightly to the right. In (c) the user stops panning and local-flattening is performed, removing perspective distortions and revealing otherwise hidden text.

For local-affine undistortion, we define a local basis at $\mathbf{p}$ with axes $\mathbf{n}$, $\mathbf{v}_{right}$ and $\mathbf{v}_{up}$. We define $\mathbf{v}_{up}$ as $\mathbf{n} \times \mathbf{v}_{right}$ could compute it separately to account for shearing of the text. We then compute an affine transformation between this local basis and a canonical basis aligned with screen-space, and render the mesh under this transformation.

To perform local-flattening around $\mathbf{p}$ we extract a subset of the mesh consisting of all the triangles within some geodesic radius of $\mathbf{p}$ and then flatten this sub-mesh using LSCM [LPRM02]. The resulting mesh $\mathcal{M}^F$ is arbitrarily rotated, so we perform a registration step to align $\mathcal{M}^F$ with the current view of $\mathcal{M}$.

Letting $\mathbf{t_p}$ be the triangle in $\mathcal{M}$ containing $\mathbf{p}$ and $\mathbf{t'_p}$ the corresponding triangle in $\mathcal{M}^F$, we project both $\mathbf{t_p}$ and $\mathbf{t'_p}$ into the current viewing camera and align the resulting screen-space triangles using Ordinary Procrustes Analysis [CT01].

4.2. Provenance

The provenance feature allows the user to click on some point on the mesh and view the image from the original collection which best observes that point. The image is aligned with the current view of the mesh so that the user can compare the two easily. We use a similarity transform for alignment since the purpose of this feature is to display the original data with minimal processing, and using a more complex registration would defeat this purpose.

The best image for each vertex is defined as the image with maximal weight under the Unstructured Lumigraph blending scheme [BBM*01]. These weights are precomputed during the 3D reconstruction step.

We then define the best image for any point $\mathbf{p}$ on $\mathcal{M}$ to be the image assigned to the vertex $\mathbf{v}$ nearest to $\mathbf{p}$. The weighting scheme ensures that this image fully observes all triangles incident on $\mathbf{v}$, so it is guaranteed to observe $\mathbf{p}$;

Now, when the user clicks on a point on screen we back-project this point onto the mesh to obtain a point $\mathbf{p}$ lying in

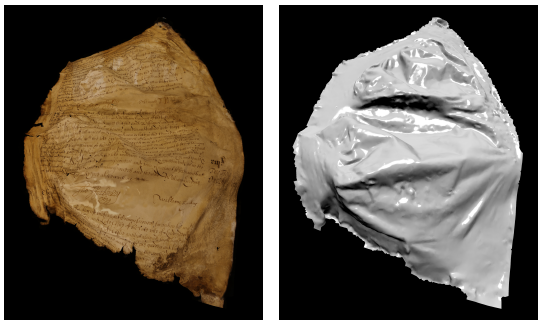
some triangle $\mathbf{t_p}$, with an associated image I. We register I with the current view of the mesh by aligning the projection of $\mathbf{t_p}$ in I with its projection in the viewing camera, again using Procrustes Analysis.

5. Pre-Processing

To explore a manuscript in our viewer, we first capture a set of high-resolution images and then perform two pre-processing steps: reconstruction, and computation of the text-flow vector field. Between 40 and 60 20 MP images are typically required to thoroughly cover the whole surface and reach all folds and creases of a crumpled document, and generate a high-quality reconstruction.

5.1. Reconstruction

We first use Wu's VisualSFM software [Wu11, Wu07, WACS11] and Furukawa's PMVS algorithm [FP10] to calibrate the images and generate a dense point cloud, which we then mesh using Poisson Surface Reconstruction [KBH06]. These are all state-of-the-art algorithms which produces high-quality results consistently. Finally, high-resolution texture maps are generated using the method of Schmitt & Yemez [SY99] by combining the input images using Unstructured Lumigraph blending [BBM*01], in which each image is assigned a weight at each vertex based on the angle at which it observes the vertex, the resolution of the image, and the visibility of the vertex. At each vertex we also store the index of the image with the greatest weight for use in the provenance feature. Note that our system does not rely on this pipeline and will work with models generated by other reconstruction methods. Figure 8 shows an example reconstruction of a parchment.



(a) (b)

Figure 8: Reconstruction of a page, rendered with and without texture to show 3D shape.

5.2. Text-Flow Field

To compute the vector field V, we first apply Bossen & Heckbert's re-meshing algorithm [BH96] to $\mathcal{M}$ to generate a second mesh $\mathcal{M}^S$ containing k evenly distributed vertices. Then we compute k sample-points, $\mathbf{s}_1 ... \mathbf{s}_k$, by projecting the vertices of $\mathcal{M}^S$ onto $\mathcal{M}$. At each sample point $\mathbf{s}_i$ with normal vector $\mathbf{n}_i$, we define an initial basis B_i^0 by rotating the standard basis $(\hat{i}, \hat{j}, \hat{k})$ to align the y-axis with $\mathbf{n}_i$ and the x and z-axes with the tangent plane of $\mathcal{M}$ at $\mathbf{s}_i$.

Next, we determine the angle θ_i such that rotating B_i^0 by θ_i about $\mathbf{n}_i$ will align the x-axis with the direction of the text. We use the observation, similarly to Schneider et. al. [SBR07], that the distribution of the horizontal integration of a patch of text exhibits the highest variance when the text is properly aligned (Figure 9).

We extract a subset, $\mathcal{M}_i$, of $\mathcal{M}$ consisting of all the triangles which lie completely within a geodesic radius r of $\mathbf{s}_i$, and then apply LSCM [LPRM02] to $\mathcal{M}_i$ to obtain a flattened sub-mesh $\mathcal{M}_i^F$. To find the optimal rotation for this patch, we render $\mathcal{M}_i^F$ and use Otsu's method [Ost79] to generate a binary image I_{seg} with the text roughly segmented from the background. This segmentation need only be rough since we never need to accurately localize individual characters. We search for the optimal rotation angle between maximum and minimum possible rotations θ_{min} and θ_{max}. These are pre-selected by a user as sensible bounds to optimize for speed and disambiguate upside-down text. For each possible angle $\theta \in [\theta_{min}, \theta_{max}]$ we rotate I_{seg} about its centre by θ to generate a rotated patch I_θ, and then compute vector h_θ by integrating I_θ along the horizontal axis. We select the angle θ for which the variance of h_θ is maximal. Results of this process are shown in Figure 9.

By rotating B_i^0 about $\mathbf{n}_i$ by θ_i, we generate a local text-aligned basis with axes $\mathbf{v}_{i,normal}$, $\mathbf{v}_{i,right}$, and $\mathbf{v}_{i,up}$. The vectors $\mathbf{v}_{i,right}$ at each sample point make up the vector field V. We smooth V with a mesh-based Gaussian kernel according to the connectivity of $\mathcal{M}^S$, and interpolate its values within triangles using barycentric coordinates. Figure 10 shows a visualization of such a vector-field.

The text-flow field works well for guiding the navigation of the viewer. We also investigated using V in a global flattening approach but found it to be difficult to incorporate with a global parameterisation algorithm, and unreliable when incorporated into a mass-spring approach.

For arbitrary 3D models without text, we compute a sensible proxy for V which runs around the object, similar to lines of latitude on the earth. In the future, for shapes where this is ill-defined the proxy V could be defined as the isolines of a geodesic distance field computed using the method of Surazhsky et.al. [SSK*05].

6. Results

We show visualisation results using our application on reconstructions from two parchments provided by London Metropolitan Archives. The meshes contain approximately 50,000 triangles, and the textures are in the order of 100

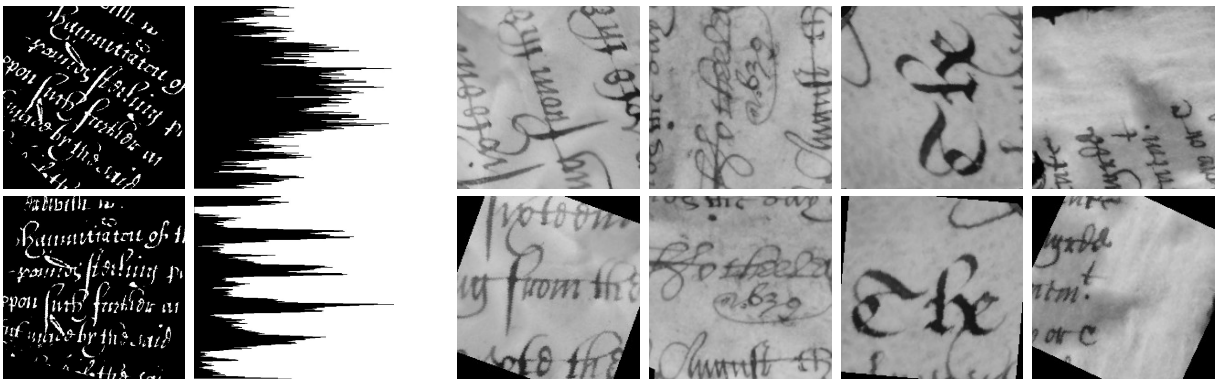


Figure 9: *Left:* Horizontal integration of a segmented text patch for various rotations. The distribution of the horizontal projection shows stronger peaks and higher variance as the text lines are more closely aligned with the horizontal axis. *Right:* Each pair shows a document patch before and after the computed optimal rotation, with $(\theta_{min}, \theta_{max}) = (-135, -45)$

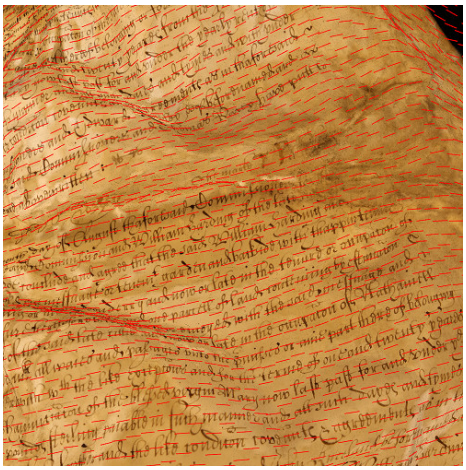


Figure 10: A visualization of a text flow field. At each sample point $\mathbf{s}_i$, a red line is rendered parallel to the computed value of $\mathbf{v}_{i,right}$.

MP. Figure 11 shows example results of our system in local-affine mode (left) and local-flattening mode (right). In the local-affine mode the text is correctly aligned in the viewing camera so that it runs from left to right, however distortions are still visible in very crumpled regions. In Figure 11e a large amount of text is completely obscured from view by a fold, in Figures 11a and 11c the text around the fold is not obscured but is distorted and foreshortened by perspective, and in Figures 11g and 11i, even though most of the text is visible, the distortions and occlusions caused by the sharp creases and fine wrinkles have a significant detrimental effect on legibility.

We can see how the local-flattening mode solves these problems. In Figure 11f the page is unfolded and a large previously occluded chunk of text becomes visible, in Figures 11b and 11d, the foreshortening and perspective distortion is removed, and in Figures 11h and 11j the fine creases and wrinkles are opened out revealing some previously occluded text and significantly reducing the overall distortion.

6.1. Provenance

In Figure 12 , we can see an example of the importance of provenance. The word circled in red looks very unusual and could reasonably be thought to be an error introduced by the reconstruction process. By inspecting the same region of the page as seen by one of the source images, we can see that the reconstruction is actually an accurate rendition of the text.

6.2. User Evaluation

The key success metric for our system is the ease with which an expert palaeographer can read the text, specifically the confidence with which they can identify certain words and letters. This quality is difficult to measure, so we instead evaluated our system with three users with experience in transcribing medieval scripts. The users were asked to spend time transcribing texts presented in three modalities: a globally flattened static image, the viewer with only local-affine undistortion enabled, and the fully functional viewer. Static images were used which did not contain large stretch distortions or distortions due to reconstruction artefacts. We then interviewed the participants to try to understand their experience using each modality.

We found that a static image is preferable in areas where the text was originally fairly undistorted since a global approach effectively flattens such areas, and the simple 2D pan-and-zoom style of interaction with images is simple and familiar. However in highly crumpled areas, the feedback indicated that text was harder to decipher in a static image, and being able to manipulate the page in 3D was useful. One participant commented that they *"would not trust the [global] image on its own"* and that it was *"helpful to have the option of looking into creases"* in local-affine mode. One participant observed that the local-flattening mode allows for easier

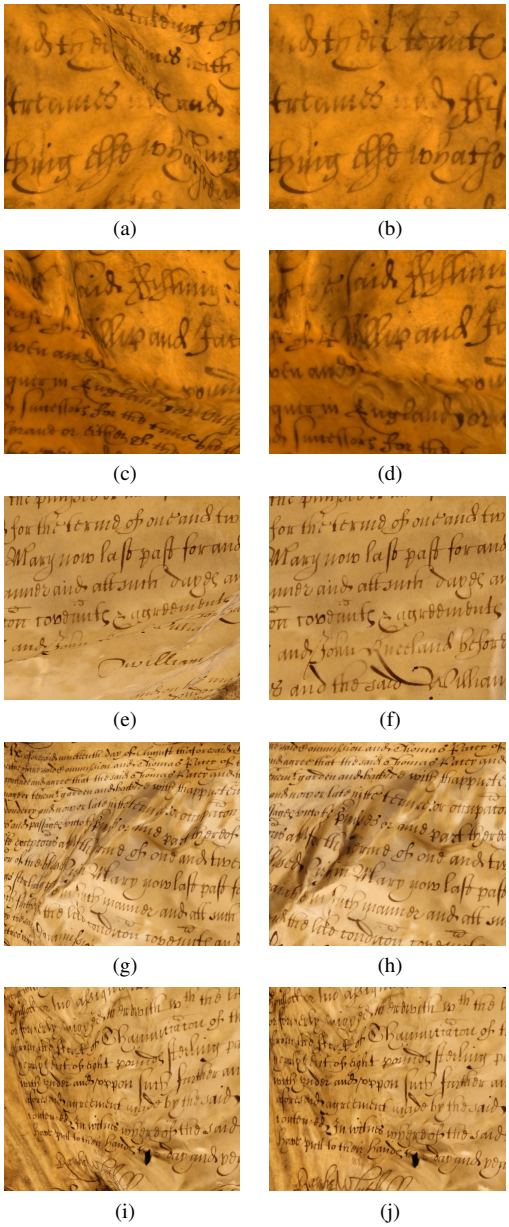


Figure 11: Sections of the mesh rendered in local-affine mode (left) and local-flattening mode (right). Note that the original wrinkles remain visible since the shading is baked into the texture

reading the majority of the time but that in certain crumpled areas, navigating back and forth over a crease allowed them to view a difficult word from a number of angles, mimicking the standard palaeographic method of physically manipulating the manuscript to observe that word from multiple viewpoints. They also pointed out that exploring a crumpled area in local-affine mode gave them *"a better sense of the level of damage"* which could then influence they way in which

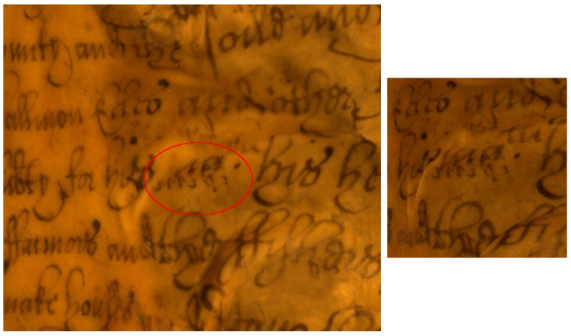


Figure 12: The text circled in the view of the 3D mesh (left) looks as though it may be an erroneous artefact of the reconstruction pipeline. However, inspection of the corresponding image (right) shows it to be accurate.

they transcribed that area in local-flattening mode.

Generally the participants indicated that having both modes available and being able to go back and forth between each was useful and allowed them to read the text with more confidence.

They also found the provenance feature to be useful for the reasons which we intended. The participants confirmed that it was useful for transcribing difficult areas where the reconstruction was unclear, and also because it allows the transcriber to "trust what they see" in the reconstruction.

6.3. Other Objects

We also demonstrate the viewer's ability to navigate other types of objects. Figure 13 shows a 3D model of a statue

Figure 13: Inspecting a 3D model of a statue. Left: 3D view of the statue's head as the user navigates over the surface. Right: The head unwrapped into a plane.

inspected in our viewer. This model is a closed surface so a single global flattening does not even exist. Globally parameterising a closed surface requires it to be cut into subsets which are parameterised separately. Our local approach effectively performs this role by selecting a single subset around the current viewing target.

7. Conclusion

We have presented an interactive system for exploring highly distorted documents in which the user is always shown an undistorted view of the region of the document they are inspecting. We have demonstrated that the current flattening methods can fail in cases of highly distorted documents and proposed an alternative approach in which local regions of the text are undistorted dynamically. We have shown how this interactive, local flattening is preferable to static, global flattening for exploring documents with large distortions.

We proposed two modes of undistortion: a *local-affine* mode where the local region of text is aligned with the screen by an affine transformation, and a *local-flattening* mode which also flattens that local region by parameterising it in 2D. We have shown how the local-flattening mode is superior to the local-affine mode as it unfolds creases in the page to expose text that is otherwise foreshortened or hidden, but also that both undistortion modes are useful in the transcription process when used in conjunction. The provenance feature was also show to be helpful for resolving ambiguities and giving the transcriber a greater sense of trust in the reconstruction. One of our participants was the palaeographer who will be performing the transcription of the Great Parchment Book, and as a consequence of this preliminary study our system will be used to aid that transcription process.

8. Acknowledgements

This work was supported by London Metropolitan Archives and the UCL EngD VEIV Centre for Doctoral Training.

References

[BBM*01] BUEHLER C., BOSSE M., MCMILLAN L., GORTLER S., COHEN M.: Unstructured lumigraph rendering. In *Proc. 28th ACM Conference on Computer Graphics and Interactive Techniques* (2001), pp. 425–432. 4, 5

[BH96] BOSSEN F., HECKBERT P. S.: A pliant method for anisotropic mesh generation. In *Proc. International Meshing Roundtable* (1996), pp. 63–74. 5

[BP05] BROWN M. S., PISULA C. J.: Conformal deskewing of non-planar documents. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* (2005), vol. 1, pp. 998–1004. 2

[BS01] BROWN M., SEALES W.: Document restoration using 3d shape: a general deskewing algorithm for arbitrarily warped documents. In *Proc. IEEE International Conference on Computer Vision* (2001), vol. 2, pp. 367–374. 2

[BSY*07] BROWN M., SUN M., YANG R., YUN L., SEALES W.: Restoring 2d content from distorted documents. *IEEE Trans. Pattern Analysis and Machine Intelligence 29* (2007), 1904–1916. 2

[CT01] COOTES T., TAYLOR C.: Statistical models of appearance for computer vision. *World Wide Web Publication, February* (2001). 4

[FP10] FURUKAWA Y., PONCE J.: Accurate, dense, and robust multiview stereopsis. *IEEE Trans. Pattern Analysis and Machine Intelligence 32*, 8 (2010), 1362–1376. 5

[KBH06] KAZHDAN M., BOLITHO M., HOPPE H.: Poisson surface reconstruction. In *Proc. 4th Eurographics Symposium on Geometry processing* (2006), pp. 61–70. 5

[LPRM02] LÉVY B., PETITJEAN S., RAY N., MAILLOT J.: Least squares conformal maps for automatic texture atlas generation. In *Proc. 29th ACM Conference on Computer Graphics and Interactive Techniques* (2002). 2, 4, 5

[Ost79] OSTU N.: A threshold selection method from gray-level histograms. *IEEE Trans. Systems, Man and Cybernetics 9*, 1 (1979), 62–66. 5

[PCCS11] PIETRONI N., CORSINI M., CIGNONI P., SCOPIGNO R.: An interactive local flattening operator to support digital investigations on artwork surfaces. *IEEE Trans. Visualization and Computer Graphics 17*, 12 (2011), 1989–1996. 2

[SBR07] SCHNEIDER D., BLOCK M., ROJAS R.: Robust document warping with interpolated vector fields. In *Proc. 9th International Conference on Document Analysis and Recognition* (2007), vol. 1, pp. 113–117. 2, 5

[SdS01] SHEFFER A., DE STURLER E.: Parameterization of faceted surfaces for meshing using angle-based flattening. *Engineering with Computers 17*, 3 (2001), 326–337. 2

[SLMB05] SHEFFER A., LÉVY B., MOGILNITSKY M., BOGOMYAKOV A.: Abf++: fast and robust angle based flattening. *ACM Trans. Graph. 24*, 2 (2005), 311–330. 2

[SSK*05] SURAZHSKY V., SURAZHSKY T., KIRSANOV D., GORTLER S. J., HOPPE H.: Fast exact and approximate geodesics on meshes. *ACM Trans. Graph. 24*, 3 (2005). 5

[SSS06] SNAVELY N., SEITZ S. M., SZELISKI R.: Photo tourism: exploring photo collections in 3d. *ACM Trans. Graph. 25*, 3 (2006), 835–846. 2

[SY99] SCHMITT F., YEMEZ Y.: 3d color object reconstruction from 2d image sequences. In *Proc. International Conference on Image Processing (ICIP)* (1999), vol. 3. 5

[SYY*05] SUN M., YANG R., YUN L., LANDON G., SEALES B., BROWN M.: Geometric and photometric restoration of distorted documents. In *Proc. IEEE International Conference on Computer Vision* (2005), vol. 2, pp. 1117–1123. 2

[Ter11] TERRAS M.: *Artefacts and errors: Acknowledging issues of representation in the digital imaging of ancient texts*. Books on Demand, 2011, pp. 43–61. 2

[TN11] TIAN Y., NARASIMHAN S.: Rectification and 3d reconstruction of curved document images. *Proc. IEEE Conference on Computer Vision and Pattern Recognition* (2011), 377–384. 2

[TSL00] TENENBAUM J. B., SILVA V. D., LANGFORD J. C.: A global geometric framework for nonlinear dimensionality reduction. *Science 290*, 5500 (2000), 2319–2323. 3

[WACS11] WU C., AGARWAL S., CURLESS B., SEITZ S. M.: Multicore bundle adjustment. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* (2011), pp. 3057–3064. 5

[Wu07] WU C.: Siftgpu: A gpu implementation of scale invariant feature transform (sift). `http://cs.unc.edu/~ccwu/siftgpu`, 2007. 5

[Wu11] WU C.: Visualsfm: A visual structure from motion system. `http://www.cs.washington.edu/homes/ccwu/vsfm/`, 2011. 5

[ZT05] ZHANG L., TAN C.: Warped image restoration with applications to digital libraries. In *Proc. 8th International Conference on Document Analysis and Recognition* (2005), pp. 192–196. 2