Revealing hidden text in rolled and folded papyri

Daniel Baum[1] · Norbert Lindow[1] · Hans-Christian Hege[1] · Verena Lepper[2,3] · Tzulia Siopi[2] · Frank Kutz[4] · Kristin Mahlow[5] · Heinz-Eberhard Mahnke[6,7]

Abstract Ancient Egyptian papyri are often folded, rolled up or kept as small packages, sometimes even sealed. Physically unrolling or unfolding these packages might severely damage them. We demonstrate a way to get access to the hidden script without physical unfolding by employing computed tomography and mathematical algorithms for virtual unrolling and unfolding. Our algorithmic approaches are combined with manual interaction. This provides the necessary flexibility to enable the unfolding of even complicated and partly damaged papyrus packages. In addition, it allows us to cope with challenges posed by the structure of ancient papyrus, which is rather irregular, compared to other writing substrates like metallic foils or parchment. Unfolding of packages is done in two stages. In the first stage, we virtually invert the physical folding process step by step until the partially unfolded package is topologically equivalent to a scroll or a papyrus sheet folded
only along one fold line. To minimize distortions at this stage, we apply the method of moving least squares. In the second stage, the papyrus is flattened, which requires the definition of a medial surface. We have applied our software framework to several papyri. In this work, we present the results of applying our approaches to mockup papyri that were either rolled or folded along perpendicular fold lines. In the case of the folded papyrus, our approach represents the first attempt to address the unfolding of such complicated folds.

1 Introduction

Our best sources for information about the culture of our ancestors and predecessors are their writings. Some of these writings are hidden in rolled, folded or scrambled documents. Due to recent improvements with respect to sensitivity and resolution, a few spectacular disclosures of hidden texts have been possible by X-ray tomography on heavily corroded metallic foils with imprinted writings [1, 2] and on parchment written with inks containing heavy elements, typically Fe [3–5]. Most recently, even letters written with carbon ink could be identified in carbonized papyri by phase contrast tomography [6].

Besides the mere detection of the script, the reconstruction of the text as a whole poses a challenge to the mathematical treatment of the tomographic data set. In contrast to more homogeneous material, such as metal foils or even parchment, the structure of papyrus poses further challenges. First, the stripes of the cut papyrus stems, which are arranged in two layers with the orientation of the stripes being orthogonal to each other, complicate the distinction between writings and substrate for iron gall ink or less X-ray absorbent inks. Second, virtual unfolding is

✉ Heinz-Eberhard Mahnke
mahnke@helmholtz-berlin.de

[1] Zuse-Institut Berlin, Berlin, Germany

[2] Ägyptisches Museum und Papyrussammlung, Staatliche Museen zu Berlin, Berlin, Germany

[3] Humboldt-Universität zu Berlin, Berlin, Germany

[4] Freie Universität Berlin, FB Geowissenschaften, Berlin, Germany

[5] Museum für Naturkunde Berlin, Berlin, Germany

[6] Freie Universität Berlin, FB Physik and TOPOI, Berlin, Germany

[7] Helmholtz-Zentrum Berlin, Berlin, Germany

hampered by physical separation of the two layers that can often be observed.

While the recovery of text on a scroll, no matter how badly the scroll had been distorted or compressed, has been achieved in various cases (e.g. [3–6]), the recovery from folded documents represents an even bigger challenge. In the investigation presented here, we report on the result of virtual unfolding performed on a mockup papyrus that was lettered with very highly absorbent red ink, made from cinnabar and minium, to demonstrate the algorithmic procedure and to separate it from the complex problem of uncovering the script on ancient papyri.

2 Objects, samples, and micro-tomography

Many objects, excavated in different archaeological campaigns, were then often buried again in museums' depots. With regard to papyri, this applies particularly to objects that are unreadable, since mechanical unrolling or unfolding would damage them. Therefore, there is a strong need to develop a non-destructive method that allows scholars to access the hidden texts of as many objects as possible. With the resolution of commercial table-top instruments available today, the procedure may simply be to sort out candidates for tomography by X-ray fluorescence and to decide on the imaging technique most appropriate to distinguish ink from substrate. When the X-ray fluorescence spectrum provides enough evidence for a metal ion containing ink to yield a good absorption contrast like Fe (or elements with even higher atomic number Z) in, e.g. iron gall ink, then absorption tomography may be the most straightforward way to get access to the inner secrets. That this procedure could be turned into routine means was demonstrated in the Apocalypto project [3] as well as with the recent publication on the En-Gedi scroll [5], both with parchment as writing substrate.

In the case of papyrus, we have followed the procedure outlined above after a successful start on a mockup sample with modern papyrus and modern iron gall ink. The object selected is a small package belonging to the objects in the papyrus collection of the Berlin Egyptian Museum. It is a piece found in the excavation by Rubensohn in Ashmonein in 1904/5, shown in Fig. 1. XRF spectra taken with a handheld portable X-ray fluorescence system (Niton XL3t, with radiation safety chamber, X-ray beam with 3-mm spot size) on different spots of the object showed varying Fe intensity, indicative for a possible hidden writing of text with iron gall ink. For this object, tomography data were taken at the GE Phoenix

Fig. 1 Papyrus package (ÄM ref. 185/B L301 from the Rubensohn excavation in Ashmunein 1904/1905, approximately 0.5 cm *thick* and with 3.7×3.7 cm^2 planar size) (*upper left*) with XRF spectrum taken at two different spots as marked in *red* and *yellow* (*upper right*). *Lower row:* Tomographic cut showing *white* spots possibly due to Fe (*left*), a 2D reconstructed inner part of the papyrus sheet is shown to the *right*
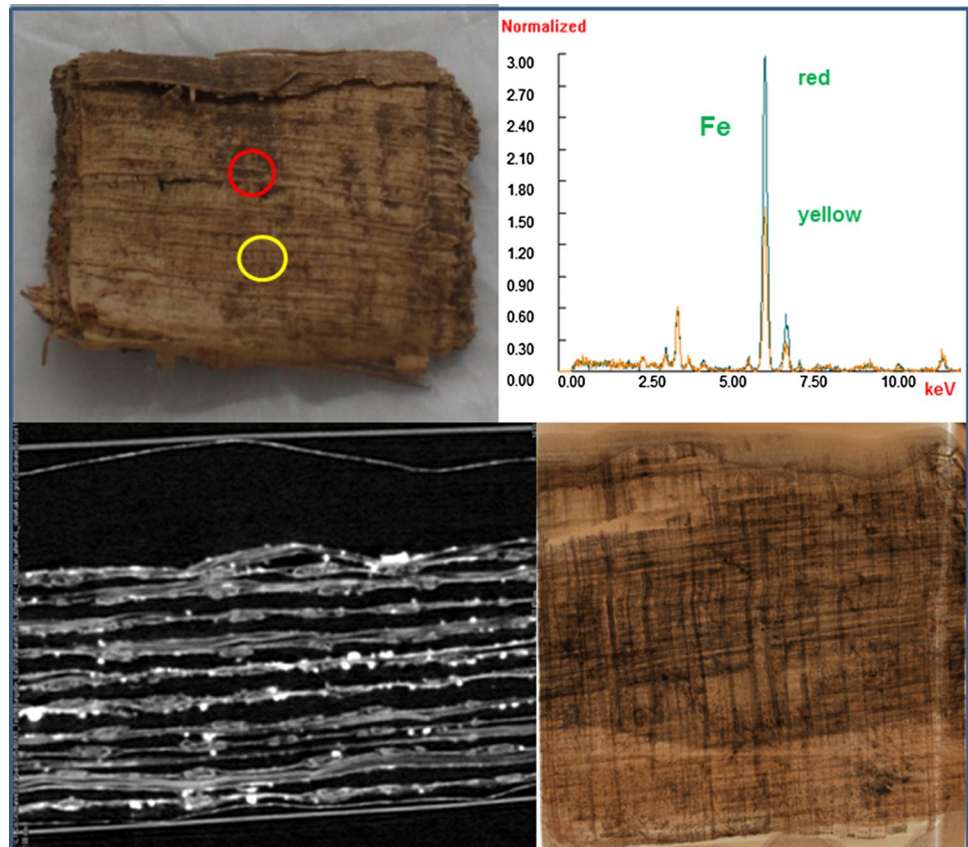
Nanotom s of the Museum für Naturkunde Berlin at 50 kV and 150 μA, using a W anode, and generating 1440 projections with an exposure time of 1 s per projection, and with an effective voxel size of 7.53 μm. To reach this high resolution despite the size of the object, the piece was scanned in four fractions. The object was put into an air-cushioned plastic box, which was then clammed in a cut plastic test-tube, mounted on the rotating chuck. To improve the signal-to-noise ratio, the X-ray intensity below the Fe absorption K-edge was reduced by putting a 50-μm thick aluminum foil between object and detector. In Fig. 1 a cut through the package clearly shows the layers of papyri in-between Japanese tissue and the plastic foil coverage of the air cushions. Extraction of an inner layer of the papyrus package utilizing the approach presented in this work yielded the image shown in the right of the figure. It nicely illustrates the structure of the papyri fiber with its quite pronounced intensities due to different thicknesses of the fibers and fiber bundles. This property, in particular, makes the separation of script from the writing material difficult.

To develop an optimized procedure for the virtual unrolling and unfolding, we have prepared sample papyri from modern material written with cinnabar (mercury red, dark red, HgS) and minium (lead red, light red, Pb_3O_4) ("mockups") for the best possible contrast in absorption tomography. The photo showing the written letters and symbols is depicted in Fig. 2. After forming a scroll, fluorescence spectra were taken at different spots to demonstrate the procedure, see the right part of Fig. 2. The choice of mercury with $Z = 80$ and lead with $Z = 82$ for maximizing the contrast for absorption is not arbitrary but reflects the mentioning and description already by Pliny the Elder [7]. In rare cases, especially for important and valuable texts, some letters, symbols or words, e.g. at the beginning of a paragraph, were emphasized by writing in red, using the expensive and more striking red of cinnabar or minium in contrast to the reddish iron oxides (often called "rubrum" for the Latin "red"). Such an emphasis was also common in later Arabic texts.

The scroll formed as shown in Fig. 2 was tomographed inside a plastic test-tube; the tendency for mechanical unrolling by its tension held the scroll in position. As already done for the ancient papyrus, the same setup was used for the acquisition of the data with computed tomography. With an effective voxel size of 9.17 μm, the scan was taken in two fractions. For a tomography on a folded package, the scroll was mechanically unrolled again, then folded according to the "magic fold", as described by Krutzsch [8] and as illustrated in the Figure S1 presented in the supplement, and mounted in an air-cushioned plastic box on the rotating chuck of the tomograph. Figure 3 compares the photographs of the scroll and the folded package with the 3D images obtained with tomography and visualized using volume rendering.

3 Algorithmic approaches for unrolling and unfolding

We have implemented two approaches that allow the virtual unfolding of folded papyrus packages. Approach 1 assumes that the papyrus has only been rolled (or folded only in one direction). This can be considered as a very simple form of fold. Approach 2 can cope with more complicated folds like, for example, the "magic fold" [8], which comprises folds along two approximately perpendicular fold lines.

For a rolled papyrus, each cross-section along the roll shows the papyrus as a spiral (Fig. 4a). To virtually unroll the papyrus roll, we define a two-dimensional (2D) surface that is placed in the three-dimensional (3D) tomogram and that describes the medial surface of the papyrus sheet. After flattening this 2D medial surface into the 2D plane together with the voxels in the close vicinity of the 2D surface, the papyrus can be visualized using standard visualization techniques like volume rendering (Fig. 4b). We define the medial surface by drawing contours in the

Fig. 2 Mockup of modern papyrus written with cinnabar and minium and XRF spectrum showing the Hg and Pb L-lines as well as the spectrum for the papyrus background well below the Hg and Pb spectrum taken at the indicated spots at the rolled papyrus

Fig. 3 Rolled and folded modern papyrus ("magic fold") (photos at the *top*) and their volume rendered 3D images at the *bottom*
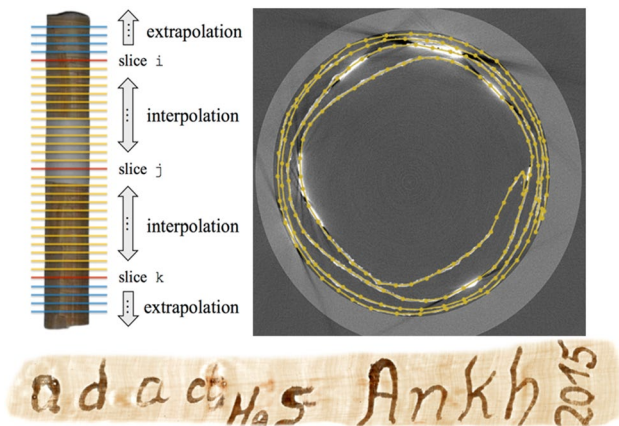

Fig. 4 a *Left*: Illustrating how the contour lines are defined for interpolating and unrolling. *Right*: Result of the linear interpolation after manually defining a few contours for the spiral contour along the rolled papyrus. **b** Top part of the virtually unrolled papyrus based on three manually placed spiral contours

cross-sections of the virtual roll, where each contour is an open polygonal line. However, instead of drawing a contour in each cross-section, we can take advantage of the fact that the cross-sections of consecutive slices show very similar spirals. This is often even true across the whole tomogram. Thus, we need to define contours only in a few cross-sections and can generate the other contours by interpolation across the whole volume along the roll. Ideally, the process starts by defining a contour somewhere in the middle of the roll. Let us denote this contour by c_j, where j is the index of the cross-section. This first contour is then copied in both directions to all other cross-sections, thereby often already giving a good initialization of the contours. If either of the interpolated contours diverges too far from the visible papyrus spiral, this contour can be corrected, thus defining a new contour. Now, assume that we have defined three contours c_i, c_j, c_k, where $i < j < k$ denote the indices of cross-sections in the tomogram (cf. red lines in Fig. 4a). Then the contours c_m, with $i < m < j$, will be the result of the interpolation between c_i and c_j (see yellow lines in Fig. 4a). Likewise, for cross-sections n, $j < n < k$, the contours c_n will be interpolated between c_j and c_k. For cross-sections p, with $0 < p < i$, the contours c_p are extrapolated from c_i (see blue lines in Fig. 4a). The same applies for contours c_q, with $k < q \leq N$, where N is the overall number of cross-sections. Of course, instead of three contours, we can define as many contours until we are satisfied with the contours in all cross-sections. The contours between two explicitly defined contours are always interpolated. Notice that we ensure that the number of points for all contours is the same. Thus, we have an implicitly defined connection of the points of consecutive slices, which allows an easy flattening of the 2D medial surface into the plane as a quadrangular mesh. In order to avoid distortions when flattening, distances between neighbored contour points are preserved as good as possible (see next section). Note that this approach cannot only deal with rolled papyri but also with papyri that have been folded in a fan-like way, that is, along one direction.

Approach 2 can deal with packages that have been folded several times from different (usually approximately orthogonal) directions. This approach iteratively unfolds the folded package until the unfolded papyrus is topologically equivalent to a rolled papyrus. At this point, we can apply a single unrolling step using Approach 1, thus completely unfolding the package. However, for the iterative unfolding steps, a new approach is needed. The goal here is to virtually undo each step of the physical folding process. Before we can address a next unfolding step, we first have to identify the direction of the fold line. This can usually be achieved by visualizing the current state of the folded package using volume rendering (see Fig. 3, bottom row) or iso-surface rendering together

Fig. 5 Iterative unfolding by defining the inner contour (*red line*) and the outer contour (*blue line*), the final step is similar to unrolling like in Fig. 4
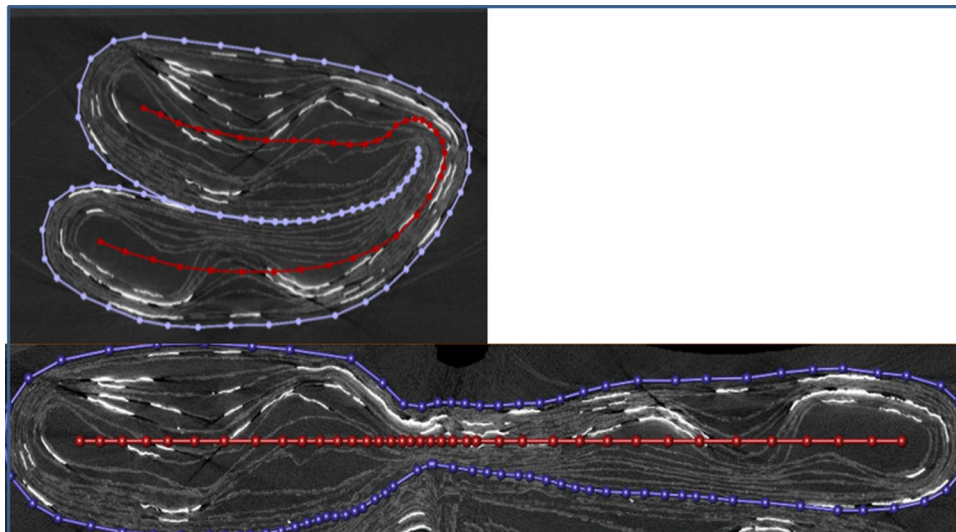
with an inspection of the cross-sections. For the actual unfolding, we apply the concept of Moving Least Squares (MLS) for image deformation similar to what was presented by Schäfer et al. [9]. In our implementation, we use two types of contours (Fig. 5). In fact, these are again sets of contours as in Approach 1, that is, for each cross-section contours are created. The contours are defined iteratively, starting again with a cross-section in the middle of the package and utilizing interpolation as far as possible. Notice that these two types of contour are different to the one we used in Approach 1, because they do not define the medial surface of the papyrus. Instead, the first contour (red contour in Fig. 5) represents the medial axis of the folded package, and the second contour (blue contour in Fig. 5) represents the outline of the package. In each iteration k of the unfolding process, we unfold the current image I_k (Fig. 5, upper image) to obtain a new image I_{k+1} (Fig. 5, lower image). To get image I_{k+1}, two steps are carried out: In the first step, the middle contour (red) is flattened such that the lengths of the line segments are preserved. Then, the second contour (blue) is deformed as rigidly as possible [9] to match the new position of the middle contour in I_{k+1}. This concludes the first step. In the second step, for each point in I_{k+1} enclosed in the transformed second contour, we look up the position in image I_k to acquire the original value at this point. This is done by again transforming each point as rigidly as possible, but this time from image I_{k+1} back to image I_k. Using this approach, we can fill in all the missing information into I_{k+1}, which allows us to proceed to the next unfolding step until we can apply Approach 1.

4 Distortion-reducing flattening

The last step in the unfolding process is always the 2D flattening of the mesh that was implicitly defined by the contours used in Approach 1. Since this mesh is defined in a 3D volume, the mesh usually cannot be flattened into a 2D plane without introducing distortions. This is even the case for rolled papyri, since the mesh usually does not exactly follow the medial surface of the papyrus, thereby introducing bulges. Further distortions are introduced when iteratively unfolding a package, despite the utilization of the MLS method, which locally minimizes distortions. The more unfolding steps are carried out, the more distortions might occur. Hence, care must be taken to reduce the distortions when flattening out the papyrus into the 2D plane. We use an approach that is similar to a mass-spring model [10, 11]. Another approach that might work equally well is Least-Squares Conformal Mapping (LSCM) [12, 13]. The first step in our algorithm is to flatten the implicitly given mesh into the 2D plane, without reducing distortions. Then we iteratively relax the mesh to dewarp it, until the gained improvements are very small or a maximum number of iterations has been achieved.

Remember that we linearly interpolate all contours between those contours that we manually defined. Also, remember that we extrapolate the first manually defined contour up to the first cross-section and the last manually defined contour up to the last cross-section. This means that if we have K manually defined contours, we can build our mesh from $N = K + 2$ contours, where the additional two contours are the ones in the first and last cross-section. Recall that we have the same number of points, M, in each

contour. Thus, our implicitly given mesh consists of $N \times M$ vertices. Let us denote these vertices by ν_{ij}, where i is the index of the contour, and j is the index of the point in contour i. By $\nu_{ij}{}^k$ we denote the position of vertex ν_{ij} after the kth iteration of the relaxation process. Initially, we place vertex ν_{11} at position $\nu_{11}{}^0 = (0,0)^T$ and vertices ν_{i1} at position $\nu_{i1}{}^0 = (0, p_i)^T$, where p_i is the distance between the first cross-section and the cross-section corresponding to contour i. Starting from these initial positions, we compute the initial positions of the other vertices as follows: $\nu_{ij+1}{}^0 = \nu_{ij}{}^0 + (l_{ij}, 0)^T$, where l_{ij} is the distance between the points j and $j+1$ in contour i in the respective cross-section, that is, in the 3D tomogram. The contours thus represent an implicitly quadrangular mesh. The initial mesh with vertex positions $\nu_{ij}{}^0$ of the unfolded package is shown in Fig. 6c. Note that the mesh has been mirrored here to match the image. To obtain a relaxed mesh, in each iteration of the relaxation process the algorithm tries to locally adjust the distances between the neighbors in the 2D mesh to the 3D distances of the respective contour points in the tomogram. In every iteration $k+1$, each vertex ν_{ij} is moved from its previous position $\nu_{ij}{}^k$ to a new position $\nu_{ij}{}^{k+1}$. This new position is determined locally by looking at all its neighbors, where a vertex is a neighbor of ν_{ij} if it belongs to one of the quadrangles around ν_{ij}. For all of these (usually 8) neighbors, we determine the vectors from those neighbors to $\nu_{ij}{}^k$, set the lengths of the vectors to the distances of the respective contour points in the tomogram, and average the vectors to obtain the displacement vector. We then move ν_{ij} to its new position $\nu_{ij}{}^{k+1}$ by adding the displacement vector to $\nu_{ij}{}^k$. Before moving a vertex, however, it is checked whether this movement would result in an intersection or in a concave quadrangle. If this is the case, we discard the movement. After each iteration, we measure the distortion as the average difference between the optimal distances of the vertex neighbors in the 2D mesh and in the 3D tomogram. We either stop after a user-defined average difference has been reached or after a maximum number of iterations, also defined by the user. Usually, 2000 iterations are sufficient. For the example shown in Fig. 6d, relaxing the mesh with 2000 iterations took only a few seconds.

5 Results

We applied Approach 1 to the rolled mockup shown in Fig. 2. However, only part of the whole physical role was imaged using computed tomography. To unroll the papyrus from the CT scan, we manually defined contours in only three cross-sections, each with approximately 180 points. This took less than 15 min of manual labor and less than one minute to algorithmically unroll the papyrus. The result of this unrolling, visualized using volume rendering, can be seen in Fig. 4b.

To unfold the mockup package shown in Fig. 3 (right), we applied Approach 2 with two iterations followed by an
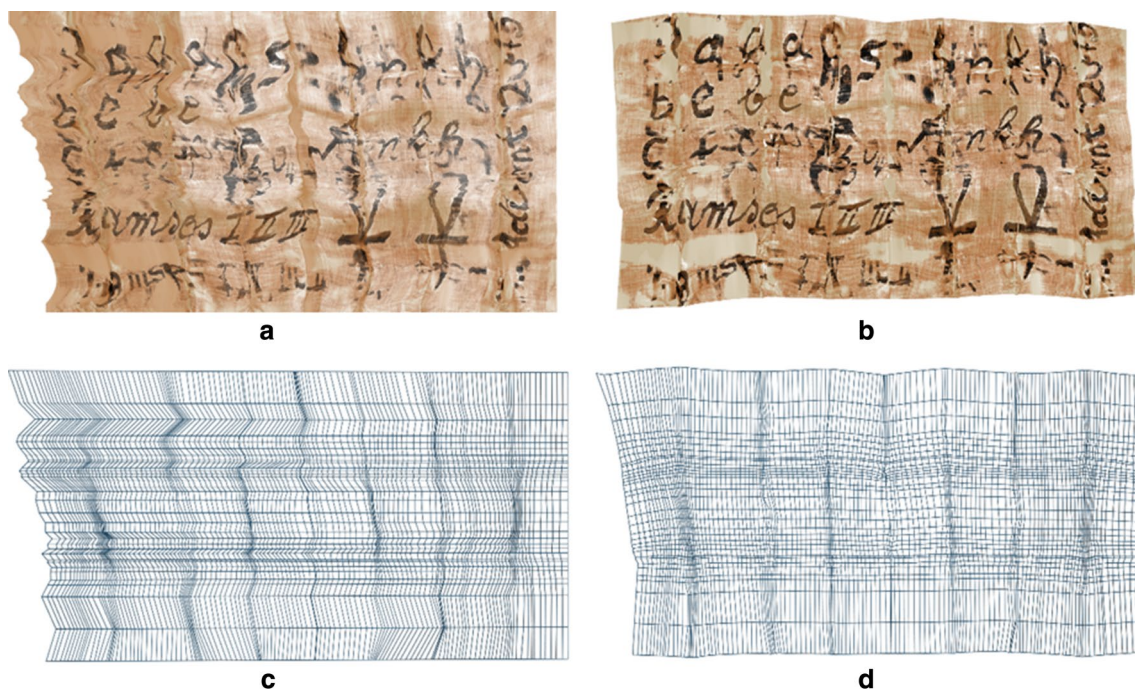
Fig. 6 **a** The virtually unfolded and unrolled papyrus with remaining distortions, indicating that the flattening still needs some improvement. **b** Unfolded papyrus after relaxing the flattened image. **c** Unfolded mesh before relaxation. **d** Unfolded mesh after relaxation

unrolling step using Approach 1. The identification of the direction of the fold line in each unfolding step was done using volume rendering of the current unfolding state of the package. In every iteration of Approach 2, three to five contours were set, each with approximately 200 points. The final unrolling, however, required the definition of many more contours, since the papyrus role was heavily distorted due to the folding. Thus, roughly 35 contours had to be set, each with approximately 200 points. Overall, the whole unfolding took 8–10 h of manual labor. The result of the procedure can be seen in Fig. 6: The image in Fig. 6a shows heavy distortions that can be reduced by relaxing the mesh as described, resulting in the much less distorted image shown in Fig. 6b.

6 Discussion and perspectives

We have successfully demonstrated that with the proposed algorithmic approaches we are able to virtually unroll a papyrus scroll and to unfold a folded package that was folded along perpendicular fold lines (see supplementary material). To the best of our knowledge, this has not been done until now. In combination with the distortion-reducing flattening, we are able to apply the approach also to real ancient samples. In the future, we will work on reducing the manual time needed for unfolding. This can be achieved by automating some of the processing steps. To this end, the described approaches represent a very flexible framework that allows the virtual unrolling and unfolding of even complicated and partly damaged papyrus packages. In addition to the mentioned algorithmic enhancements, we are confident to successfully apply this procedure to a folded papyrus package from ancient Egypt in which less contrasting ink than in our cinnabar study, like iron gall ink, will have been used as identified by XRF.

References

1. D. Neuber, C. Reinhart, Moderne Computertomographie erlaubt virtuelle Entschlüsselung einer antiken, geheimnisvollen Bleischriftrolle der Mandäer, Giesserei-Rundschau 59:5/6 (2012) 133–135. http://www.voeg.at/upload/editor/File/archiv/2012/5-6/Giesserei_5_6_2012.pdf

2. G.H. Barfod, J.M. Larsen, A. Lichtenberger, R. Raja. Revealing text in a complexly rolled silver scroll from Jerash with computed tomography and advanced imaging software. Sci. Rep., 5:17765 (2015)

3. D. Mills, O. Samko, P.L. Rosin, K. Thomas, T. Wess, G.R. Davis. Apocalypto: revealing the unreadable. Proc. SPIE 8506, Developments in X-Ray Tomography VIII, 85060 A (2012). doi:10.1117/12.928917

4. O. Samko, Y.K. Lai, D. Marshall, P. Rosin, Virtual unrolling and information recovery from scanned scrolled historical documents. Pattern Recognit. **47**(1), 248–259 (2014)

5. W. B. Seales, C. S. Parker, M. Segal, E. Tov, P. Shor, Y. Porath. From damage to discovery via virtual unwrapping: reading the scroll from En-Gedi. Sci. Adv. 2:e1601247 (2016)

6. V. Mocella, E. Brun, C. Ferrero, D. Delattre, Revealing letters in rolled Herculaneum papyri by X-ray phase-contrast imaging. Nat. Commun. **6**, 5895 (2015)

7. Pliny the Elder, Naturalis Historia, Liber XXXIII

8. M. Krutzsch. Falttechniken an altägyptischen Handschriften, Archiv für Papyrusforschung, Beiheft 24 (2008) 71–83

9. S. Schaefer, T. McPhail, J. Warren, Image deformation using moving least squares. ACM Trans. Graph (TOG) **25**(3), 533–540 (2006)

10. X. Provot. Deformation constraints in a mass-spring model to describe rigid cloth behavior, Proc. of Graphics Interface '95. Canadian Human-Computer Communications Society, 1995, 147–154 (1995)

11. M.S. Brown, W. Brent Seales. Document restoration using 3D shape: a general deskewing algorithm for arbitrarily warped documents, in Proceedings Eighth IEEE International Conference on Computer Vision (ICCV), vol. 2, IEEE (2001) 367–374

12. B. Lévy, S. Petitjean, N. Ray, J. Maillot, Least squares conformal maps for automatic texture atlas generation. ACM Trans. Graph. (TOG) **21**(3), 362–371 (2002)

13. M.S. Brown, C.J. Pisula. Conformal deskewing of non-planar documents, in "Computer Society Conference on Computer Vision and Pattern Recognition" (CVPR'05), vol. 1, IEEE (2005) 998–1004