Applications of Digital Image Processing VII

Andrew G. Tescher
Chairman/Editor

1984

Proceedings of SPIE—The International Society for Optical Engineering

Volume 504

Applications of Digital Image Processing VII

Andrew G. Tescher
Chairman/Editor

Cooperating Organizations
Image Processing Institute/University of Southern California
Optical Sciences Center/University of Arizona
Institute of Optics/University of Rochester

August 21-24, 1984
San Diego, California

Published by
SPIE—The International Society for Optical Engineering
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone 206/676-3290 (Pacific Time) • Telex 46-7053

SPIE (The Society of Photo-Optical Instrumentation Engineers) is a nonprofit society dedicated to advancing engineering and scientific applications of optical, electro-optical, and optoelectronic instrumentation, systems, and technology.

The papers appearing in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and are published as presented and without change, in the interests of timely dissemination. Their inclusion in this publication does not necessarily constitute endorsement by the editors or by SPIE.

Please use the following format to cite material from this book:
Author(s), "Title of Paper," *Applications of Digital Image Processing VII*, Andrew G. Tescher, Editor, Proc. SPIE 504, page numbers (1984).

Library of Congress Catalog Card No. 84-072640
ISBN 0-89252-539-8

Printed in the United States of America.

APPLICATIONS OF DIGITAL IMAGE PROCESSING VII

Volume 504

Conference Committee

Chairman
Andrew G. Tescher
The Aerospace Corporation

Co-Chairmen
William B. Green
System Development Corporation/Burroughs

Ali Habibi
The Aerospace Corporation

To Russell Hsing
GTE Laboratories Incorporated

André Oosterlinck
Catholic University of Leuven, Belgium

Robert A. Gonsalves
Northeastern University

Session Chairmen
Session 1—Pattern Recognition
Andrew G. Tescher, The Aerospace Corporation

Session 2—Industrial Applications
André Oosterlinck, Catholic University of Leuven, Belgium

Session 3—Hardware and Systems Issues
William B. Green, System Development Corporation/Burroughs

Session 4—Modular Configurations for Multiple Image Formats
Francis J. Corbett, Imaging Technology, Incorporated

Session 5—Analytical Methods and Concepts
Ali Habibi, The Aerospace Corporation

Session 6—Enhancement and Restoration
Robert A. Gonsalves, Northeastern University

Session 7—Miscellaneous Applications
John A. Saghri, The Aerospace Corporation

Session 8—Coding Technology
To Russell Hsing, GTE Laboratories Incorporated

APPLICATIONS OF DIGITAL IMAGE PROCESSING VII

Volume 504

INTRODUCTION

Digital image processing has continued to demonstrate the vitality of this field through the contribution of many excellent papers at this conference. The conference covered four days including eight sessions. Most major fields of the digital image processing technology were well represented. Specific topics included pattern recognition, industrial applications, hardware and systems issues, various analytical techniques including enhancement and restoration, and image compression. Similarly to previous years, a significant number of papers were from outside the USA, indicating the growing international character of this important technology. The contribution and efforts of the co-chairmen and session chairmen are gratefully acknowledged.

Andrew G. Tescher
The Aerospace Corporation

APPLICATIONS OF DIGITAL IMAGE PROCESSING VII

Volume 504

Contents

APPLICATIONS OF DIGITAL IMAGE PROCESSING VII

Volume 504

Session 1

Pattern Recognition

Chairman
Andrew G. Tescher
The Aerospace Corporation

Character Image Segmentation

Yoshitake Tsuji and Ko Asai

C&C Systems Research Laboratories
NEC Corporation

1-1, Miyazaki 4-chome, Miyamame-ku, Kawasaki-city,
Kanagawa 213, Japan

Abstract

In the Optical Character Reader (OCR) system design, the character segmentation technique is important. For example, the Automatic Mail Address Reader is required to manage printed characters of many font types and poor print quality. In this case, OCR performance will be affected by character segmentation technique.

This paper describes two new methods for character segmentation under more general conditions. The character segmentation problem can be formulated and classified as a pitch estimation problem and a character sectioning decision problem. These problems are resolved by using a statistical analysis method based on least square error function and a dynamic programing method with the minimum variance for separation between candidate positions in a line image. The effectiveness of the proposed methods has been evaluated through actual mail address segmentation experiments.

Introduction

OCR systems, such as the Automatic Mail Address Reader or the Automatic Document Reader, are required to manage printed characters of many font types and poor print quality. Moreover, printed characters in a line image are often touching each other or split because of poor print quality. Incorrect characters, such as both split and merged characters, are the cause of errors in OCR . Therefore, a reasonable and universal character segmentation technique is desired.

Several pitch based character segmentation methods have been developed in order to process the incorrect characters. These methods are basically applied to constant pitch characters. Most of those are based on standard character width and topological features detected from the character geometries. The beginning and end positions of characters are determined by using these different bits of information, one by one.[(1)-(3)] However, under general conditions existing for many font types and poor print quality too, these segmentation methods will be affected by some local fluctuations in character geometries. Therefore, it is necessary for all character sectioning positions in a line image to be decided on simultaneously, rather than for each character sectioning position to be individually determined.

On the other hand, if a line image ,which is composed of variable pitch characters, has poor print quality and printed characters are touching each other or split, the pitch based character segmentation will not be applied. In this case, character segmentation difficulties are increasing rapidly, and character segmentation problem will be considered as a character recognition problem with help from contextual information.[(6)] However,in such a case, the character segmentation performance will be dependent on the character recognition result, such as its algorithm, etc.[(3)-(6)] Accordingly, when the OCR system is required to manage printed characters of both constant pitch and variable pitch, it becomes necessary to estimate the character pitch accurately and to distinguish variable pitch data from constant pitch data. As a result, the adaptive character segmentation strategy according to the printed characters property will be applied.

In this paper, the character segmentation problem under more general conditions is considered and classified as pitch estimation and character sectioning decision problems.[(7)] A statistical analysis method, based on least square error function, is introduced for pitch estimation. By using this method, it is shown that character pitch can be exactly estimated and the property of printed characters can be automatically understood. Furthermore, in order to decide the character sectioning positions more reliably, a dynamic programing method with the minimum variance criterion is introduced for character sectioning decision, and its effective scope is discussed.

Pitch Estimation

First, consider the pitch estimation problem. An example of constant pitch characters is shown in Fig.1. In Fig.1, let P denote the character pitch. The pattern area, in which the vertical black bit lines are continious, is called the black region. This region is detected by using a known technique such as vertical histograms. The region can be regarded as indicating one character, only if printed characters are not touching each other or are not split. Let Bi (i = 1,2,.) denote the i-th black region width. Let Wi (i = 1,2,.) denote the i-th space width between the neighbouring black regions Bi,Bi+1. Let Wi(L) denote the space width between the left edge of the i-th black region and the beginning of the pitch frame including its region. Let Ui,j (i< j) denote the distance between the two left edges of the i-th black region and the j-th black region. Also, the above mentioned distance Ui,j can be defined by using the two right edges instead of the above mentioned two left edges. In addition, if distance Ui,j includes k characters, it is defined by distance Ui,j(k).

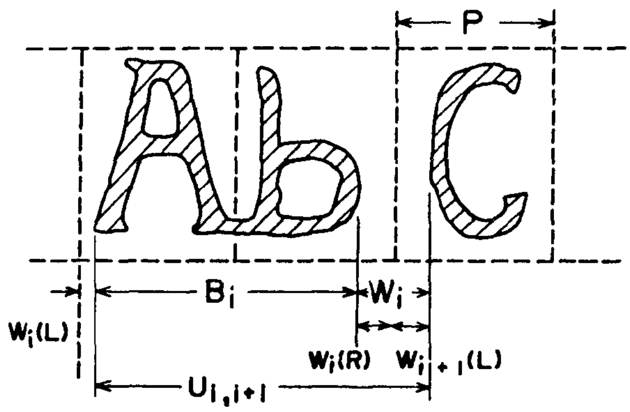


Figure 1. Constant pitch characters

Best linear unbiased estimator for constant pitch

According to the above definitions, the distance Ui,j(k) is formulated as follows.

$$U_{i,j}(k) = \sum_{t=i}^{j} (B_t + W_t)$$
$$= k \cdot P + W_i(L) - W_j(L)$$
$$= k \cdot P + w_{i,j}(k) \qquad (1)$$

where error wi,j(k) is considered a distortion in regard to a multiple of pitch k·P and can be handled as a random variable, independent from printed positions i,j, only due to print quality, font types , character categories and the optical scanner characteristics. Let Z_K denote the sample size for distance Ui,j'' Then, distance Ui,j(k) is represented by the following equation.

$$U(k,r) = k \cdot P + w(r) \quad (r= 1,2,.. ,Z_K) \qquad (2)$$

where distance U(k,r) is the r-th sample value in the distance Ui,j(k) distribution. Error w(r) can be assumed as a random variable, which has the statistical property of expectation $E[w(r)] = 0$, variance $E[w^2(r)] = V_w^2$.

For any character number k (k=1,..,n) in the distance U(k,r), from Eq.(2), it is easily shown that the quotient $\overline{U}(k,Z_K)/k$, which divides the sample mean $\overline{U}(k,Z_K)$ of distance U(k,r) by character number k, is an unbiased estimator for character pitch P. Hence, the Linear Square Error Function (LSEF) can be introduced for obtaining the best linear unbiased estimator $\hat{P}$.

$$LSEF = \sum_{k=1}^{n} C(k,Z_k) \cdot (\overline{U}(k,Z_k)/k - P)^2 \cdot /P^2 \qquad (3)$$

$$\text{where} \quad \sum_{K=1}^{n} C(k,Z_k) = 1 \qquad (4)$$

$C(k,z_K)$ is a coefficient satisfying $0 \leq C(k,Z_K) \leq 1$. For any printed characters satisfying Eq.(2), coefficient $C(k,Z_K)$ is determined as follows.

$$LSEF = 1/P^2 \cdot \sum_{K=1}^{n} \left[C(k,Z_k) \sum_{r=1}^{Z_K} (w(r) / k \cdot Z_k)^2 \right]$$

Expectation E [LSEF] is given by

$$E\left[LSEF\right] = 1/P^2 \sum_{K=1}^{n} C(k,Z_k) \, V_w^2 / k^2 \cdot Z_k \qquad (5)$$

8750226

From Eq.(5), if item $C(k,Z_K)/k^2 \cdot Z_K$ is equal in regard to character number k ($k = 1,..,n$), E LSEF becomes a minimum value . Accordingly, by using Eq.(4), item $C(k,Z_K)$ is given by

$$C(k,Z_k) = k^2 \cdot Z_k \Big/ \sum_{k=1}^{n} k^2 \cdot Z_k \qquad (6)$$

In Eq.(6), if printed characters are produced by constant pitch machines , LSEF becomes small in connection with the square of character number k and sample size Z_K . For any candidate pitch P(Ui,j), the best unbiased estimator $\hat{P}$ is obtained by P(Ui,j), having the minimum LSEF value.

Distance Ui,j Clustering

LSEF described above is determined on the assumption that distance U(k,r) can be clearly obtained. In order to obtain the distance U(k,r) distribution f(k) , the distance Ui,j distribution ,observed from line images, should be separated into individual distributions f(k), including unique character number k (k= 1,2,.. n). The best linear unbiased estimator for character pitch P will be obtained by statistical analysis method which includes both the LSEF computation and the individual distribution detection. Let s(f(k),f(k+1)) denote the boundary point between distributions f(k), f(k+1). Best linear unbiased estimator $\hat{P}$ is obtained by the following algorithm.

(Step 1) Make distance Ui,j disribution.

(Step 2) By determining a candidate pitch P(Ui,j) and each boundary point S(f(k),f(k+1)), divide distance Ui,j distribution into individual distributions f(k), as shown in Fig. 2.

(Step 3) Compute LSEF by using statistical quantities k,Z ,$U(k,Z_K)$ for each distribution f(k)(k=1,..n).

(Step 4) Repeat the LSEF computation, for all of the candidate pitch P(Ui,j) and each boundary point S(f(k),f(k+1)).

In step 4, a best linear unbiased estimator $\hat{P}$ is obtained as a candidate pitch P(Ui,j) ,which has minimum LSEF value. However, in order to use the above pitch estimation algorithm, three constraints described below, must be taken into account.

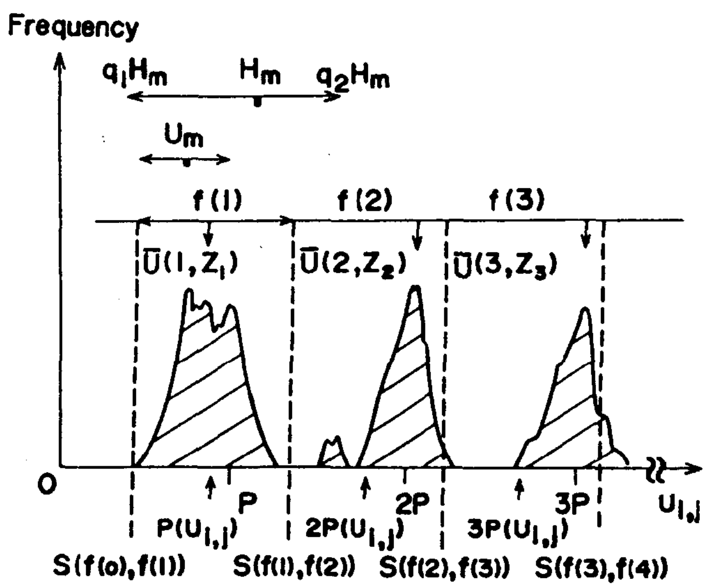


Figure 2 Distance $U_{i,j}$ clustering

Candidate pitch P(Ui,j) range

Candidate pitch P(Ui,j) range can be predicted by using a mean character height Hm value, such as

$$q_1 \cdot Hm < P(Ui,j) < q_2 \cdot Hm \qquad (0 < q_1 < q_2) \qquad (7)$$

In addition, the above mentioned range will be limited by detecting a distance Um having the most frequent distance Ui,j in order to reduce the computation time.

$$MAX (q_1 \cdot Hm, Um - q_3) < P(Ui,j) < MIN (q_2 \cdot Hm, Um + q_3) \quad (0 < q_3 < Um) \qquad (8)$$

Boundary point S(f(k),f(k+1)) assumption

Individual boundary points S(f(k),f(k+1)) exist between $k \cdot P(Ui,j)$ and $(k+1) \cdot P(Ui,j)$, once a candidate pitch P(Ui,j) is determined. However, the computation time required for detecting the best boundary points is increased. The reason is that the best boundary points must be obtained by moving them independently from each other and computing LSEF . Therefore, it is assumed that each boundary point S(f(k),f(k+1)) is a middle point between $k \cdot P(Ui,j)$ and $(k+1) \cdot P(Ui,j)$, based on practical considerations.

Distance Ui,j observation

Generally, character pitch can be exactly estimated, according to an increase in the sample size for distance Ui,j. It is possible that distances Ui,j (j=i+1,i+2,i+3,...) are used in order to increase the sample size. On the other hand, if the sample size is small, punctuation marks, such as comma and period, or large spaces influence the pitch estimation performance. Accordingly, if these regions were to be detected by a simple method, then pitch estimation could be processed more accurately.

Pitch discrimination

Consider the pitch discrimination when dealing with both constant pitch and variable pitch characters. As the result of the above mentioned method used for pitch estimation, the character pitch can be obtained as best linear unbiased estimator $\hat{P}$, if a line image is composed of constant pitch characters. At the same time, distance Ui,j distribution can be broken down into individual distributions f(k) (k= 1,2,3,..). Each variance $V^2(k)$ for divided distribution f(k) is considered one important factor for understanding characteristics of printed characters. In the case of constant pitch, each variance $V^2(k)$ is regarded as a random interference around value k.p. It is considered that the variance $V^2(k)$ is constant and independent in regard to character number k. In the case of variable pitch, individual variances $V^2(k)$ are not constant and are dependent on character number k. Therefore, printed characters can be classified as constant and variable by using the discrimination function (DF), as follows.

$$DF = \text{square root} \left\{ 1/\hat{P}^2 \cdot \sum_{K=1}^{n} C(k,Z_k) \cdot V^2(k) \right\} \qquad (9)$$

where

$$C(k,Z_k) = k^2 \cdot Z_k \; / \sum_{K=1}^{n} k^2 \cdot Z_k$$

When satisfying inequality $DF < \alpha$ ($0<\alpha<1$, α is a threshold value), the line images can be handled as constant pitch and the pitch based character segmentation can be applied. On the other hand, if DF is larger than α, the line images must be regarded as variable pitch. Accordingly, by inequality DF, a character segmentation method is enabled which adapts to the property of printed characters.

Character sectioning decision

In order to decide character sectioning positions, the Dynamic Programing (DP) method with minimum variance criterion is introduced. The effective scope in which this can be applied will be discussed.

Minimum variance criterion

A pitch based character segmentation can be basically processed by using the spaces and the estimated character pitch $\hat{P}$. In Fig.3, let x(k,i) ($i = 1,..,h_K$) denote the i-th candidate position in stage k (k=0,..,n). The distance between two candidate positions x(k,i), x(k+1,j)($j = 1,..,h_{K+1}$) is denoted by di,j(k,k+1). The mean value or the variance for m-r+1 distances ,which are measured between stage r and stage m ($r < m$), are denoted by M(r,m;i,j), $V_d^2(r,m;i,j)$, respectively. According to the above definitions, the criterion for deciding character sectioning positions is introduced as follows.

$$F(r,m;i,j) = (1-q(DF)) \cdot V_d^2(r,m;i,j) + q(DF)(M(r,m;i,j) - \hat{P})^2 \qquad (10)$$

In the above criterion, q(DF) is a coefficient satisfying $0 < q(DF) < 1$. the first item is represented as the variance for m-r+1 distances di,j(k,k+1)(k= r,..m-1). The second item is represented as the control item to select a series of candidate positions which locate near the multiple position of the estimated pitch $\hat{P}$. If the above mentioned DF value satisfying inequality $DF < \alpha$ is larger , then the criterion F(r,m;i,j) should reduce the influence on the second item by adjusting coefficient q(DF). For computing a series of candidate positions from a line image by using criterion F(r,m;i,j), it is obviously necessary to determine the stages able to include the candidate positions.

Stage determination

As shown in Fig.3, all of the regions in a line image are classified into two categories, called admissible or prohibitive region. Admissibility for the region is determined as follows. Any space region is regarded as an admissible region. The black region is regarded as a prohibitive region, when black region width Bi satisfies inequalty $Bi < (1+ q_x) \cdot \hat{P}$. Parameter q_x ($0<q_x<1$) is regarded as the value depending on DF value. On the other hand, the black region is interpreted as touching characters when black region width Bi satisfies inequality $Bi \geq (1+ q_x) \cdot \hat{P}$. In such a black region, the two subregions in the neighorhood of the black region edges are regarded as prohibitive and the remaining subregion is regarded as admissible. (See Fig.3) This constraint is to exclude the possibility that the character sectioning positions are determined near the left or right edges of the black region. The conditions concerned with the stage described below can be effectively operated in the admissible regions.

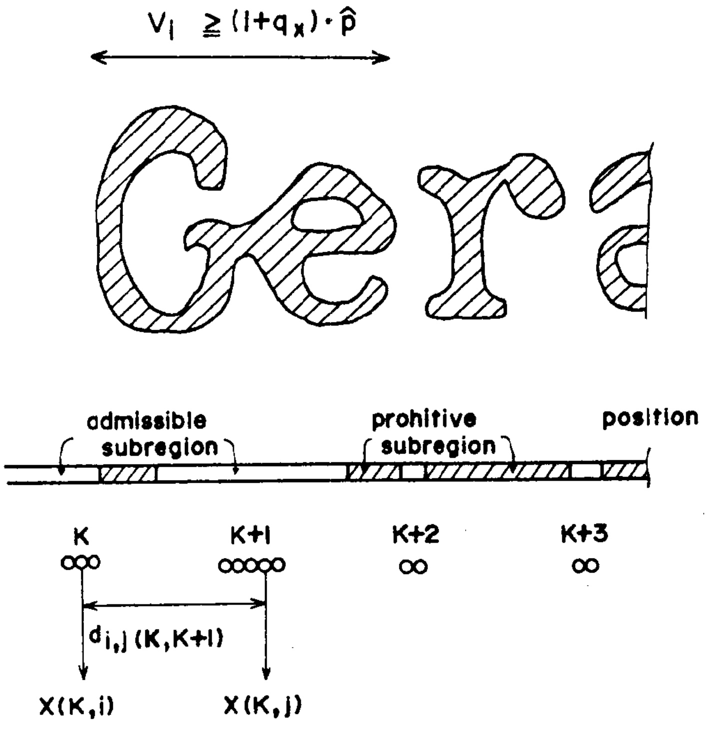


Figure.3 Stage determination

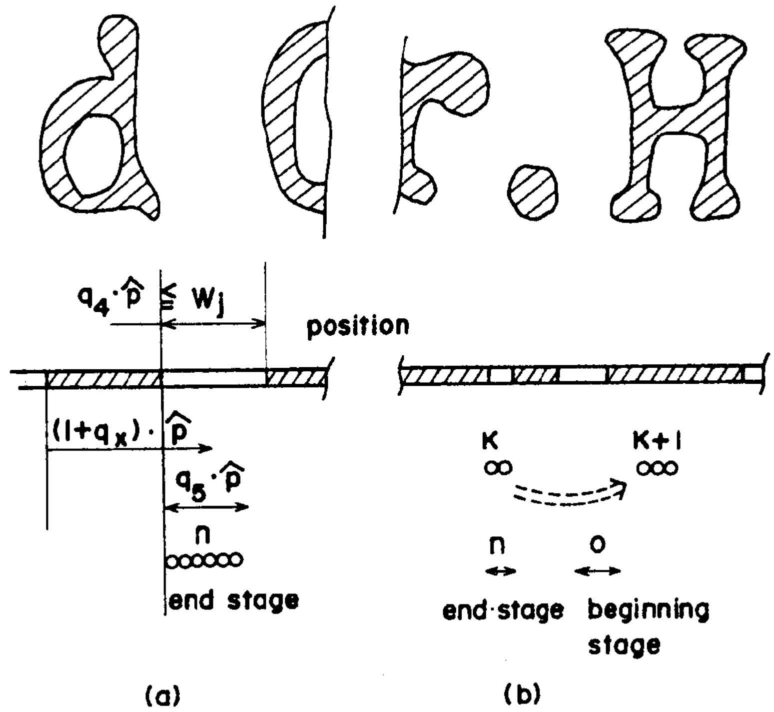


Figure.4 End stage n detection

Beginning stage condition

Once the end stage described below is determined in a line image, the beginning stage is automatically determined by using the space.

Neighboring stages condition

The relation between the two candidate positions x(k,i) and x(k+1,j) is given by

$$| x(k+1,i) - x(k,j) - \hat{p} | < q_x \cdot \hat{p} \qquad (11)$$

End stage condition

Two kinds of end stages k(=n) are detected in the space region. One is determined by estimated pitch P before the evaluation process for deciding the character sectioning positions. (See Fig.4-a.) Another is determined during the evaluation process. In such a case, it is considered that the constant pitch property on the line image is partially broken. (See Fig.4-b.)

In the evaluation process with criterion F(r,m;i,j), if the candidate positions enter into an end stage n, a candidate position, which has the minimum criterion F(0,n;i,j) value, is selected out of the candidate positions able to enter into an end stage n. As a result, the end position x(n,i) for the last character is decided. On the other hand, if stage k+1 is determined by using inequality (11) and is included in a prohibitive region, stage k is handled as an end stage n due to disturbing the constant pitch property.

Optimal path computation using DP

Character sectioning positions are obtained by using the DP algorithm with the above criterion F(r,m;i,j). For any candidate position x(k+1,j)($j=1,..,h_{k+1}$)in stage k+1, an optimal candidate position $x^*(k,i)$,which is connected with the candidate position x(k+1,j), is obtained by recursive equations described below.

$$d_{i,j}(k,k+1) = x(k+1,j) - x(k,i) \quad (12)$$

$$S(k+1,j) = S^{*}(k,i) + d_{i,j}(k,k+1) \quad (13)$$

$$M(0,k+1;t,j) = \frac{1}{K+1}\left\{k \cdot M^{*}(0,k;t,i) + d_{i,j}(k,k+1)\right\} \qquad (t = 1,..,h_0) \quad (14)$$

$$F(0,k+1;t,j) = (1-q(DF)) \cdot \left\{\frac{1}{K+1} \cdot S(k+1;j) - M^{2}(0,k+1;t,j)\right\} + q(DF) \cdot \left\{M(0,k+1;t,j) - \hat{P}\right\}^{2} \qquad (t = 1,..,h_0) \quad (15)$$

By using Eqs.(12) - (15), an optimal candidate position $x^{*}(k,i)$ is as follows.

$$x^{*}(k,i) = \underset{i \in (1,\cdot\cdot,h_K)}{ARG}\left[\,Min\ F(0,k+1;t,j)\right] \quad (16)$$

Optimal candidate position $x^{*}(k,i)$, optimal criterion value $F^{*}(0,k+1;t,j)$, optimal mean $M^{*}(0,k+1;t,j)$ and optimal squared summation $S^{*}(k+1,j)$ are stored in any candidate position x(k+1,j) . The above mentioned recursive equations are used to calculate whole candidate positions x(k,i) (k=0,..,n, $i = 1,..,h_K$) from beginning stage 0 to end stage n. As a result, character sectioning positions are obtained as an optimal path, which has minimum F(0,n;i,j) value.

Scope limitations for adapting DP with minimum variance criterion

If the above mentioned DF value, used to classify printed characters, is larger than a threshold value , DP with minimum variance criterion will not be applied to decide the character sectioning positions. For instance, printed variable pitch characters or handwritten characters , which do not have the character frame constraint, cannot be used. Generally, the character segmentation problem for variable pitch will be considered as a character recognition problem, in the presence of split or merged characters . However, in economical practice , a simple character segmentation method using the spaces is applied instead of using the more complicated segmentation method. It is based on the assumption that printed quality for variable pitch is higher than the printed quality for constant pitch.

Experiments

Actual mail addresses were scanned by CCD image scanner (75μm resolution). The CCD image scanner output is an analog signal corresponding to the amount of light reflected from each pixel on the source lines. Each pixel in a line image is converted into a binary value (black or white). 115 samples were chosen from the actual mail samples to use as experimental data . Address and name lines detection and skew detection were performed on the basis of the partial horizontal histograms. Table 1 shows the experiment data properties.

	Constant	Variable
Mail piece number	65	50
Word number	790	552
Character number	4907	4399
Total merged characters	2004 40.8%	64 1.5%
Mean character number per mail item	80.9	

	Constant	Variable
2 merged characters	613 (25.0%)	13 (0.6%)
3 merged characters	113 (6.9%)	6 (0.4%)
4 merged characters	54 (5.3%)	2 (0.2%)
5 merged characters	6 (0.7%)	2 (0.2%)
over 6 merged characters	26 (4.7%)	0 (0 %)

Table 1 experimental data properties

Clustering examples for distance $U_{i,j}$ distribution are shown in Fig.5. The values for DF are shown in Fig. 6. From Fig.6, Constant pitch data can be distinguish from variable pitch data in DF value 20%.

In the character sectioning experiment, two methods were used acoording to the character pitch estimation result. If DF for the mail sample is less than 20 %, the DP method with minimum variance criterion is used. On the other hand , if DF is larger than 20 %, a simple method using the spaces is applied. The correct segmentation rate in 115 pieces of mail was 99.3 % (per character) and 98.5 % (per word). Some character segmentation results are shown in Fig.7. In Fig.7, vertical lines represent both character image edges. Specifically, dotted vertical lines are sectioning lines which separate touching characters.

Conclusion

The character segmentation problem under more general conditions can be formulated and classified as pitch estimation and character sectioning decision. In order to resolve the character segmentation problem, the statistical analysis method with linear square error function and DP method with minimum variance criterion for the distance between the candidate positions in a line image are introduced. The effectiveness of these methods has been evaluated through an actual mail address segmentation experiment.

These methods have the following three advantages.

(1) Best linear unbiased estimator for pitch estimation can be obtained by the statistical analysis method for the distance $U_{i,j}$ with criterion LSEF.

(2) The property of printed characters can be determined by variance discrimination DF. Accordingly, the adaptive strategy for character sectioning decision can be applied.

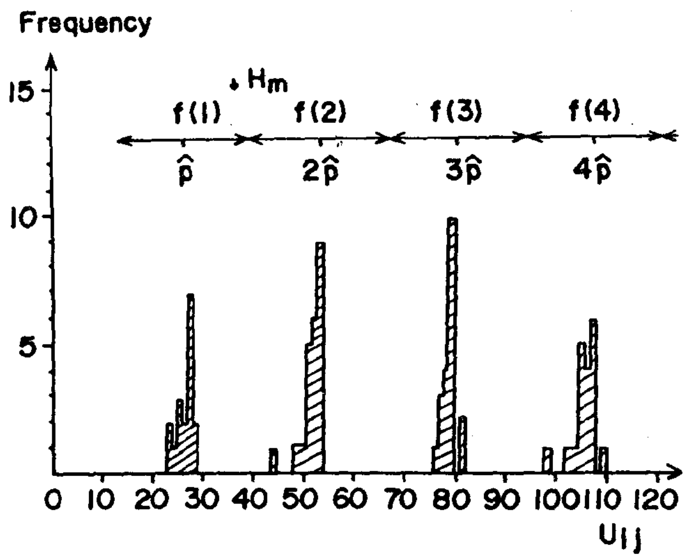
(3) The DP method, with minimum variance criterion, can be used as the pitch based character segmentation in order to absorb some local fluctuations in characters.

Acknowledgement

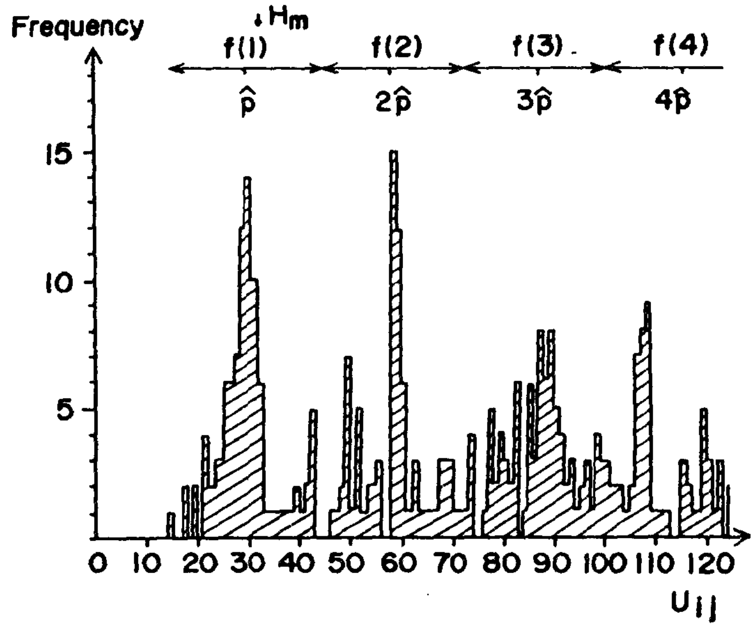
The authors would like to thank Mr. Hanaki and Mr. Miura of NEC Corporation for their guidance and encouragement.

References

1. R.L Hoffman, J. W. Mccullough:" Segmentation Method For Recognition of Machine-Printed Character", IBM J. Res & Dev, Vol 15, March 1971.
2. R.G.Casey, C.R.Jih:" A Processor-Based OCR System", IBM J. Res & Dev,Vol 27, July 1983.
3. R.kooi,W.Lin: "An On-line Mini Computer-based System For Reading Printed Text Aloud", IEEE Trans vol SMC-8 January 1978.
4. V.A.Kovalevsky:" Image Pattern Recognition", Springer-Verlag.
5. R.G.Casey, G.Nagy:" Recursive Segmentation and Classification of composite character patterns ",6th IJCPR 1982 pp.1023-1026
6. J.Schurman: "Reading Machines", 6th IJCPR 1982 pp.1031-1044
7. Y.Tsuji, K.Asai: " Character Segmentation", paper of Technical Group on Pattern Recognition and Learning, PRL 83-66(Japan. 1984) (In Japanese)



(a) Constant pitch



(b) Variable pitch

Figure. 5 Examples showing Distribution of Distances $U_{i,j}$ (j = i+1, i+2, i+3)

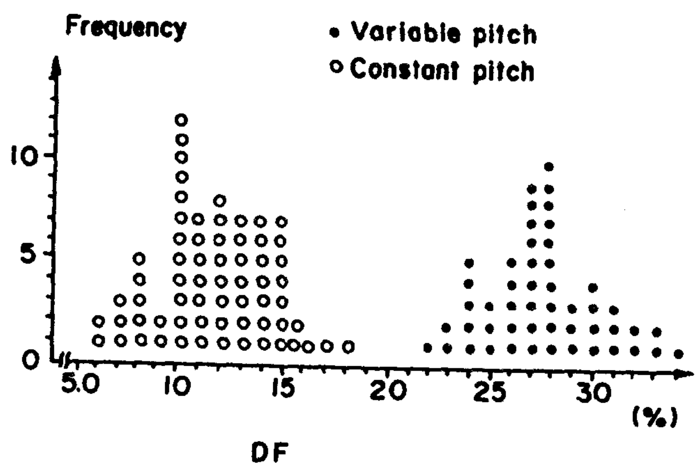
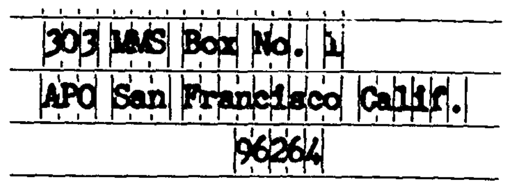
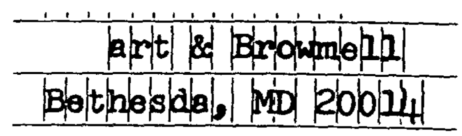


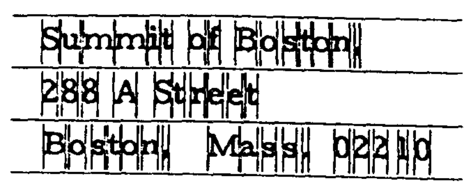
Figure 6 DF Value for Pitch Estimation



(a) Constant pitch



(b) Constant pitch



(c) Variable pitch

Figure 7 Character segmentation results

MODEL BASED SEGMENTATION OF FLIR IMAGES

Bir Bhanu
Department of Computer Science, University of Utah
Salt Lake City, UT 84112
and
R.D. Holben
Ford Aerospace & Communications Corporation
Newport Beach, CA 92660

Abstract

In the automatic recognition of tactical targets in FLIR images, it is desired to obtain an accurate and precise representation of the boundary of the targets. It is very important since the features used in the classification of the target are normally based on the shape and gray scale of the segmented target and therefore the performance of a statistical or a structural classifier critically depends on the results of segmentation. Generally, only the gray scale of the image is used to extract the target from the background. The segmentation thus obtained normally depends upon several parameters of the technique used. It is possible to obtain better segmentation by using other sources of information present in the image such as contextual cues, temporal cues, gradient, a priori information etc. In this paper we consider specifically the use of gray scale together with the edge information present in the image to obtain more precise segmentation of the target than obtained by using gray scale or edge information alone. A model of FLIR images based on gray scale and edge information is incorporated in a gradient relaxation technique which explicitly maximizes a criterion function based on the inconsistency and ambiguity of classification of pixels with respect to its neighbors. Four variations of the basic relaxation technique are considered which provide automatic selection of threshold to segment FLIR images. A comparison of these methods is discussed.

Introduction

FLIR images exhibit thermal radiations emitted by a target. Since the target is in direct contact with its environment, there is an exchange of heat between the target boundary surface and the immediate surroundings (background). A very important reason for the loss of contrast between target and background with range to the observer is due to the atmospheric scatter and absorption. As a result the boundary of the target in the image is not very sharp and there is an intensity gradient across it. For the automatic recognition of tactical targets in FLIR images, it is desired to obtain an accurate and precise representation of the boundary of the target. It is very important since the features used in the classification of the target are normally based on the shape and gray scale of the segmented target and therefore the performance of a statistical or a structural classifier critically depends on the results of segmentation. Generally, only the gray scale of the image is used to extract the target from the background. The segmentation thus obtained normally depends upon several parameters of the technique used. The problem addressed here is: can a better segmentation be obtained more efficiently than a segmentation obtained by using gray scale alone? It is possible to obtain a better segmentation by using other sources of information present in the image such as contextual cues, temporal cues, gradient, a priori information etc. In this paper we consider specifically the use of gray scale together with the edge information (gradient of the gray scale) present in the image to obtain a more precise segmentation of the target than a segmentation