Pattern Recognition 42 (2009) 3348 – 3354

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Adaptive shape prior for recognition and variational segmentation of degraded historical characters

Itay Bar-Yosef[a,*], Alik Mokeichev[a], Klara Kedem[a], Itshak Dinstein[b], Uri Ehrlich[c]

[a]Department of Computer Science, Ben-Gurion University, Israel
[b]Department of Electrical and Computer Engineering, Ben-Gurion University, Israel
[c]Department of Jewish thought, Ben-Gurion University, Israel

ARTICLE INFO

Article history:
Received 16 August 2008
Accepted 15 October 2008

Keywords:
Segmentation
Degraded character recognition
Historical documents
Level set
Shape prior

ABSTRACT

We propose a variational method for model based segmentation of gray-scale images of highly degraded historical documents.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 boundary. Therefore, we are not limited to any particular shape in the shape model set. In addition, we demonstrate the application of our shape prior to degraded character recognition. Experiments show that our method achieves very accurate results both in segmentation of highly degraded characters and both in recognition. When compared with manual segmentation, the average distance between the boundaries of respective segmented characters was 0.8 pixels (the average size of the characters was $70 * 70$ pixels).

1. Introduction

Much effort has been devoted in the last few years to the digitization of books and documents in order to archive them in digital form. A common need in libraries and archives is to improve the readability of historical documents. This has high cultural and scientific values, e.g., enhancing degraded documents, improving OCR accuracy and facilitating paleographic researchers for which 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 However, segmentation of historical documents is difficult due to varying contrast, smudges, faded ink, and the presence of bleed-through text. In order to overcome these difficulties, binarization algorithms designed especially for historical documents have been proposed in the last few years [1,2].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 Fig. 1.

A small number of scientific papers have dealt with 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. An 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 gradient vector flow (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 adaptive 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 integrate the shape prior into a variational segmentation model and also demonstrate how the proposed shape prior can significantly improve degraded character recognition.

* Corresponding author.
E-mail addresses: itaybar@cs.bgu.ac.il (I. Bar-Yosef), mokeiche@cs.bgu.ac.il (A. Mokeichev), klara@cs.bgu.ac.il (K. Kedem), dinstein@ee.bgu.ac.il (I. Dinstein), ehrlich@exchange.bgu.ac.il (U. Ehrlich).

0031-3203/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2008.10.005

Fig. 1. Example of highly degraded text.

The rest of the paper is organized as follows: In Section 2 we present our shape prior and describe the prior construction process. Section 3 demonstrate the application of our shape prior with respect to recognition. Section 4 gives a short introduction on active contours and level set methods, and outlines the variational formulation of the shape prior segmentation. Experimental results are presented in Section 5, and in Section 6 we conclude our work and outline the future work.

2. Definition and construction of the shape prior

Accurate binarization of highly degraded characters is often practically impossible without using any prior knowledge of the character's shape or identity. A simple method to obtain the degraded character's identity is matching based on normalized cross-correlation (NCC) which has been proven to be robust against vast gray-scale degradations. The main drawbacks of the NCC is its weakness against local geometric distortions and affine-transformations [7]. However, as commonly seen in historical documents, the main differences in character shapes are due to the writer's writing style, thus the main challenge is to overcome these local differences while affine-distortions (orientation and scale) are minor and can be neglected. We note that affine-invariance can be achieved by using more sophisticated approaches such as the one proposed by Wakahara et al. [7].

In the following, we propose a novel shape model which is adapted to the character's boundaries very accurately. The shape prior construction process described below assumes a given gray-scale character image and a small set of shape models corresponding to the identity of that character. In Section 3 we relax the assumption of known character's identity and show how to utilize the shape prior for character recognition. The creation of the shape models used for prior construction is further discussed in Section 2.3.

Given a gray-scale image of a certain character and a set of N respective shape models, we align each shape to the gray-scale image based on NCC (each model is aligned according to their maximum correlation lags). Naturally, each model fits the gray-scale image differently (see example in Fig. 2). We exploit this fact to create a new shape model constructed of the models' fragments that best fit to the character's boundary. The main advantage of our approach is that our shape model captures local information of the gray-scale character image with respect to the given models. We further use this shape model for both accurate segmentation and recognition of highly degraded characters (see Sections 3 and 4).

2.1. Confidence map construction

Given a gray-scale image and a set of N respective shape models, we define a mechanism that provides local information on the degree of fitting of each model to the gray-scale image. We will further utilize this information for local comparison between the models in order to find those fragments that best fit to the 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. Around each point along model's boundary, we place a small window (Fig. 3(a)) and calculate the normalized correlation between the corresponding portion of the model and that of the image (Fig. 3(b)). The calculated value is then assigned to the corresponding contour point (Fig. 3(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. [8]. 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 model. For the set of N shape models, we define s_i to be the *confidence map* of the ith model. Fig. 3(d) presents the confidence map of the shape model shown in white in Fig. 3(a). 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.

2.2. Shape representation

After calculating the confidence map of the N respective shape models, we determine for each pixel in the image the most suitable model, i.e., the one with the highest confidence value. The shape prior is composed from pixels of the models such that each model contributes only pixels that best fit the raw data.

For this purpose we use the signed distance function (SDF) for shape representation. This representation has gained much popularity in the recent years, mainly in the field of variational segmentation. We will elaborate on the advantages of this representation in Section 4. The idea is to represent the shape's contour C by embedding it in a higher dimension level set function Φ, as follows:

$$\Phi(x) = \begin{cases} -D(x,C), & x \text{ is inside } C \\ D(x,C), & x \text{ is outside } C \\ 0, & x \in C \end{cases} \tag{1}$$

where $D(x,C)$ denotes the Euclidean distance from x to the closet point on C. The contour C can be reconstructed from such representation by taking its zero level set $C = \{x|\Phi(x) = 0\}$. An example for SDF is shown in Fig. 4(b). Denote by Φ_i the SDF of the ith model, and denote by Φ_p the composite shape prior for the character. For each pixel (x,y) in the image, the model with the highest confidence value determines Φ_p:

$$\Phi_p(x,y) = \Phi_k(x,y), \quad k = \underset{i}{\operatorname{argmax}}(s_i(x,y)) \tag{2}$$

In a similar way to Φ_p, we define the *prior confidence map* η, which measures the reliability of the composite shape prior at each pixel:

$$\eta(x,y) = s_k(x,y), \quad k = \underset{i}{\operatorname{argmax}}(s_i(x,y)) \tag{3}$$

Fig. 5 shows the shape prior Φ_p and the confidence map $\eta(x,y)$ based on the gray-scale image and the three shape models shown in Fig. 2. Notice the high correspondence between the shape prior fitting to the gray-scale image, and the confidence map. In Section 4