LAMP-TR-066	4400019848
CAR-TR-963	N660010028910/IIS9987944
CS-TR-4219	February 2001

A Downhill Simplex Algorithm for Estimating Morphological Degradation Model Parameters

Tapas Kanungo and Qigong Zheng

Language and Media Processing Laboratory
Center for Automation Research
University of Maryland
College Park, MD 20742-3275
{kanungo,qzheng}@cfar.umd.edu

Abstract

Noise models are crucial for designing image restoration algorithms, generating synthetic training data, and predicting algorithm performance. However, to accomplish any of these tasks, an estimate of the degradation model parameters is essential. In this paper we describe a parameter estimation algorithm for a morphological, binary image degradation model. The inputs to the estimation algorithm are i) the degraded image, and ii) information regarding the font type (italic, bold, serif, sans serif). We simulate degraded images and search for the optimal parameter by looking for a parameter value for which the neighborhood pattern distributions in the simulated image and the given degraded image are most similar. The parameter space is searched using the Nelder-Mead downhill simplex algorithm. We use the p-value of the Kolmogorov-Smirnov test for the measure of similarity between the two neighborhood pattern distributions. We show results of our algorithm on degraded document images.

This research was funded in part by the Science Application International Corporation under Contract 4400019848, the Defense Advanced Research Projects Agency under Contract N660010028910, and the National Science Foundation under Grant IIS9987944.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **FEB 2001**	2. REPORT TYPE	3. DATES COVERED **00-02-2001 to 00-02-2001**

4. TITLE AND SUBTITLE **A Downhill Simplex Algorithm for Estimating Morphological Degradation Model Parameters**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Language and Media Processing Laboratory,Institute for Advanced Computer Studies,University of Maryland,College Park,MD,20742-3275**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**15**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

LAMP-TR-066 4400019848
CAR-TR-963 N660010028910/IIS9987944
CS-TR-4219 February 2001

A Downhill Simplex Algorithm for Estimating Morphological Degradation Model Parameters

Tapas Kanungo and Qigong Zheng

A Downhill Simplex Algorithm for Estimating Morphological Degradation Model Parameters

Tapas Kanungo and Qigong Zheng

Center for Automation Research
University of Maryland
College Park, MD 20742-3275
{kanungo,qzheng}@cfar.umd.edu

Abstract

Noise models are crucial for designing image restoration algorithms, generating synthetic training data, and predicting algorithm performance. However, to accomplish any of these tasks, an estimate of the degradation model parameters is essential. In this paper we describe a parameter estimation algorithm for a morphological, binary image degradation model. The inputs to the estimation algorithm are i) the degraded image, and ii) information regarding the font type (italic, bold, serif, sans serif). We simulate degraded images and search for the optimal parameter by looking for a parameter value for which the neighborhood pattern distributions in the simulated image and the given degraded image are most similar. The parameter space is searched using the Nelder-Mead downhill simplex algorithm. We use the p-value of the Kolmogorov-Smirnov test for the measure of similarity between the two neighborhood pattern distributions. We show results of our algorithm on degraded document images.

This research was funded in part by the Science Application International Corporation under Contract 4400019848, the Defense Advanced Research Projects Agency under Contract N660010028910, and the National Science Foundation under Grant IIS9987944.

1 Introduction

Numerous document image degradation models have been proposed in the literature [1, 7, 8]. However, prior to using these models, it is important to i) validate the models — that is, verify that the simulations generated by these models are similar to real-world examples, and ii) provide algorithms for estimating the model parameters from real samples. The issue of validation was addressed by Kanungo *et al.* [5, 6] by converting the validation problem into a hypothesis testing problem and then using a permutation test to test the null hypothesis that a synthetic sample of degraded characters and another sample of real degraded characters come from the same underlying distribution. Lopresti *et al.* [10] instead proposed to study the differences in the error characteristics of the OCR output for the real and synthetic samples. This method, however, considers the degradation coupled with the OCR system and not just the degradation process.

The issue of model parameter estimation has been studied to a lesser extent. Kanungo and Haralick [4] reported results of some preliminary experiments that they conducted to estimate the degradation model parameters using an objective function based on the power function. They assumed that that an ideal document image and the corresponding degraded image were given. Baird [2] used the same power function approach to estimate the parameters of another physics-based degradation model, and Sural and Das [13] estimated the parameters of a two-state Markov chain document degradation model using the power function approach. The drawback of all the above estimation approaches is that they assume that the degraded image and the ideal image are perfectly aligned and that character-level geometric groundtruth (bounding boxes) is available. This, however, is not easy to achieve since pixel-level alignment of documents with arbitrary warping present, due to changes in printer and scanner speeds, is difficult. A preliminary version of this paper appeared in [9].

In this paper we propose a parameter estimation algorithm that does not require the degraded and ideal images to be aligned and does not require character-level geometric groundtruth either. The algorithm is based on computing differences between the distributions of neighborhood patterns in the degraded and synthetic images. In Section 2 we describe our document degradation model. We outline the estimation algorithm in Section 4 and provide simulation results in Section 5.

2 The Morphological Document Degradation Model

In this section we briefly describe a document degradation model for the local degradation that are introduced when documents are printed, scanned and digitized [5, 7, 8].

The model accounts for (i) the pixel inversion (from foreground to background and vice versa) that occurs independently at each pixel due to light intensity fluctuations, pixel sensitivity, and thresholding level, and (ii) the blurring that occurs due to the point-spread function of the optical system of the scanner. We model the probability of a background pixel flipping as an exponential function of its distance from the nearest boundary pixel. The parameter α_0 is the initial value for the exponential, and the decay speed of the exponential is controlled by the parameter α. The foreground and background 4-neighbor distance are computed using a standard distance transform algorithm [3] .

The flipping probabilities of the foreground pixels are similarly controlled by β_0 and β. The parameter η is the constant probability of flipping for all pixels. Finally, the last parameter k, which is the size of the disk used in the morphological closing operation [3], accounts for the correlation introduced by the point-spread function of the optical system.

The degradation model thus has six parameters: $\Theta = (\eta, \alpha_0, \alpha, \beta_0, \beta, k)$. These parameters are used to degrade an ideal binary image as follows:

1. Compute the distance d of each pixel from the character boundary.

2. Flip each foreground pixel with probability $p(0|1, d, \alpha_0, \alpha) = \alpha_0 e^{-\alpha d^2} + \eta$.

3. Flip each background pixel with probability $p(1|0, d, \beta_0, \beta) = \beta_0 e^{-\beta d^2} + \eta$.

4. Perform a morphological closing operation with a disk structuring element of diameter k.

The application of the various steps of the model is illustrated in Figure 1. The procedure described above works on bit-mapped images. Since there is no restriction on the size of the image that can be degraded, or the language of the written text, an entire document page image can be degraded using this model.

3 Neighborhood Pattern Distributions

Our estimation algorithm is based on the assumption that if the degradation parameters are estimated correctly, the local degradations in a simulated image generated using the estimated parameters will look similar to those of a real image. The way we capture this fact is by looking at the distribution of neighborhood patterns.

Let P be a set of neighborhood bit patterns and p be an arbitrary element in the set P. For example, p could be a 3×3 neighborhood with all 1s, or it could be a 5×5 neighborhood with a 1 in the middle and 0s everywhere else. Now we define the neighborhood pattern distribution of an image R. Let H_R denote a neighborhood pattern distribution, so that $H_R(p)$, where $p \in P$, is the number of times the pattern p occurs in the binary image R. Using mathematical morphology [3] we can define $H_R(p)$ more precisely: $H_R(p) = \#\{R \ominus p\}$.

We conducted three experiments to study whether the pattern distributions could discriminate various font and language characteristics. In particular, we studied whether the change in i) fonts, ii) font size, or iii) text or text language probabilities could be detected using the neighborhood pattern distributions in ideal images. In Figure 2 and Table 1 we show subimages of same text typeset in serif, sans serif, bold, and italic fonts. We find that the Kolmogorov-Smirnov test can easily detect the differences in the neighborhood pattern distributions. In Figure 3 and Table 2 we show that even if we change the font size of serif text, the neighborhood pattern distributions are quite indistinguishable. Finally, in Figure 4 and Table 3 we show that if we replace the original text with another from the same source, and keep the font characteristics identical, the

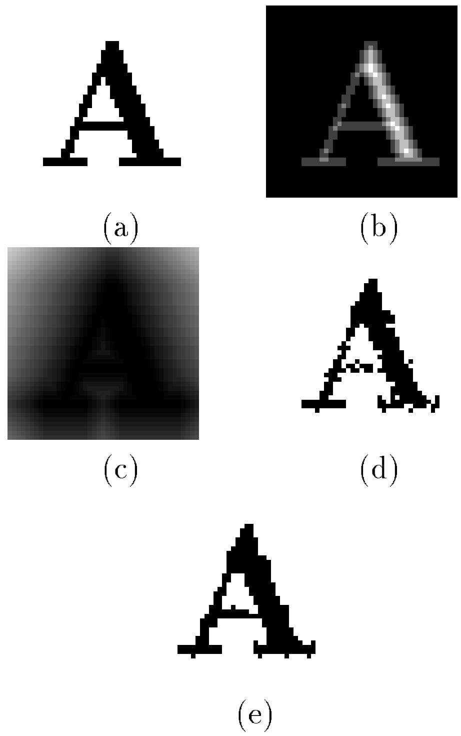


Figure 1: Local document degradation model: (a) Ideal noise-free character; (b) Distance transform of the foreground; (c) Distance transform of the background; (d) Result of the random pixel-flipping process (the probability of a pixel flipping is $p(0|d,\beta,f) = p(1|d,\alpha,b) = \alpha_0 e^{-\alpha d^2}$; here $\alpha = \beta = 2$, $\alpha_0 = \beta_0 = 1$); (e) Morphological closing of the result in (d) by a 2×2 binary structuring element.

Kolmogorov-Smirnov test cannot detect the difference. This is quite an interesting result because it says that in order to compare the noise pattern distributions of two images, the two images need not have the same underlying ideal image.

Table 1: Kolmogorov-Smirnov Test Statistics and Significance level (T, P-value).

(T, P-value)	Serif	Sans Serif	Serif Bold	Serif Italic
Serif	(0.0,1.0)	(0.193,0.00)	(0.120,0.00)	(0.078,0.09)
Sans Serif	(0.193,0.00)	(0.0,1.0)	(0.096,0.02)	(0.25,0.00)
Serif Bold	(0.12,0.00)	(0.096,0.02)	(0.0,1.0)	(0.19,0.00)
Serif Italic	(0.078,0.09)	(0.25,0.00)	(0.19,0.00)	(0.0,1.0)

4 The Estimation Algorithm

Let I be the ideal image and R be the given degraded image. The problem is to estimate the model parameter θ such that if we degrade I with the model with parameter fixed

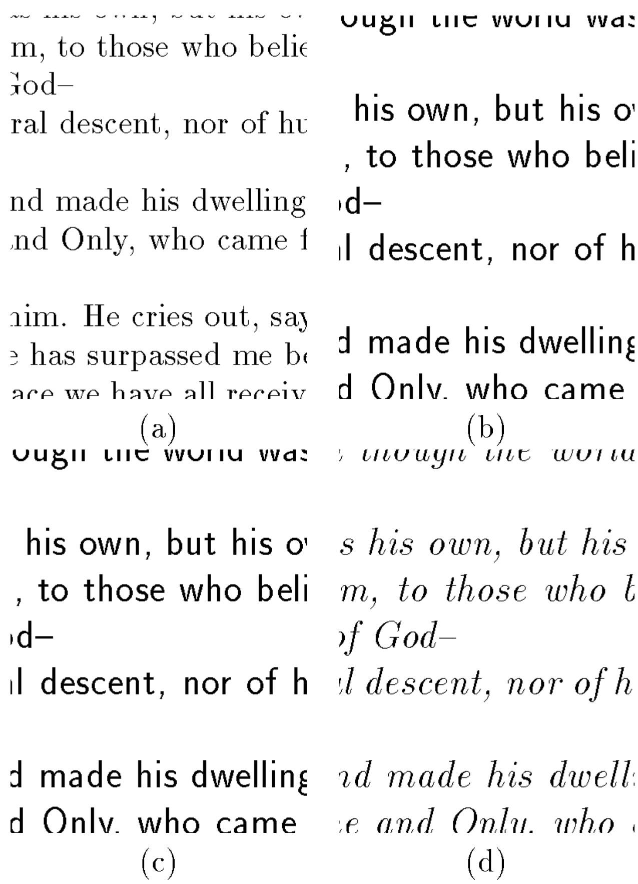


Figure 2: Text typeset in Computer Modern Roman font. (a) Serif text; (b) Sans Serif text; (c) Serif bold text; (d) Serif Italic text.

at θ, we will get an image S_θ that looks similar to R. For our purposes, we say that two images R and S are similar if the corresponding neighborhood pattern distributions H_R and H_{S_θ} are similar. We use the Kolmogorov-Smirnov test [11] to test the similarity of the two neighborhood pattern distributions. Let $KS(H_R, H_{S_\theta})$ denote the KS test p-value for the null hypothesis that the two distributions are same. We will use this p-value as the objective function that the estimation process tries to maximize. That is,

$$\hat{\theta} = \max_{\theta} KS(H_R, H_{S_\theta}). \tag{1}$$

Notice that S_θ is computed by simulation. Thus the derivatives of the objective function cannot be computed in closed form. Hence standard derivative approaches to maximizing KS are not applicable. We therefore used the Nelder-Mead derivative-free optimization algorithm [12] to maximize KS. There is no reason to believe that KS is unimodal over the model parameter space; the Nelder-Mead algorithm provides us with a local maximum. To circumvent this problem we do multiple random starts and then pick the solution corresponding to the highest maximum value.

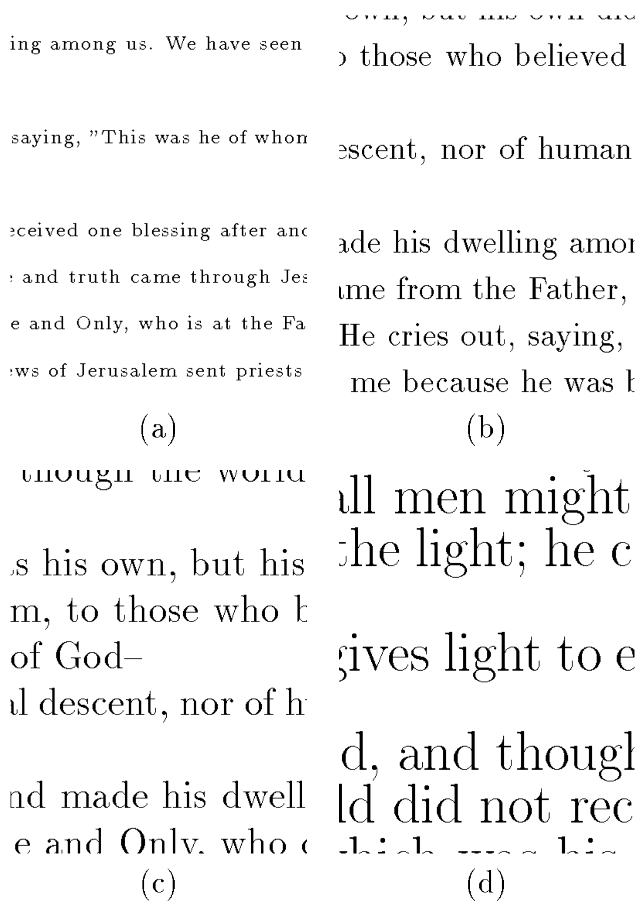


Figure 3: Text in various font sizes. (a) 6pt font; (b) 8pt font; (c) 12pt font; (d) 17pt font.

5 Protocol and Results

We start with a 400×400 ideal binary image I such as that shown in Figure 5(a). The given degraded image R shown in Figure 5(b) was created using the model parameters $\theta = (0.0, 0.6, 1.5, 0.8, 2.0, 3)$. The neighborhood pattern set P was chosen to be all the possible binary patterns in a 3×3 window. Thus P has 512 patterns. The neighborhood pattern distribution corresponding to Figures 5(a)–(c) are shown in Figures 6(a)–(c). Notice that some patterns occur more frequently than others, and that the distributions of the ideal and degraded images are different. The search was done for $\alpha_0, \alpha, \beta_0, \beta$; η and k were assumed known. The Nelder-Mead algorithm was started 10 times with random start locations. The objective function value $(1 - p\text{value})$ is plotted as a function of iterations in Figure 7. The best optimal solution is found to be $\hat{\theta} = (0.0, 0.64, 1.57, 0.96, 2.02, 3)$. In Figure 5(c) we show the image $R_{\hat{\theta}}$ generated using the optimal solution $\hat{\theta}$. Notice that the neighborhood pattern distribution corresponding to the estimated image, which is shown in Figure 6(c), is quite similar to the histogram of the original degraded image shown in Figure 6(b). Note that the ideal image need not correspond to the degraded image. In fact, one can use any other ideal image that has i) the same font type as that of the degraded image (the font size can be different, however), and ii) language properties similar to those of the degraded text.

Table 2: Kolmogorov-Smirnov test statistics and significance level (T, P-value).

(T, P-value)	6pt	8pt	12pt	17pt
6pt	(0.0,1.0)	(0.053,0.472)	(0.057,0.382)	(0.086,0.045)
8pt	(0.053,0.472)	(0.0,1.0)	(0.062,0.268)	(0.076,0.101)
12pt	(0.057,0.382)	(0.062,0.268)	(0.0,1.0)	(0.049,0.572)
17pt	(0.086,0.045)	(0.076,0.101)	(0.049,0.572)	(0.0,1.0)

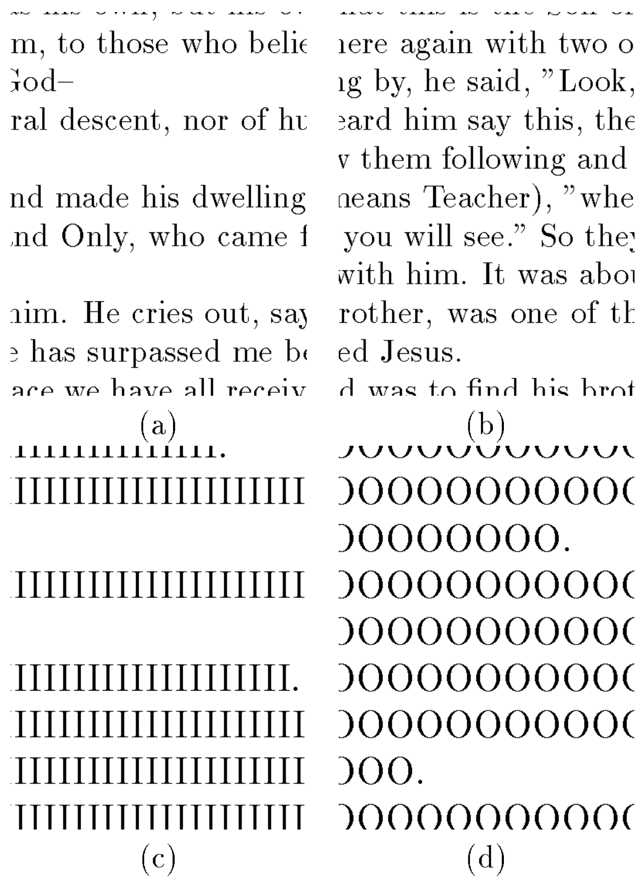
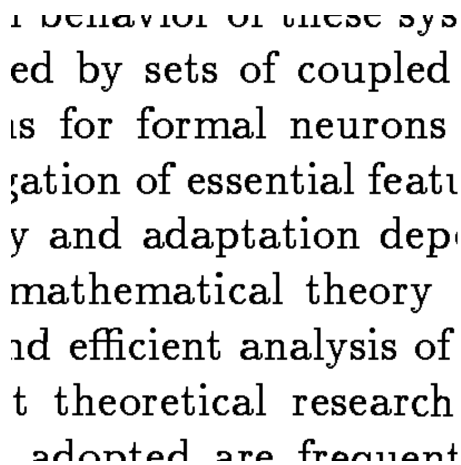


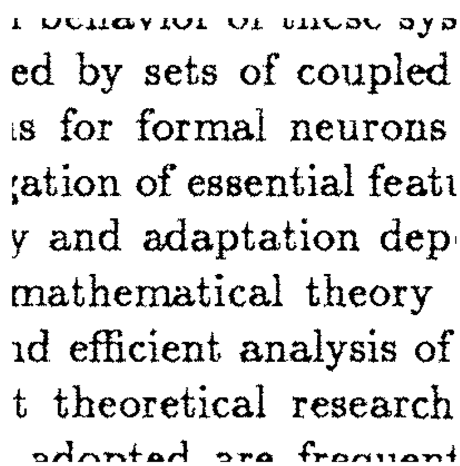
Figure 4: Texts in 12pt serif Computer Modern Roman font. (a) A fragment of text from one document. (b) Another fragment from a similar document. (c) A fragment of text containing only 'I's. (d) A fragment of text containing only 'O's.

Table 3: Kolmogorov-Smirnov test statistics and significance level (T, P-value).

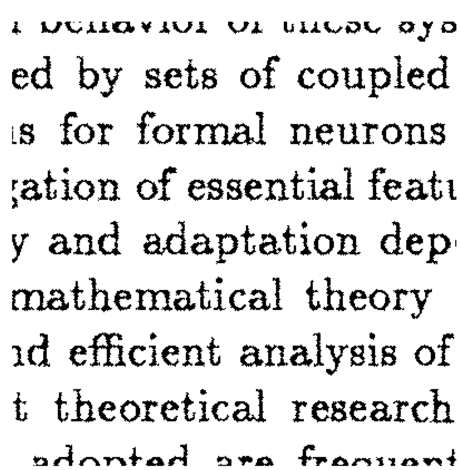
(T, P-value)	Fig 4(a)	Fig 4(b)	Fig 4(c)	Fig 4(d)
Fig 4(a)	(0.0,1.0)	(0.025,0.996)	(0.392,0.00)	(0.384,0.00)
Fig 4(b)	(0.025,0.996)	(0.0,1.0)	(0.412,0.0)	(0.404,0.00)
Fig 4(c)	(0.392,0.00)	(0.412,0.0)	(0.0,1.0)	(0.075,0.118)
Fig 4(d)	(0.384,0.00)	(0.404,0.0)	(0.075,0.118)	(0.0,1.0)



(a)

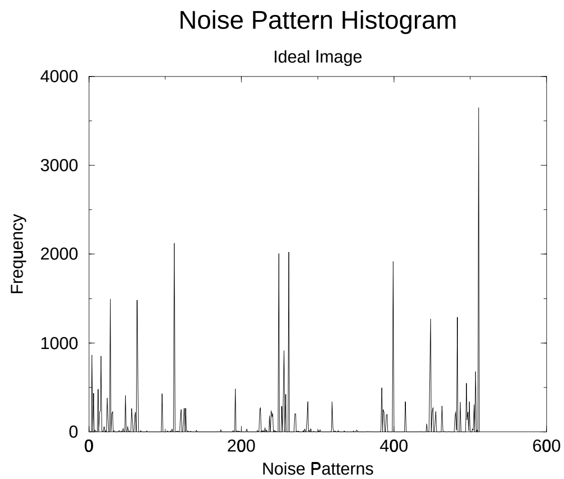


(b)

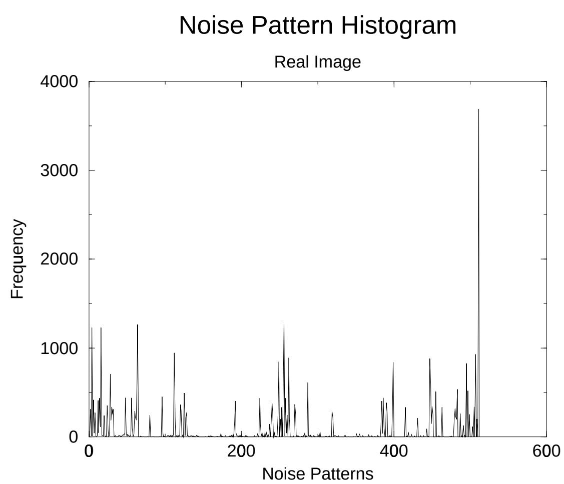


(c)

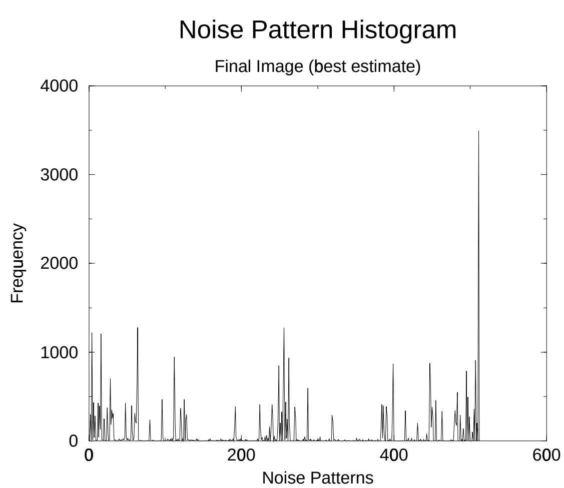
Figure 5: (a) A typical ideal image. (b) A degraded image with parameters $(0.0, 0.6, 1.5, 0.8, 2.0, 3)$. (c) Image generated using the estimated parameters $(0.00.64, 1.57, 0.96, 2.02, 3)$.



(a)



(b)



(c)

Figure 6: Neighborhood pattern distributions corresponding to Figures 5(a)–(c). Each bin along the x-axis corresponds to a different 3×3 neighborhood pattern.

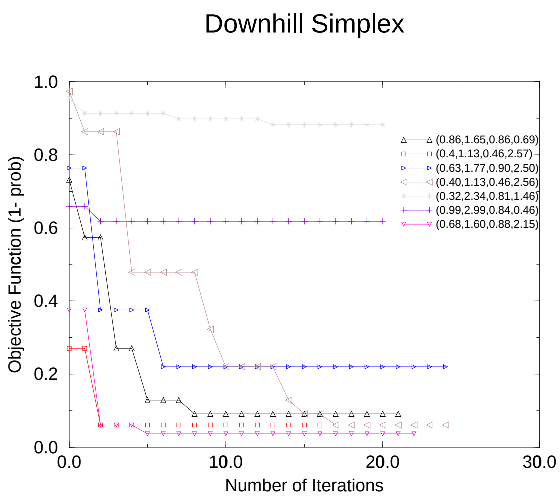


Figure 7: Downhill simplex convergence for different (random) starting locations.

6 Conclusion

We have described an algorithm for estimating the parameters of a degradation model. The algorithm assumes that we know or can estimate the font type (serif, sans serif, bold, italic) of the degraded image and then typeset an arbitrary ideal text image in the same font. The ideal image is then degraded with various parameters of the degradation model. For each parameter value the neighborhood pattern distributions of the ideal and the degraded images are compared using the Kolmogorov-Smirnov test. The parameter value that maximizes the p-value is used as an estimate of the model parameters. The search for the optimal parameters is done using the Nelder-Mead algorithm.

Acknowledgments

This research was funded in part by Science Applications International Corporation under Contract 4400019848, the Defense Advanced Research Projects Agency under Contract N660010028910, and the National Science Foundation under Grant IIS9987944.

References

[1] H. Baird. Document image defect models. In *Proc. of IAPR Workshop on Syntactic and Structural Pattern Recognition*, pages 38–46, Murray Hill, NJ, June 1990.

[2] H. S. Baird. Document image quality: Making fine discriminations. In *Proc. of International Conference on Document Analysis and Recognition*, pages 459–462, Bangalore, India, September 1999.

[3] R. M. Haralick and L. G. Shapiro. *Robot and Computer Vision* (vols. 1 and 2). Addison-Wesley, Reading, MA, 1992.

[4] T. Kanungo and R. M. Haralick. Morphological degradation parameter estimation. In *Proc. of SPIE Conference on Nonlinear Image Processing*, volume 2424, pages 86–95, San Jose, CA, February 1995.

[5] T. Kanungo, R. M. Haralick, H. Baird, W. Stuetzle, and D. Madigan. A statistical, nonparametric methodology for document degradation model validation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:1209–1223, 2000.

[6] T. Kanungo, R. M. Haralick, H. S. Baird, W. Stuetzle, and D. Madigan. Document degradation models: Parameter estimation and model validation. In *Proc. of International Workshop on Machine Vision Applications*, Kawasaki, Japan, December 1994.

[7] T. Kanungo, R. M. Haralick, and I. Phillips. Global and local document degradation models. In *Proc. of International Conference on Document Analysis and Recognition*, pages 730–734, Tsukuba, Japan, October 1993.

[8] T. Kanungo, R. M. Haralick, and I. Phillips. Non-linear local and global document degradation models. *International Journal of Imaging Systems and Technology*, 5:220–230, 1994.

[9] T. Kanungo and Q. Zheng. Estimation of morphological degradation model parameters. In *Proc. of IEEE International Conference on Speech and Signal Processing*, Salt Lake City, UT, May 2001.

[10] Y. Li, D. Lopresti, G. Nagy, and A. Tomkins. Validation of document defect models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18:99–107, 1996.

[11] F. J. Massey. The Kolmogorov-Smirnov test for goodness. *J. of Amer. Stat. Assoc.*, 46:68–78, 1951.

[12] J. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7:308–313, 1965.

[13] S. Sural and P. K. Das. A two-state Markov chain model of degraded document images. In *Proc. of International Conference on Document Analysis and Recognition*, pages 463–466, Bangalore, India, September 1999.