Matching and Searching the Dead Sea Scrolls

Taivanbat Badamdorj
School of Engineering
Tel Aviv University

Adiel Ben-Shalom
School of Computer Science
Tel Aviv University

Nachum Dershowitz
School of Computer Science
Tel Aviv University

Abstract—The Dead Sea Scrolls are of immense historical significance. Unfortunately, the scrolls have deteriorated over the millennia and continue to deteriorate since their discovery. Thus, it is of paramount importance to preserve for posterity the current state of the material as best as possible. This goal is being achieved in various ways. One way is via the ongoing digitization efforts of the Israel Antiquities Authority, which complement the older infrared images of plates of fragments done under the auspices of the Palestine Archaeological Museum.

Each parchment or papyrus fragment is being photographed at high resolution on a black felt background and images are being made available to all on the IAA Leon Levy Dead Sea Scrolls Digital Library. At the same time, we are in the midst of an international project with the goal of designing and building algorithmic tools that will relate the different images of scroll fragments with each other and with their transcriptions. As part of that effort, we are in the process of incorporating a deep-learning based segmentation method into the pipeline, which will allow one to manipulate images of the individual fragments themselves. Previous segmentation efforts succeeded in removing most of the shadows from the older images but failed to remove from the foreground of the new images those parts that show the Japanese tissue paper used by conservators to strengthen the edges of the fragment and hold it in place.

We solve the problem of identifying and removing the tissue from the segmented images. This advance dramatically improves the effectiveness of our matching algorithms for searching among the old plates for the location of the newly-digitized fragments. In particular, the improved matching has allowed us to locate two fragments whose positions on the old plates were unknown. The matching algorithm is being incorporated in the new platform and will begin serving scroll scholars in the very near future.

I. INTRODUCTION

The Dead Sea Scrolls (DSS) were discovered during the years 1947–1956 in the caves of the Judean desert. They date to the centuries around the turn of the eras, and hold great historical, religious and linguistic significance. The tens of thousands of parchment and papyrus fragments include the oldest known manuscripts of many works later included in the Hebrew Bible, along with non-canonical and extra-biblical manuscripts in Hebrew, Aramaic, and Greek.

Shortly after the scrolls were discovered, grayscale infrared images were taken of each plate on which they were stored by Najib Anton Albina, the photographer of the Palestine Archaeological Museum (PAM). These often contain dozens of small fragments.

As the reconstructions of scrolls changed, the place of fragments on plates changed and new photo series were taken. Many fragments appear on at least half a dozen images and no complete record is extant for each photo of each fragment.

Presently, high-quality multispectral images are being taken at the Israel Antiquities Authority (IAA) by Shai Halevi, and are offered to scholars and the public on the net (https://www.deadseascrolls.org.il).

An important aspect of the digitization project is also to identify the location of each fragment in older images, especially the PAMs. Once we find a fragment in the PAM plates, one can then assess the level of degradation endured by the fragment in the interim. See Fig. 1.

To this end, a system that segments, matches, and aligns the DSS fragments is under development at Tel Aviv University [3]. It will be incorporated in the ongoing international Scripta Qumranica Electronica project [1], which will link the images of the IAA with the *Qumran-Wörterbuch* produced over the past decades at Die Akademie der Wissenschaften zu Göttingen [2].

The main platform for the project's future operation is what we call the *scrollery*, a standardized environment for collaborative production and presentation of scholarly editions of the scrolls. Besides accessing both the image and the readings of each individual fragment, a scholar will be able to create and store her own new reading and align them with a material reconstruction of the fragments using the images located by this algorithm.

The segmentation algorithm we describe below produces "masks" that – for each fragment – cover the irrelevant parts of the image. So the scrollery can display that part alone, and give the scholar the ability to move it around, turn it over, paste it virtually to other fragments, and so on.

Given an image of a new fragment, the Tel Aviv system first segments it using simple morphological image operations. For the older, grayscale plates, it segments the fragments from their white background using a deep-learning based method to create a "pool" of candidates. Shadows are also removed. Using a cascaded approach, it compares the size and shape of the query image to each candidate. If the shape and scale of the candidate are roughly the same as those of the query, the textures of the two fragments are compared. Later, a dense matching is used to align them. See Fig. 2.

The main drawback of that system is its failure to match the smaller fragments. And since many of the scroll fragments are small (often less than 1 cm^2 in size), this represents a serious limitation. Studying the failures, it became clear that a major cause of the poor results is that the current segmentation algorithm treats the Japanese tissue paper used by conservators to strengthen the edges of the fragment and hold it in place

Fig. 1. Examples of recent color images on the left and an old IR image on the right. The two arrows indicate matches. (All images are courtesy Leon Levy Dead Sea Scrolls Digital Library, Israel Antiquities Authority; color photographer Shai Halevi, infrared by Najib Anton Albina.)
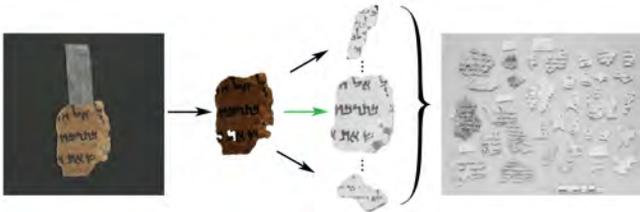

Fig. 2. System overview: We segment the new images, and the old images. Then we compare each fragment on the old plates to the new image using the scale, shape, and texture tests. The correct match is seen in the middle. Images courtesy of Leon Levy Dead Sea Scrolls Digital Library, Israel Antiquities Authority; color photographer Shai Halevi, infrared by Najib Anton Albina.

as foreground.

In this paper, we offer a simple solution that produces better segmentation of the new images. We verify the improvement by measuring the overall effectiveness of the combined system in the task of locating fragments in old plates. By ranking matches, we obtain a robust system that is impervious to large numbers of candidates. The outcomes of the wholesale matching of the new and old images will be incorporated in the toolkit that is currently being developed for both scholars and laypeople. The system has already located several "lost" fragments.

II. PROPOSED METHOD

A. Segmentation of new images

The main contribution of this paper is the devising of a cleaner segmentation method, one that also removes the Japanese tissue showing in many of the images of new fragments (see Fig. 3). This material is used for conservation and preservation purposes to strengthen weak areas in the fragments. Previously, the Japanese tissue was usually left together with the foreground; that is, it was considered a part of the fragment. In some cases for smaller fragments, the tissue took up as much as 40% of area of the segmentation. This led
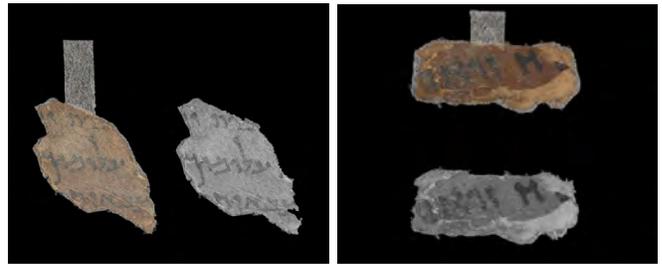

Fig. 3. Two examples of tissue removal. Images courtesy of Leon Levy Dead Sea Scrolls Digital Library, Israel Antiquities Authority, photo: Shai Halevi.

to unreliable results when we compared between the texture of the new and old fragment in the last step of our method.

1) Segmentation: We first convert the image to grayscale and set a low threshold to separate the foreground and background. We need to set a low threshold since the fragments were placed on a dark felt surface. Following this step we find the connected components of the image. We know that the fragment will always be the component that falls in the centre of the image, so we mark a boundary around its connected component. This leaves us with a cropped image of the fragment without other components such as the color patch and ruler.

The next step is to separate the fragment from the background. This is not easy since there are holes in some fragments, which result in black background pixels, and we need to know if a dark pixel is ink and is part of a letter or if it is background seen through a hole. To solve this, we use GrabCut [6], which is a state-of-the-art background removal tool that uses graph cuts to separate object from background. We apply only the first, automatic step of the GrabCut algorithm, and the results appear excellent.

At this stage of the process, we consider the Japanese tissue to be part of the fragment. It is removed in the next step.

2) Removing Japanese tissue: We tried two methods for isolation of the Japanese tissue in the new images.

In the first, we used a simple brightness threshold to create a binary mask for the tissue. The tissue is a light color, so all pixels with red, green, and blue channels with value greater than 108 were selected, followed by a morphological close operation to remove the small holes in the resulting mask. The mask was used to remove the tissue from the fragment image. The fragment was then further smoothed using a morphological open operation to remove long thin lines.

In the second method, we use a two color separation in the Lab space. We first transform the fragment from the RGB space to Lab space. In the latter, the brightness of each color is ignored, and it is easier to distinguish between two color classes. We cluster the absolute values of the transformation into two color classes using k-means. The first class represents the color of the fragment and the other color is the background color, which includes the Japanese tissue.

TABLE I
SUMMARY OF FRAGMENT OVERLAP BETWEEN PLATES: EACH CELL
CONTAINS THE NUMBER OF GROUND TRUTH MATCHES BETWEEN THE OLD
PLATE IN THE GIVEN ROW AND THE NEW PLATE IN THE GIVEN COLUMN.

Plate	4Q57_363	4Q57_382	4Q57_387
M42.002	5	✗	✗
M42.000	✗	8	✗
M42.022	✗	13	✗
M42.162	✗	3	✗
M42.029	✗	✗	18

3) Recto-verso alignment: We take the binary mask of the recto and verso of each fragment, and flip the verso mask left to right. We then find the orientation of the recto and the verso and use this to find the angle between the two orientations. The orientation is a scalar that specifies the angle between the abscissa and the major axis of the ellipse that has the same second-moments as the fragment.

B. Segmentation of old images

Segmenting the older images is a harder task since the images of the old plates have a lower resolution, have multiple fragments on them, and the fragments themselves can have shadows. We use a u-net architecture [5] with a training set of 21 manually labeled plates. Details may be found in [3].

We found experimentally that training multiple models and taking multiple segmentations of the same plate gives better results because some models correctly label fragments that other models may miss entirely. In addition, the models often mislabel fragments that are close to one another as one fragment. Most of the time, they are connected together with a thin line. So we perform a morphological open with a disk structuring element on the segmentations to create new connected components.

C. Locating fragments

1) Scale test: We apply a scale test to determine which fragments are roughly the same size. The new images were taken at an aspect ratio approximately three times larger than that of the old images, so we scale accordingly, and measure the absolute differences in height and width.

2) Shape test: We use Hamming distance as a measure of how similar two fragments are in terms of shape. Given a new image and an old image, we resize the new image to the size of the old one, take the binary masks of both images, and find the Hamming distance between the two masks. For binary masks S'_{i0} and P'_{j0}, their Hamming distance is

$$\|S'_{i0} - P'_{j0}\| = \sum_{\alpha=1}^{t}\sum_{\beta=1}^{p}\left(\frac{s'_{\alpha,\beta}\ \boldsymbol{xor}\ p'_{\alpha,\beta}}{t \cdot p}\right)$$

For simplicity, we refer to the Hamming distance as "shape distance".

3) Texture test: To compare the texture of two fragments, we use SIFT-flow [4]. First we calculate the SIFT descriptor at every pixel of the new and old images. Then the algorithm attempts to find the flow between the two images $\boldsymbol{w}(\boldsymbol{p}) = (u(\boldsymbol{p}), v(\boldsymbol{p}))$. That is, for each pixel $\boldsymbol{p} = (x, y)$ in the new image, find the displacement $\boldsymbol{w}(\boldsymbol{p}) = (u(\boldsymbol{p}), v(\boldsymbol{p}))$ such that $(x+u(\boldsymbol{p}), y+v(\boldsymbol{p}))$ is the same SIFT feature in the old image. This is done by optimizing an energy function using dual-layer loopy belief propagation. The energy function is as follows:

$$E(\boldsymbol{w}) = \sum_{\boldsymbol{p}} \min\{\|s_1(\boldsymbol{p}) - s_2(\boldsymbol{p} + \boldsymbol{w}(\boldsymbol{p}))\|_1, t\} +$$
$$\sum_{\boldsymbol{p}} \eta(|u(\boldsymbol{p}) + v(\boldsymbol{p})|) +$$
$$\sum_{(\boldsymbol{p},\boldsymbol{q})\in\varepsilon} \min\{\alpha|u(\boldsymbol{p}) + u(\boldsymbol{q})|, d\} +$$
$$\min\{\alpha|v(\boldsymbol{p}) + v(\boldsymbol{q})|, d\}$$

where s_1 and s_2 are two SIFT images to match. The set ε contains all the spatial neighborhoods. Here, η is a parameter to ensure that flow vectors are small.

Using SIFT-flow allows us to rank the candidates, and also to do image registration on the correct fragments. An example of image registration is shown in Fig. 4.

Fig. 4. Registration: New image of fragment on left, old image in the middle, and the registration of the two on the right. Images courtesy of Leon Levy Dead Sea Scrolls Digital Library, Israel Antiquities Authority; color photographer Shai Halevi, infrared by Najib Anton Albina.

We use the minimum of the energy required to align two images as a texture similarity metric, from here on referred to as the SIFT-flow distance.

III. EXPERIMENTS

A. Dataset details

We worked with 77 new images of fragments taken from three plates, and 111 old fragments appearing on five different PAM plates. Some new fragments correspond to multiple old fragments. We know that roughly, 4Q57_363 corresponds to PAM plate M42.002, 4Q57_382 corresponds to PAM plates M42.000, M43.022 and M43.162, and 4Q57_387 corresponds to PAM plate M43.029. The overlap between each new and old plate is summarized in Table I.

B. Ranking

In [3] the matching was done with some prior knowledge about where each fragment was originally located. We are

interested in searching without a priori knowledge about the location of the fragment. Thus, the ranking must be robust enough that the correct fragment is near the top of the rankings no matter the number of candidates.

We search for each new fragment in all five of the old plates and rank the returned pool of candidates by their SIFT-flow distance. We also try ranking by shape distance.

We try decreasing the size of the query and candidate images before calculating the SIFT-flow distance. Smaller images allow us to compute the SIFT-flow distance faster. Since we would like to search through thousands of images, we would like the computation to be as fast as possible without sacrificing MRR (mean reciprocal rank) and recall. Initially, we resized both images to 512×512. On a personal computer, it took 20 seconds to calculate the SIFT-flow distance, whereas resizing to 100×100 pixels decreases the calculation time to only 1 second.

The direction of the alignment when using SIFT-flow is also important. We can align the old image of the fragment to the new image or vice versa. Experimentally, we found that trying to align the new image to the old image of the fragment gives subpar results due to the noisy segmentation of the old fragments. Thus, we only align the old images to the new images when calculating SIFT-flow distance.

For this task, we manually labeled 73 correct pairs. After removing double occurrences (some fragments on the new images have multiple matches on the old photos), we ended up with 52 pairs. So, when we find two correct matches for a new image, we only count one of them. We report results for these 52 new fragments.

C. Experimental setup

Let T_{scale}, T_{shape} be thresholds on the scale, and shape distances, respectively, such that candidates with any distance higher than either of the set thresholds are discarded. We set T_{scale} to $400[px]$, and T_{shape} to 0.3

IV. Results

A. Removal

The Lab method performs better in general, as the thresholding method performs poorly for lighter fragments. After our initial experiments with the first five plates, we moved to the Lab method.

B. Ranking

Our main focus was on the smaller fragments. Previously unmatched fragments such as the fragment on the right in Fig. 3 are not only found but are also at the top of the results. We nearly double the MRR, and achieve very high recall. Ranking by SIFT-flow distance gives us by far the highest recall at 1 (!), but we also note that ranking only by shape distance – given that a candidate also meets the size requirements – gives fair results. In fact, from Table II, we see that ranking by shape after removing the Japanese tissue is better than ranking by SIFT-flow distance without removing

TABLE II
RANKING RESULTS, IMAGE SIZES: 512×512. '✗' DENOTES THAT THE JAPANESE TISSUE (JT) WAS REMOVED

JT	Ranking metric	Recall@1	Recall@3	Recall@10	MRR
✓	SIFT-flow	33%	49%	61%	0.440
✓	shape	27%	41%	55%	0.375
✗	SIFT-flow	**72%**	**86%**	**86%**	**0.857**
✗	shape	41%	64%	78%	0.599

the tissue. Here, all images were resized to 512×512 [px] before SIFT-flow distance calculation.

Since removing the Japanese tissue gives us much better results, all further experiments were carried out after the tissue was automatically removed.

We find that, although MRR and recall at 1 decrease when we decrease the size of the images, recall at 10 does not. At 50×50 pixels, we achieve comparable results to the 512×512 pixels case in almost one-hundredth of the time.

From this point on, we resize all images to 100×100 when calculating SIFT-flow distance.

TABLE III
RANKING RESULTS AT VARIOUS IMAGE SIZES, RANKED USING SIFT-FLOW DISTANCE AND AFTER REMOVING TISSUE.

Time [s]	Size [px]	Recall@1	Recall@3	Recall@10	MRR
∼ 20	512×512	72%	86%	86%	0.857
∼ 4	250×250	72%	86%	88%	0.855
∼ 1	100×100	66%	88%	88%	0.831
∼ 0.3	50×50	60%	86%	88%	0.793

V. Practical Application

A. Verification task

After these improvements, we tested the system in a more practical setting. The task is to find the location of fragments on the PAM plates given their recently acquired images. For testing purposes, we know a priori which set of PAM plates contain the fragments that we would like to search for. However, we do not know exactly on which PAM plate each fragment belongs. We only know that it is on one of them.

We have two new plates, P106 and P107, with 38 individual images, 32 of which are of well-preserved fragments. (We don't consider fragments that are poorly preserved to test our system.) Thanks to prior work of experts, we knew that each fragment appears on one of the following PAM plates: M40613, M40601, M40618, M41455, M41453, M41664, M42396, M41974, M41516, M43293, M43294, and M42395. These 12 plates provided a total of around 500 candidate fragments.

An example of a correct match from the system can be seen in Fig. 5. The system highlights the fragment that it has found to make it easy for scholars to verify the results.

Considering the success of the matching, we increased the number of plates to 350 (provided by the IAA), showing the robustness of ranking. These 350 plates contain a total
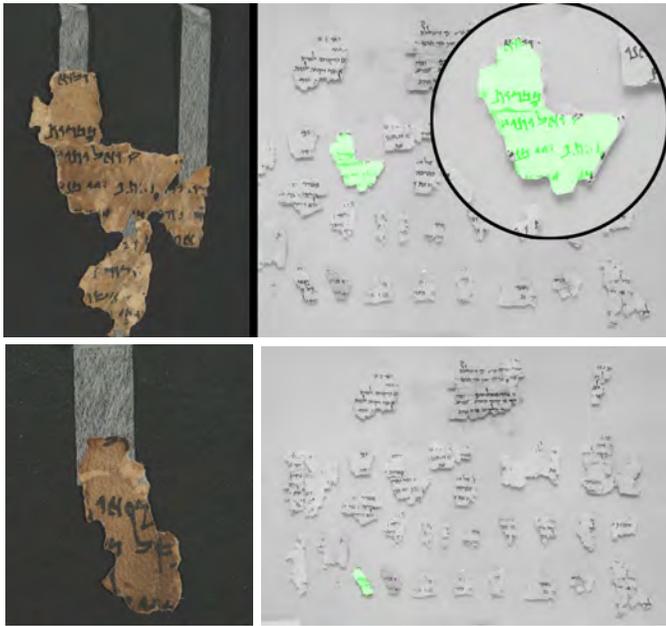
Fig. 5. Examples of new fragments located on an old PAM plate: The system highlights the location of the fragment in green. Inset on the upper plate is the enlarged match. Note that the matching new image has an extra piece connected with tissue paper. Images courtesy of Leon Levy Dead Sea Scrolls Digital Library, Israel Antiquities Authority; color photographer Shai Halevi, infrared by Najib Anton Albina.

TABLE IV
VERIFICATION OF FRAGMENTS FROM P106 AND P107.

Number of candidates	Recall@1	Recall@3	Recall@10	MRR
∼ 500	87%	90%	90%	0.890
∼ 6000	84%	84%	84%	0.848

of over 6000 candidate images for matching. Because of the morphological open operation, some candidates are spurious, not actual fragments. We overcompensate with the number of possible fragments, and show that the system will still find the most correct match.

B. Results

We were able to correctly match and locate 29 out of the 32 well-preserved fragments on plates P106 and P107. As can be seen on the second line of Table IV, the system is remarkably impervious to the large increase in the number of candidates. It also doesn't matter that many of the candidates are spurious.

More importantly, we also located two fragments in previously unknown locations – fragment 3 from P106 was found in plate M40626, and fragment 19 from P107 was found in plate M40964. The fragments are shown in Fig. 6.

VI. CONCLUSION

We have demonstrated the practical effectiveness of the proposed system for searching visually through the Dead Sea Scroll repository to locate manuscript fragments in older
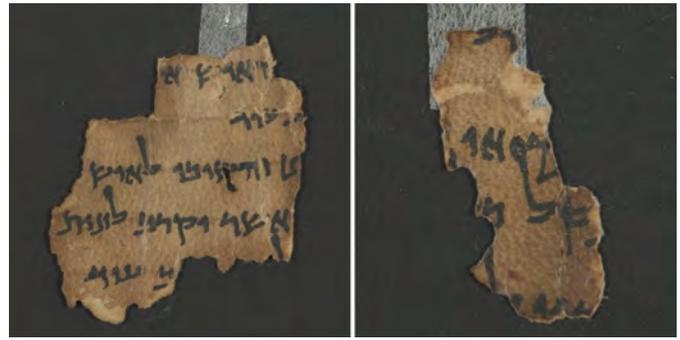

Fig. 6. Two newly located fragments: P106, fragment 3 (right) and P107, fragment 19 (left) were found on previously unlisted plates. Images courtesy of The Leon Levy Dead Sea Scrolls Digital Library; Israel Antiquities Authority, photo: Shai Halevi.

images of plates of fragments. We have already found the locations of two newly-imaged fragments. In upcoming work, in coöperation with the IAA, we plan to use the system we have developed to search for all of the many thousands of newly-imaged fragments among the old infrared images, thereby creating a full and accurate catalog for scholars, linking the new high-quality images with the older plates, which sometimes show fragments in a better state of preservation.

REFERENCES

[1] B. Brown-deVost, "Scripta Qumranica Electronica (2016–2021)," Hebrew Bible and Ancient Israel (HeBAI) 5 (2017): 307–315.

[2] R. G. Kratz, A. Steudel, and I. Kottsieper, *Hebräisches und aramäisches Wörterbuch zu den Texten vom Toten Meer: Einschliesslich der Manuskripte aus der Kairoer Geniza [Hebrew and Aramaic Dictionary of the Dead Sea Scrolls – Including the Manuscripts from the Cairo Geniza]*, volume 1, de Gruyter, Berlin/Boston, 2017.

[3] G. Levi, P. Nisnevich, A. Ben-Shalom, N. Dershowitz, and L. Wolf, "A Method for Segmentation, Matching and Alignment of Dead Sea Scrolls", Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, CA, pp. 208–217., March 2018.

[4] C. Liu, J. Yuen, and A. Torralba, "SIFT Flow: Dense Correspondence across Scenes and Its Applications," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 5, pp. 978–994, May 2011.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, vol. 9351, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 234–241.

[6] C. Rother, V. Kolmogorov, and A. Blake, GrabCut: Interactive Foreground Extraction Using Iterated Graph Cuts, ACM Trans. Graph., vol. 23, no. 3, pp. 309–314, Aug. 2004.