Global and Local Shape Prior for Variational Segmentation of Degraded Historical Characters

Itay Bar-Yosef,
Alik Mokeichev,
Klara Kedem

Department of Computer Science,Ben-Gurion university, Israel

itaybar@cs.bgu.ac.il
mokeiche@cs.bgu.ac.il
klara@cs.bgu.ac.il

Uri Erlich

Department of Jewish Thought, Ben-Gurion university, Israel

ehrlich@bgu.ac.il

Itshak Dinstein

Department of Electrical and Computer Engineering,Ben-Gurion university, Israel

dinstein@ee.bgu.ac.il

Abstract

We propose a variational method for model based segmentation of highly degraded gray scale images of historical documents. Given a training set of characters (of a certain letter), we construct a small set of shape models that cover most of the training set's shape variance. For each gray scale image of a respective degraded character, we construct a custom made shape prior using those fragments of the shape models that best fit the character's boundaries. Therefore, we are not limited to any particular shape in the shape model set. Experiments show that our method achieves very accurate results in segmentation of highly degraded characters. When compared with manual segmentation, the average distance between the the boundaries of respective segmented characters was 0.8 *pixels (the average size of the characters is* $70 * 70$ *pixels).*

Keywords: Segmentation, historical documents, level set, shape prior

1. Introduction

Much effort has been devoted in the last few years to the digitization of paper collections in order to archive them as digital documents. A common need in libraries and archives is to improve the readability of historical documents. This has high cultural values as well as scientific ones, e.g., enhancing degraded documents, improving OCR accuracy and aiding paleographic researchers for whom the accuracy of the segmentation process is of high research value. However, segmentation of historical documents is difficult due to varying contrast, smudges, faded ink, and the presence of bleed-through text. In order to overcome these difficulties, binarization algorithms designed especially for historical documents have been proposed in the last few years [1],[2]. Although for some cases impressive results were obtained, these methods often fail when dealing with extremely degraded documents. The main reason is that most of the binarization methods are based on global or local statistics derived from the gray scale image. These statistics alone are not sufficient in complex cases, where characters are broken and/or partially erased as shown in Figure 1.

Figure 1. Example of highly degraded text.

A small number of works have been reported on restoration of degraded historical documents. Very few among them considered gray scale images. Droettboom [3] and Hobby [4] dealt with binary documents. Droettboom et al. [3] proposed a method based on graph combinatorics to merge broken characters of historical documents. The goal of their algorithm was to find an optimal way to join connected components on a given page, that maximizes the mean confidence of all characters. Hobby et al. [4] proposed to improve the quality of degraded text images by image matching techniques. Similar symbols were clustered, and a prototype of each cluster was generated to replace the cluster symbols. Active contour model for restoration of degraded gray scale characters was developed by Allier et al. [5]. The authors incorporated a single shape model to a parametric active contour based on a GVF representation [6] of the image and of the shape prior. Although impressive results have been reported, their approach does not provide satisfying results when applied on highly degraded text image with large variability of the characters shapes.

In this paper we present a variational approach for accurate segmentation of highly degraded characters. Our main contribution is a novel shape prior that is customized to the character's gray scale image and is not limited to any particular shape of the training set. We finally integrate this shape prior into a variational model.

The rest of the paper is organized as follows: Section 2 gives a short introduction on active contours and level set methods. In Section 3 we present our shape prior and describe the prior construction process. The variational formulation of the shape prior segmentation is outlined in Section 3.4, and our experimental results are presented in Section 4. We conclude our work and outline future work in Section 5.

2. Active Contours and Level-Set framework

In the last decade, variational methods have gained much popularity in the computer vision community and affected considerably the approach to image segmentation. First introduced by Kass et al. [7], parametric active contours are dynamic curves which are attracted to image boundaries based on a flow that minimizes a certain energy functional. This flow is derived by calculating the Euler-Lagrange equation of the energy functional and gradient descent methods are then used to evolve the active contour. Casseles et al. [8] introduced the *geodesic active contours* and showed that Kass's model can be formulated as finding a curve C with minimal geodesic length, i.e.:

$$min \int g(|\nabla I(C(s))|)|C'(s)|ds \qquad (1)$$

where g is a non decreasing function such that g(x) $\rightarrow 0$ as x $\rightarrow \infty$, and g(x) $\rightarrow 1$ as x $\rightarrow 0$. The steady state of the active contour can be reached by evolving each contour point according to the following flow:

$$\frac{\partial C}{\partial t} = g\kappa\vec{N} - (\nabla g \cdot \vec{N})\vec{N} \qquad (2)$$

where κ is the contour curvature, and $\vec{N}$ is the unit normal to the curve. The first term of Eq.2 is a smoothing term, which controls the speed of the evolving contour and slows it down in places with high gradient (where g $\rightarrow 0$). The second term is an attraction force, which attracts the contour to high gradient boundaries.

The authors [8] implemented their model based on the *level set* method, a numerical technique for front propagation proposed by Osher and Sethian [9]. The main idea of the level set method is to embed the evolving contour C as the zero level set of an implicit function Φ (usually the signed distance function), defined in a higher dimension C(t)=$\{$(x,y)$|\Phi$(t,x,y)=0$\}$. Then, the curve's evolution equation $\frac{\partial C}{\partial t} = F\vec{N}$ can be re-written in a level set formulation $\frac{\partial \Phi}{\partial t} = F|\nabla\Phi|$, where F is the evolution speed function. By this, topological changes of the evolving contour are handled naturally and efficient numerical techniques can be applied [9].

Although the geodesic active contour has been widely used in different domains, it has two major drawbacks. First, the evolution is based on local information, therefore it is very sensitive to local minima. Second, when the boundaries have gaps or weak edges, the attraction term in Eq.2 is often not strong enough to avoid edge-linkage. In order to solve these problems, some authors proposed to integrate shape prior knowledge into the geodesic active contour. Chen et al. [10] proposed to use the average shape of a training set as a shape prior. A statistical approach is presented in the work of Leventon et al. [11], where the shapes of the training set are represented by their signed distance function and a shape prior is constructed from a reduced dimension space. For a recent review on incorporation of shape priors to active contours, see [12]. Most of the prior based approaches treat the shape prior as a global term and are limited to a predefined set of permitted shapes. The main contribution of this paper is that we choose from each shape model only the fragments which fit the image best, and create a composite shape prior customized to the character's gray scale image. Therefore, we are not limited to any particular shape of the training set.

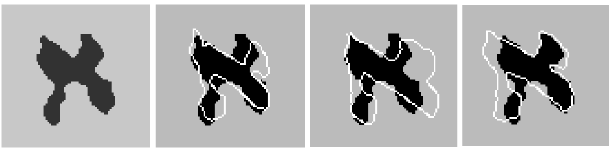
In the following section, we present our shape prior model and describe in detail its construction process.

3. Definition and Construction of the Shape Prior

Throughout this paper we assume that the main differences in character shapes are due to the writer's writing style, and that orientation and scale differences are minor and can be neglected. Therefore, we estimate directly the pose and identity of the character by using simple template matching. In the rest of this paper, we assume that the pose and the identity of each character are given.

Given a gray scale image of a character and a small set of shape models, we align each shape to the gray scale image according to a simple template matching process based on *normalized cross correlation* (each model is aligned according to their maximum correlation lags). Naturally, each model fits the gray scale image differently (see example in Figure 2). We exploit this fact to create a new shape prior constructed of the models' fragments that best fit to the character's image. The main advantage of our approach is that our shape prior captures local information for each character's image as opposed to traditional approaches where the shape prior is used as a global constraint.

Throughout this paper, we use the letter "Aleph" for demonstration since it has more intricate shape in the Hebrew alphabet.



(a) (b) (c) (d)

Figure 2. (a) Gray scale image of the Hebrew letter "Aleph". (b-d) Different models aligned based on the maximum normalized cross-correlation.

3.1. Model Set Creation

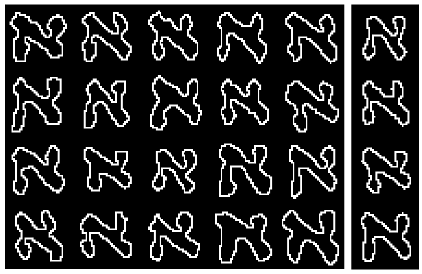
Given a training set of n characters of the same letter, we construct a small set of shape models (for computation efficiency) that cover most of the shape's variance. The shape models are constructed as follows:

1. Compute the signed distance functions (SDF) of the training characters, and rearranged each SDF as a column vector. The SDF is defined as negative inside the character, and positive outside of it.
2. Apply principal component analysis (PCA) to reduce the SDF vectors dimensionality. In our experiments, SDF vectors of 4900 components ($70 * 70$ images) were represented with 95% accuracy by 10 principal components.
3. Cluster the n 10-dimensional vectors into k clusters.
4. Transform the k cluster centers back to the SDF's original dimension.
5. Use the zero level sets of the k obtained SDFs as shape models.

Figure 3(a) shows a set of twenty characters from which we construct four models through the PCA clustering. Figure 3(b) shows the constructed models.

3.2. Confidence Map Construction

Given a gray scale image and the small set of shape models, we now define a mechanism that provides local information on the degree of fitting to the gray scale image. For a given gray scale image and an aligned shape model, we create a *confidence map* in two steps. First, we superimpose the model on the gray scale image. For each point on the model's boundary, we place a small window around it (Figure 4(a)) and calculate the normalized correlation between the corresponding portion of the model and that of the image (Figure 4(b)). The calculated value is then assigned to the corresponding contour point (Figure 4(c)). In order to compare between the different models locally, we propagate the score from each model's boundary to the rest of the image domain. For that purpose, we apply the fast image inpainting algorithm proposed by Oliveira et al. [13]. Their algorithm iteratively updates unknown pixel values with the average value of their known surrounding neighbors. Propagating the fitting score from the boundary, results in a confidence map which measures at each pixel, the local fitting to the shape model. For the constructed set of k models described in the previous section, we define s_i to be the *confidence map* of the i^{th} model. Figure 4(d) presents the confidence map of the model shown in white in Figure 4(a).



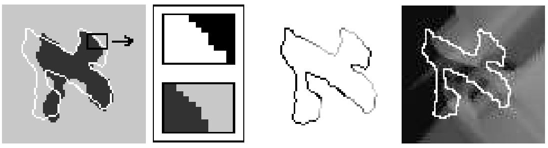
(a) (b)

Figure 3. (a) A training set of 20 characters from which we construct a compact set of 4 models. (b) The constructed models

In the following section, we describe our shape representation model and explain how we use the confidence maps of the different models to construct the customized shape prior.

3.3. Shape Representation

After calculating the confidence map of the constructed models, we determine for each pixel in the image the most suitable model according to the one with the highest confidence value. We need a representation that will enable us to converge to the desired image boundary. An appropriate choice for that purpose is the Gradient Vector Flow (GVF), proposed by Xu et al. [6]. Given an input image I, the idea is to smoothly diffuse the image's gradient vectors (∇I) to distinct parts of the image. The vector



(a) (b) (c) (d)

Figure 4. Calculation of the confidence map. (a) A small rectangular window (in black) is placed on each boundary point of the model. (b) Normalized correlation is then calculated on the extracted image and model region. (c) Scores on the model's boundary. (d) The resulting confidence map s_i after propagating the information from the model's boundary (the model's boundary is superimposed in white). Both in (c) and (d) bright regions indicate high fitting score, while dark regions indicates poor fitting.

flow $\vec{\nu}$ which minimizes the following functional is called the gradient vector flow of the image.

$$\int\int \mu(\nabla\nu^2) + |\nabla I|^2|\nu - \nabla I|^2 dxdy \qquad (3)$$

The GVF $\vec{\nu}$ have the following nice property: In places where the image gradient ∇I is high, $\vec{\nu}$ will tend to be equal to that gradient. In places with low gradient, $\vec{\nu}$ will smoothly vary toward the boundaries with high gradient. As a result, $\vec{\nu}$ determines for every point in the image how to travel in order to reach the nearest boundary.

Denote by $\vec{V_i}$ the GVF of the i^{th} model, and denote by $\vec{P}$ the composite shape prior for the character. For each pixel (x, y) in the image, the model with the highest confidence value determines $\vec{P}(x, y)$:

$$\vec{P}(x,y) = \vec{V_k}(x,y), \;\; k = argmax_i(s_i(x,y)) \qquad (4)$$

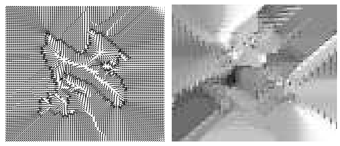
In a similar way, we define the *prior confidence map* $\eta(x, y)$, which measures the reliability of the local shape prior at each pixel:

$$\eta(x,y) = s_k(x,y), \;\; k = argmax_i(s_i(x,y)) \qquad (5)$$

Figure 5 shows the shape prior $\vec{P}$ and the confidence map $\eta(x, y)$ based on the gray scale image and the three models shown in Figure 2. In the following section, we integrate the shape prior $\vec{P}$ into the geodesic active contour.

3.4. Incorporating the Shape Prior into the Active Contour

We formulate a variational segmentation model which is based on our composite shape prior $\vec{P}$ and the geodesic active contour described in Section 2. The shape prior $\vec{P}$ is a vector flow which should be followed in order to reach



(a) (b)

Figure 5. (a) The shape prior $\vec{P}$, calculated based on the gray scale image and the three models in Figure 2. (b) The prior's confidence map.

the closest boundary. A natural choice for the evolution equation based on the shape prior is:

$$\frac{\partial C}{\partial t} = g\kappa\vec{N} - (\vec{P}\cdot\vec{N})\vec{N} \qquad (6)$$

This flow is the classical geodesic flow. The right term of the equation will be maximized when the prior $\vec{P}$ and the inward normal have the same direction. The first term acts as a regularization term which smooths the evolving contour. We use the variational level set formulation as presented in Chen et al. [10]. The level set equations for the shape flow in Eq.6 and the geodesic active contour flow in Eq.2, are:

$$\frac{\partial\Phi}{\partial t} = (g\kappa|\nabla\Phi| - \vec{P}\cdot\nabla\Phi)\delta(\Phi) \qquad (7)$$

$$\frac{\partial\Phi}{\partial t} = (g\kappa|\nabla\Phi| - \nabla g\cdot\nabla\Phi)\delta(\Phi) \qquad (8)$$

where $\delta(\cdot)$ is a smoothed Dirac function defined as in [10] and $\delta(\Phi)$ is the contour measure on the zero level set. κ, the curvature of the zero level set, and the edge indicator function g, are defined as:

$$\kappa = div(\frac{\nabla\Phi}{|\nabla\Phi|}), \;\; g = \frac{1}{1+|\nabla I|^2} \qquad (9)$$

The incorporation of the shape flow with the geodesic flow is based on the following observations:

- The prior confidence map η (see Section 3.3) determines whether we should follow the shape prior, i.e., the shape prior should be used only in regions where the η value is large.
- Regions with low η value indicate that none of the shape models was fitted to this region. In such cases there are two options, either that the real boundary passes in a different location, and information from the image is needed to attract the active contour to the correct place, thus the geodesic active contour should be dominant, or that the area is damaged and information from the image is not useful. In the latter case we need a mechanism which will constrain

the contour's evolution in order to remain as close as possible to a typical shape.

For this purpose we define a global shape constraint $\gamma(x, y)$ which acts as an adaptive balloon force, and will constrain the evolution of the active contour by controlling the amount of deformation locally. After aligning the shape models (represented by their *SDF*), we define $\mu(x, y)$ and $\sigma(x, y)$ to be the average and standard deviation of the aligned *SDF* respectively. Regions with high σ values are allowed for more deformations, while regions with small σ values are more likely to remain close to the average μ. The function which limits the evolution of the active contour is defined as:

$$\gamma(x,y) = \begin{cases} c \;, & if\ \mu(x,y) + 2 \times \sigma(x,y) > 0 \\ -c \;, & if\ \mu(x,y) - 2 \times \sigma(x,y) < 0 \\ g \;, & otherwise \end{cases} \tag{10}$$

where g is defined in Eq.9. The meaning of this function is as follows. γ creates a narrow band of permitted deformations by adding a strong balloon force outside the band (the constant c) which contracts the active contour, and a strong negative balloon force which limits the active contour by inflating it back when it reaches the limit of the allowed deformations. The final evolution equation composed by both the shape prior flow and the geodesic flow is:

$$\frac{\partial \Phi}{\partial t} = [\gamma\kappa|\nabla\Phi| - (1-\eta)\nabla g \cdot \nabla\Phi - \eta(\vec{P} \cdot \nabla\Phi)]\delta(\Phi) \tag{11}$$

To demonstrate our approach, we corrupted the image shown in Figure 6 by deleting part of it and adding an extra patch to its upper-right region. As can be seen, our model successfully segmented the character by attracting to the preserved boundaries while keeping the final contour close to a typical shape of the character.

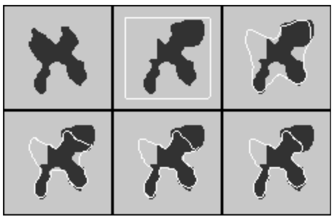


Figure 6. From left to right, top to bottom: Synthetic character; Some part was erased and an extra patch was added to the character; Four different stages of the curve evolution.

4. Experimental Results

We conducted our experiments on a set of four degraded documents from an antique Jewish prayer book composed by the famous scholar *Rav Seadia Gaon*, between the eleventh and the thirteenth century. We tested our segmentation model based on ten Hebrew character types. For each type, we extracted a set of fifteen characters for the training set, and a set of ten degraded characters for the test set. All the extracted characters were segmented manually in order to create the training set and to evaluate the segmentation results.

First, the GVF field ($\vec{V_i}$), and its SDF were calculated for each model. Then, the model construction process described in Section 3.1 was applied. In all our experiments, we used only four models for the shape prior construction. Last, we calculated the function γ based on the SDF's of the training set models [1].

For each of the 100 degraded characters, we calculated $\vec{P}$ and η and then initialized the active contour around the character and evolved it according to Eq.11. In order to evaluate our segmentation results, each of the degraded characters was also segmented manually. Figure 7 shows several examples of degraded character images along with the segmentation results with our shape prior. Segmentation with the geodesic active contour without prior (Eq.8) are shown in the third column, and the manual segmentation is shown in the last one. It is obvious that without incorporating the shape prior, accurate segmentation is impossible.

We evaluated our segmentation results based on the average distance of each segmented character's boundary from the boundary of the manual segmented character. The results are summarized in Table 1, where the first four entries display the results for the images in Figure 7, and the last entry displays the average result for the set of 100 degraded characters.

5. Conclusions and Future Work

In this paper we propose a novel shape prior which is incorporated in a variational level set formulation. We first create a set of shape models based on clustering of the PCA representation of the training set SDF's. Then, we compose a shape prior $\vec{P}$ that adopts from each model only the fragments in which fit best the image. Combined with an adaptive balloon force (γ), the prior terms are incorporated into a level set equation which shows very accurate results in segmenting highly degraded characters. The average distance between a segmented character using our model and the respective manually segmented one was 0.8 pixels.

Although in this paper we assume to know the true identity of the character, we believe that this assumption is not necessary. An accurate shape prior can be constructed by using shape models from different letters, since we ex-

[1] The calculations of $\vec{V_i}$, γ and the model construction process (PCA analysis of the SDF's) were off line

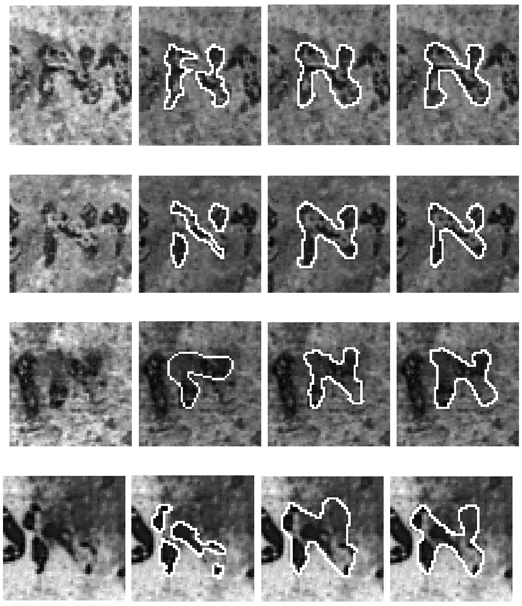


Figure 7. From left to right: Degraded character, segmentation using geodesic active contour, segmentation with our model, and manual segmentation. Similar results were obtained for the other characters. The images are numbered from top to bottom.

ploit only the models' fragments that best fit to the image.

Table 1. Evaluation of our segmentation model based on the average distance (in pixels) between the boundaries of the manual segmentation and of the respective segmented characters (the average size of the characters is $70 * 70$ pixels). The first 4 entries display the result of the images in Figure 7. The last entry displays the average in pixels of the entire set of 100 degraded characters.

Image No.1	0.3
Image No.2	0.6
Image No.3	0.8
Image No.4	1.4
Average of 100 images	0.8

6. Acknowledgments

This work was partially supported by the Lynn and William Frankel Center for Computer Sciences, the Israel Ministry of Science and Technology, and by the Paul Ivanier Center for Robotics and Production Management at Ben-Gurion University, Israel.

7. References

References

[1] B. Gatos, I. Pratikakis, J. S. Perantonis, "Adaptive degraded document image binarization", *Pattern Recognition*, (39), pp.317-327, 2006

[2] I. Bar-Yosef, "Input sensitive thresholding for ancient Hebrew manuscript". *Pattern Recognition Letters* 26(8), pp. 1168-1173, 2005

[3] M. Droettboom, "Correcting broken characters in the recognition of historical printed documents", *Proceedings of the 3rd ACM/IEEE-CS joint conference on Digital libraries*, pp. 364-366, 2003.

[4] J. D. Hobby, T. K. Ho. "Enhancing Degraded Document Images via Bitmap Clustering and Averaging", *Fourth International Conference Document Analysis and Recognition*, pp. 394-400, 1997.

[5] B. Allier , N. Bali and H. Emptoz, "Automatic accurate broken character restoration for patrimonial documents". *International Journal on Document Analysis and Recognition*, (8), pp.246-261, 2006

[6] C. Xu and J. L. Prince, "Snakes, Shapes, and Gradient Vector Flow", *IEEE Transactions on Image Processing*, (7), pp.359-369, 1998

[7] A. kass, M. witkin, D. terzopoulos "Snakes: Active contour models", *International Journal of Computer Vision*, (1), pp.321–332, 1988

[8] V. Caselles, and R. Kimmel, and G. Sapiro, "Geodesic Active Contours", *International Journal of Computer Vision*, (22), pp. 61–79, 1997

[9] S.Osher, J.A. Sethian, "Fronts propagating with curvature dependent speed: algorithms based on hamilton-jacobi formulation", *Journal of computational physics*, (79) ,pp.12–49, 1988

[10] Y. Chen , H.D. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K.S. Gopinath , R.W. Briggs, and E.A. Geiser, "Using prior shapes in geometric active contours in a variational framework", *International Journal of Computer Vision*, (50), pp.315–328, 2004

[11] M.E. Leventon, W.E.L. Grimson, O. Faugeras, "Statistical shape influence in geodesic active contours", *CVPR*, pp.316–323,2000

[12] D. Cremers, M. Rousson, and R. Deriche,"A review of statistical approaches to level set segmentation: Integrating color, texture, motion and shape, *International Journal of Computer Vision*, (72), pp.195–215, 2007.

[13] M. Oliveira, B. Bowen, R. McKenna, and Y. Chang, "Fast digital image inpainting". *Proceedings of the International Conference on Visualization, Imaging and Image Processing*, pp. 261–266, 2001.