A - 17

Proceedings of the
IASTED International Conferences

Signal and Image Processing (SIP'99)

Editor: N.M. Namazi

October 18-21, 1999
Nassau, The Bahamas

A Publication of the International Association of Science
and Technology for Development - IASTED

IASTED/ACTA Press
Anaheim ♦ Calglary ♦ Zürich

ISBN: 0-88986-267-2

Proceedings of the IASTED International Conference
Signal and Image Processing (SIP'99)
October 18-21, 1999, Nassau, Bahamas

JOINT BLIND RESTORATION AND SEGMENTATION OF BLURRED TEXT CHARACTERS

LUIGI BEDINI and ANNA TONAZZINI

Istituto di Elaborazione della Informazione
Consiglio Nazionale delle Ricerche
Via S. Maria, 46
I-56126 PISA (Italy)

ABSTRACT

Character segmentation of damaged printed texts is a very critic task, especially when the degradation causes the characters to touch and merge one another. With particular reference to ancient printed texts, we model the degradation as a unknown space-variant blur operator and try to jointly estimate the blur parameter and recover the undegraded image. Since the latter can be considered as a two-level image, we propose to integrate techniques of image restoration with techniques of image segmentation, based on Markov Random Field models. The problem is formulated as the minimization of a cost function which accounts for data consistency and for constraints derived from the adopted image model. A solution strategy is proposed where steps of image estimation iteratively alternate with steps of estimation for the degradation operator. To cope with the problem of space-variant blurs, we propose a recursive procedure that, starting with the estimation of a single blur mask for the whole image, refines the estimate in those zones of the image where suitable validation tests, based also on a linguistic analysis, reveal an error.

Keywords: Printed text segmentation, Image segmentation, Blind image restoration, Markov Random Fields.

1. INTRODUCTION

The computerized retrieval and restoration of textual information contained in damaged printed documents has recently attracted much attention within various national programs regarding the saving and the exploitation of the cultural heritage [1].

Standard optical character recognition (OCR) functions, now available as commercial software packages, cannot be applied to these images, due to their extremely poor quality. Indeed, the available images are usually very degraded because of many non quantifiable factors, such as ageing of the paper and microfilm support, smearing of the inks, deterioration due to usage, aberrations of the optical imaging system, low resolution of the digital acquisition, and so on. These various degradation factors globally act as an unknown space-variant blur operator that, especially when strong, makes the various characters to spread and overlap one another. Even more sophisticated OCR systems, based for instance on Hidden Markov models and/or neural networks [2,3], can fail in these situations, since the blurring may cause the segmentation step to produce joined and/or broken characters [4,5]. Hence, enhancement techniques are necessary, in order to remove the blur and, at the same time, separate each character from the others and from the background.

To simplify the problem, we assume that the images can be partitioned into sub-images where the space-variant blur can be approximated as a space-invariant blur. Thus, we propose to integrate techniques of blind image restoration [6,7,8] with techniques of image segmentation [9,10], so as to recover the ideal undegraded image in an already segmented form, where the pixels are grouped into mutually exclusive regions, that are homogeneous with respect to a discrete and finite set of assigned gray levels or labels. As a sub-product, we also obtain an estimate of the blur mask for each sub-image considered.

To cope with the problem of space-variant blurs, we propose a recursive procedure that, starting with the estimation of a single blur mask for the whole image, refines the estimate, and then improves segmentation, in those zones of the image where suitable validation tests, based on a linguistic analysis, reveal errors on the result of the subsequent OCR process.

2. THE PROPOSED METHOD

The problem of recovering a segmented image and jointly removing the blur is highly ill-posed in that it does not admit a unique solution. To make the problem well-posed we exploit regularization techniques [11,12] based on the definition of suitable mathematical models, and reformulate the problem as an optimization problem to be solved by minimizing a suitable cost function or energy, both with respect to the image f and the blur mask d.

By adopting a Multi-Level Logistic (MLL) model for a 8-neighbours neighbourhood system, which is a typical Markov Random Field (MRF) model for piecewise constant images [9,10], we consider a cost function U(f|g,d) which accounts for consistency with the observed image g, smoothness constraints on the gray levels of pixels belonging to homogeneous regions in the image, and geometrical constraints on the character morphology. More specifically, our models exploit some relevant characteristics of printed texts, such as the fact that the ideal image is essentially a two-level image, one level corresponding to the background and the other to the characters, and the fact that the characters present sharp and regular boundaries. In formulas, the cost function that we adopt is given by:

$$U(f|g,d) = \|g - H(d)\, f\|^2 + \sum_{c \in C} V_c(f) \tag{1}$$

where f and g are the lexicographic vector form for f and g, respectively, H(d) is a block Toeplitz matrix, whose elements derive, according to a known rule, from the blur mask d, and $V_c(f)$ are potential functions that enforce the interaction of cliques c of adjacent pixels, where the clique size, shape and orientation depend on the order of the chosen neighbourhood system. These potential can be easily designed to describe both the local smoothness constraint typically enforced by the MLL model and the peculiar features of printed texts above described. With respect to the constraints to be enforced on the blur mask, since we know that it acts as a low-pass filter, we assume that its elements are positive and of unitary sum.

Our blind segmentation problem becomes then:

$$\min_{f,d} \|g - H(d)\, f\|^2 + \sum_{c \in C} V_c(f) \tag{2}$$

subject to the extra constraints:

$$\sum_{i,j} d_{i,j} = 1 \quad \text{and} \quad d_{i,j} \geq 0 \quad \forall\, i,j \tag{3}$$

The solution strategy to problem (2)-(3) consists of the alternate execution of steps of image estimation and steps of estimation for the degradation operator, according to the following iterative scheme [13,14,15]:

$$f^{(k)} = \arg\min_f \|g - H(d^{(k)})\, f\|^2 + \sum_{c \in C} V_c(f) \tag{4a}$$

$$d^{(k)} = \arg\min_d \|g - H(d)\, f^{(k)}\|^2 \tag{4b}$$

where the constraints (3) are imposed on the solution after each iteration.

In order to perform the minimization (4a), and then to estimate the segmented image, we adopt a simulated annealing (SA) algorithm with Gibbs sampler [16,12]. This SA is periodically interrupted to produce a new estimate of the blur mask, via the gradient descent solution of the least-squares problem (4b).

By introducing the probability distribution:

$$P_\tau(f|g,d) = \frac{1}{Z} \exp\left[\frac{-U(f|g,d)}{\tau}\right] \tag{5}$$

where τ is a positive temperature parameter, the SA procedure above described is implemented according to the following iterative scheme:

1. set $k=0$, $d^{(k)}$, $f^{(k)}$, τ_k
2. set $t=1$, $f^{(t)} = f^{(k)}$

   for $t=1, L$

   compute $f^{(t+1)}$ according to $P_{\tau_k}(f|g,d^{(k)})$
3. set $f^{(k+1)} = f^{(L)}$
4. compute

   $$d^{(k+1)} = \arg\min_d \|g - H(d)\, f^{(k+1)}\|^2$$

   subject to the extra constraints:

   $$\sum_{i,j} d^{(k+1)}_{i,j} = 1 \quad \text{and} \quad d^{(k+1)}_{i,j} \geq 0 \quad \forall\, i,j$$
5. set $k=k+1$, go to step 2 until a termination criterion is satisfied

It is easy to see that the convergence of the algorithm, i.e. stabilization of the solutions, is ensured since, as τ_k approaches zero, $P_{\tau_k}(f|g,d^{(k)})$ becomes peaked around $f^{(k)}$. In fact, when this condition occurs, f do not change any more. By virtue of stabilization of f, d stabilizes as well, since it is computed deterministically as the minimizers of a cost function with now fixed parameters.

Due to the fact that the blur mask is usually of small size (e.g. 5∞5) and the image takes on only two gray levels, the computational complexity of this SA is mainly related to the size of the image treated. Although this size is usually high, considering that the MRF models to be adopted for describing the text characters entail short-range interactions among pixels, a very high degree of parallelization can be exploited for the implementation of the whole procedure.

3. EXPERIMENTAL RESULTS

We analyzed the performance of our procedure for blind segmentation both from a quantitative and a qualitative point of view. For a quantitative analysis, we first tested it on binary synthetic images artificially degraded by different amounts of blur and noise. For the real case of images of ancient printed documents, the procedure has been applied to several zones that are different for size and position in the considered documents.

In all experiments we started the process with the Dirac function and a roughly segmented version of the degraded image as initial guesses for the blur mask and the restored image, respectively. The initial value of the temperature for the SA algorithm was τ_0=500, and the law for decreasing it was chosen to be exponential, given by formula τ_{k+1}= 0.85 τ_k. At each temperature we performed 20 updates of the image field which gives a length for the Markov chains of 20N∞M, where N∞M is the size of the image. The minimization, with respect to the blur mask d, of the function in step (4b), can be performed by setting to zero its gradient. Thus, the solution could be obtained by solving a linear system. Although the matrix of the system is very small, owing to the usually small size of the blur masks considered, it is clear that it depends on the current intensity map and it is not immediate to verify whether it is non-singular. For this reasons, instead than solving the system by direct inversion of this matrix, we prefer to adopt a conjugate gradient of dimension equal to the mask size to minimize the original function. In all trials convergence to the final values of the parameters and stabilization of the reconstructions were reached in less than 30 iterations of the whole procedure. Figure 1 shows an example of the typical results obtained for synthetic images.

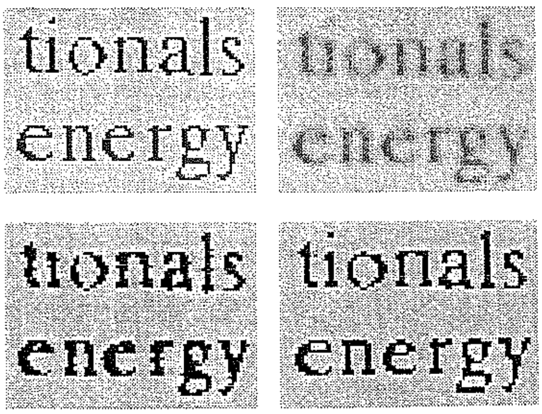


Fig. 1 -Blind segmentation of a 60∞80 synthetic image: (a) original image; (b) degraded image; (c) thresholded image; (d) segmented image.

In this case the original, ideal image was artificially degraded by convolution with the following 5∞5 blur mask:

0.000000	0.019231	0.019231	0.019231	0.000000
0.019231	0.067308	0.086538	0.067308	0.019231
0.019231	0.086538	0.153846	0.086538	0.019231
0.019231	0.067308	0.086538	0.067308	0.019231
0.000000	0.019231	0.019231	0.019231	0.000000

plus addition of Gaussian, white noise with zero mean and standard deviation σ=5. For comparison, Figure 1(c) shows the result of a simple thresholding, performed at the gray level which gives the best trade-off between disjoint and unbroken characters. The blur mask jointly estimated with the segmented image is the following:

0.015461	0.024145	0.039906	0.020161	0.014395
0.022793	0.050014	0.083375	0.054704	0.016783
0.015497	0.081305	0.140602	0.080030	0.025298
0.014023	0.056744	0.090220	0.058598	0.014566
0.012187	0.026634	0.046452	0.034366	0.017466

which has a root mean squared error of about 0.003 with respect to the ideal one.

As examples of the real images we treated, we present the results obtained on two portions of a digitized microfilm of the First Book of the Girolamo Cardano's "Opera Omnia". In the first case we assumed for the unknown blur mask a size of 5∞5 and obtained the segmentation shown in Figure 2(b) and the following estimated blur coefficients:

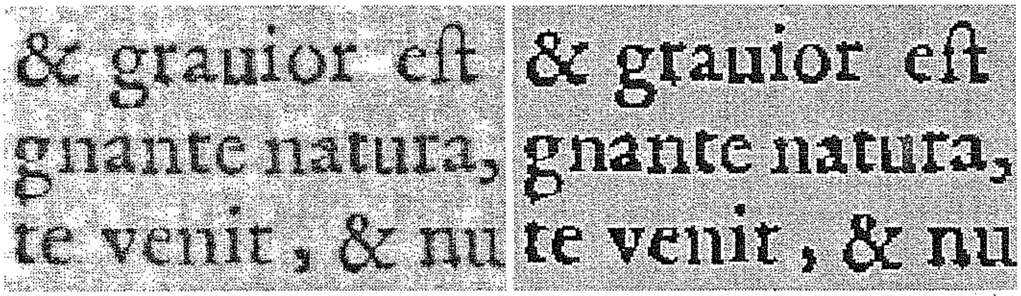
0.000000	0.007842	0.020513	0.005781	0.000000
0.000000	0.043513	0.123653	0.042452	0.003902
0.017202	0.107473	0.305856	0.125099	0.015545
0.000000	0.032287	0.109840	0.041221	0.000000
0.000000	0.000000	0.016008	0.012388	0.007752

The low values of the coefficients in the external border clearly indicate that the size of the blur mask has been overestimated. In fact, we repeated the experiment by assuming a 3∞3 blur mask, and obtained a very similar segmented image and an estimated blur mask with the following coefficients:

0.046973	0.136471	0.067762
0.093726	0.291336	0.155779
0.049481	0.106917	0.069465

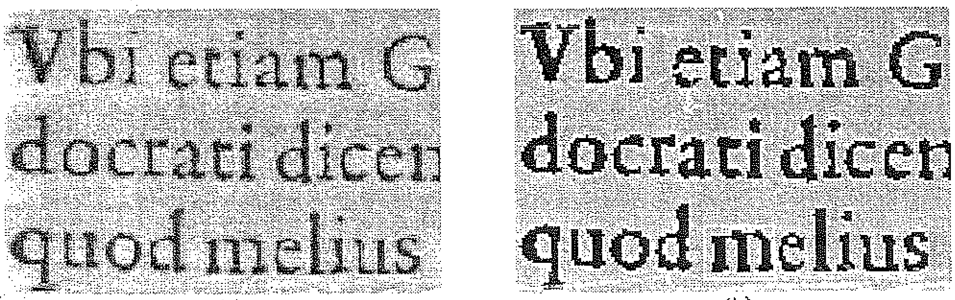
which closely resemble the central ones of the previously estimated 5∞5 blur mask. Thus, we could conclude that, for blind segmentation purposes, the assumed dimension of the blur mask can be overestimated without affecting the quality of the solutions obtained. This is advantageous because it reduces the number of the parameters to be chosen. On the other hand, considering the very small size of the blur masks usually adopted, the increasing of the computation time is insignificant.

With respect to the quality of the segmentation obtained, we can observe that all the characters have been correctly separated, while some characters appear to be broken (observe the "e" and the last "a" of the second row, and the first "n" of the third row). For the recognition of both the "e" and the "a" this defect is not critic, in that the peculiar structure of this two characters has been well preserved. Vice versa, the broken "n" could give rise to the recognition of two separated "i". The blind segmentation of the second real image, shown in Figure 3, gave a similar problem with the "m" of the third row. In this case, to obtain a correct restoration, we applied to the image the refining procedure above mentioned, by re-processing the only portion of the original image containing the "m". The final segmentation, with the "m" marked, is shown in Figure 3(b).



(a) (b)

Fig. 2 - Blind segmentation of a 110∞190 portion of a page of the First Book of the Opera Omnia of Cardano: (a) original, degraded image; (b) segmented image.



(a) (b)

Fig. 3 - Blind segmentation of a 110∞165 portion of a page of the First Book of the Opera Omnia of Cardano: (a) original, degraded image; (b) segmented image.

4. CONCLUSIONS

We have proposed a technique, based on Markov Random Field image models, for segmentation of text characters in presence of unknown space-variant blur and additive noise. This technique has been applied to highly degraded ancient printed documents, with the aim at facilitating subsequent phases of recognition and classification of the characters themselves. We formulated our technique as the alternate, iterative minimization of a cost function with respect to the image field and the degradation operator. Several results of both simulated and real experiments have been shown to validate the method.

The procedure has a high computational complexity which prevents its use to a vast extent for the processing of large amounts of documents, unless parallel architectures are exploited. Fortunately, most of the algorithms which constitute the procedure are local and distributed and then suitable for parallel implementations.

Based on the results we have obtained by analyzing and processing several portions of different documents, we conclude that the effectiveness of the proposed technique depends on the severity of the damage affecting the texts. In the case of texts strongly degraded, it is likely to find that some character are not correctly separated. In that case we found that good results can be obtained by refining the estimate of the blur mask in the zones of the text where the characters were incorrectly segmented. Thus we are currently studying a method based on the integration of the module for text segmentation, herein described, with modules for character recognition and linguistic analysis. In this method, the zones of the text where one of the two latter modules detects an error are forwarded to the module for text segmentation which operates the refining of the blur mask estimate and then produces a new restoration.

References

1 L. Bedini, A. Bozzi, A. Tonazzini, Digital techniques for character recognition in old documents, *ERCIM News*, 28, 1997, 24.

2 H. I. Avi-Itzhak, T. A. Diep and H. Garland, High Accuracy Optical Character Recognition Using Neural Networks with Centroid Dithering, *IEEE Trans. Pattern Anal. Machine Intell.*, 17(2), 1995, 218-224.

3 K. Aas and L. Eikvil, Text page recognition using grey-level features and hidden Markov models, *Pattern Recognition*, 29(6), 1996, 977-985.

4 S. D. Yanowitz and A. M. Bruckstein, A new method for image segmentation, *CVGIP*, 46, 1989, 82-95.

5 T. Taxt, P. J. Flynn and A. K. Jain, Segmentation of document images, *IEEE Trans Pattern Anal. Machine Intell.*, 11, 1989, 1322-1329.

6 A. K. Katsaggelos and K. T. Lay, Maximum likelihood blur identification and image restoration using the EM algorithm, *IEEE Trans. Signal Processing*, 39, 1991, 729-732.

7 R. L. Lagendijk, A. M. Tekalp and J. Biemond, Maximum likelihood image and blur identification: a unifying approach, *Opt. Eng.*, 29, 1990, 422-435.

8 J. Zhang, The mean field theory in EM procedures for blind Markov Random Field image restoration, *IEEE Trans. Image Proc.*, 2, 1993, 27-40.

9 S. Lakshmanan and H. Derin, Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing, *IEEE Trans. Pattern Anal. Machine Intell.*, 11, 1989, 799-813.

10 S. Z. Li, *Markov Random Field Modeling in Computer Vision*, Tokyo, Springer-Verlag, 1995.

11 L. Bedini, I. Gerace, E. Salerno and A. Tonazzini, Models and algorithms for edge-preserving image reconstruction, *Advances in Imaging and Electron Physics*, 97, P.W. Hawkes Ed., 1996, 86-189.

12 S. Geman and D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal. Machine Intell.*., 6, 1984, 721-740.

13 G. R. Ayers and J. G. Dainty, Iterative blind deconvolution method and its applications, *Opt. Lett.*, 13, 1988, 547-549.

14 A. Tonazzini, L. Bedini, Using intensity edges to improve parameter estimation in blind image restoration, SPIE's Int. Symposium on Optical Science, Engineering, and Instrumentation, Bayesian Inference for Inverse Problems (SD99), 19-24 July 1998, San Diego, *Proceedings of SPIE*, 3459, 1998, 73-81.

15 Y. You and M. Kaveh, A regularization approach to joint blur identification and image restoration, *IEEE Trans. Image Proc.*, 5, 1996, 416-428.

16 E. Aarts and J. Korst, *Simulated Annealing and Boltzmann Machines*, John Wiley & Sons, 1989.