„Algorithmische Anwendungen“

von

Prof. Dr. Heiner Klocke

Projekt
„Bildsegmentierung“

Gruppenbezeichnung: F_ROT
AI7 – WS05/06
02.02.2006

Namen:

Poltavets, Volodymyr	**11030683**
Wildt Graziani, Eduardo	**11030709**
Kunzonzila Nzuzi, Darcy	**11030725**

Projektthema: Bildsegmentierung

Unsere Aufgabe ist die Entwicklung eines Programms, das verschiedene Objekte auf einem Bild durch Segmentierung erkennt und die über folgende Funktionalität verfügt:

- Laden eines Bildes,
- Bilden des Histogramms,
- Ausgabe des Histogramms mit der Anzahl der Auftritte des jeweiligen Grauwertes,
- Möglichkeit einen Schwellwert einzugeben oder durch das Otsu-Verfahren berechnen zu lassen,
- Ausgabe des segmentierten Bild,
- Ausgabe des berechneten Schwellwerts,
- Ausgabe der Laufzeit des Otsu-Verfahrens und der Segmentierung.

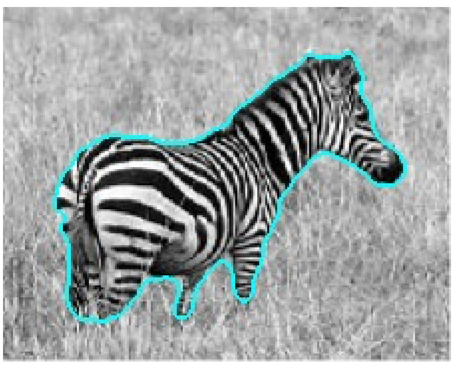
Definition:

Die Segmentierung ist die Aufteilung eines Bildes in Bereiche oder Objekten, die sich aufgrund einheitlicher Merkmale von dem Bildhintergrund und anderen Objekten unterscheiden. Nach der Segmentierung wissen wir, welcher Bildpunkt zu welchem Objekt gehört.

Die Einsatzgebiete von Segmentierungstechniken reichen von Überwachungsaufgaben und Qualitätskontrolle über Auswertung von Bildern aus der Medizin, Vermessung von Objekten und Datenkompression bis hin zu Gest- und Gesichtserkennung, um nur einige zu nennen.

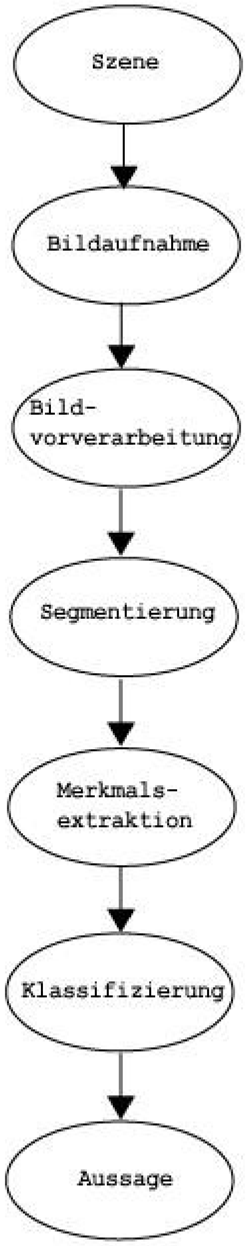
Die Bildsegmentierung ist die Voraussetzung für die vielfältige Weiterverarbeitung von Bildern. Bei der Bildanalyse spielt die Bildsegmentierung eine sehr wichtige Rolle. Denn ein Bild enthält verschiedene Objekte, die einzeln angesprochen und bearbeitet werden sollen.

Bildsegmentierung:



Die Bildanalyse besteht aus mehreren Teilen: Szene, Bildaufnahme, Bildvorverarbeitung, Segmentierung, Merkmalsextraktion, Klassifizierung und Aussage.

Bildanalyse:



Die Szene stellt das reale beobachtete Objekt dar.

Mit geeigneten Sensoren wird nun ein Bild aufgenommen (Fotografie, Radarabtastungen, Röntgen, usw.).

Bei der Bildvorverarbeitung wird die Aufnahme verbessert (zum Beispiel: Helligkeitskorrektur), damit die folgenden Schritte effektiver durchgeführt werden können. Welche Vorverarbeitung günstig ist, hängt es von den konkreten Verfahren in den folgenden Schritten ab.

Im Segmentierungsschritt werden die Pixel des Bildes in zusammengehörige Gruppen (Segmente) eingeteilt.

Bei der Merkmalsextraktion werden nun für diese Segmente bestimmte Eigenschaften festgelegt, beispielsweise Fläche, Exzentrizität der Form (eher rund oder eher länglich), mittlerer Farbwert oder andere.

Mithilfe dieser Eigenschaften und einem vorher bestimmten Regelwerk kann nun in der Klassifikation jedes Segment in eine von mehreren Klassen eingeordnet werden.

Durch die Interpretation dieses Ergebnisses kann nun eine Aussage getroffen werden (zum Beispiel: in der Gesichtserkennung (zu wem gehört das Gesicht?), in der Texterkennung (was steht auf dem Bild?), usw.).

Es sind viele Verfahren zur automatischen Segmentierung bekannt. Grundsätzlich werden sie oft in pixel-, kanten-, regionen-, texturorientierten und modellbasierten Verfahren eingeteilt:

- Pixelorientierte Verfahren:
Erkennung eines Objektes durch die Eigenschaft jedes einzelnen Bildpunkt (Pixel), z.B. Schwellwertverfahren

- Kantenorientierte Verfahren:
Suchen im Bild nach Kanten oder Objektübergängen, z.B. Laplace-Operator

- Regionenorientierte Verfahren:
Betrachtung von Punktmengen als Gesamtheit (zusammenhängende Objekte), z.B. Split and Merge

- Modellbasierte Verfahren:
Suchen nach Objekten durch ihr Modell (z.B. Form), z.B. Hough-Transformation

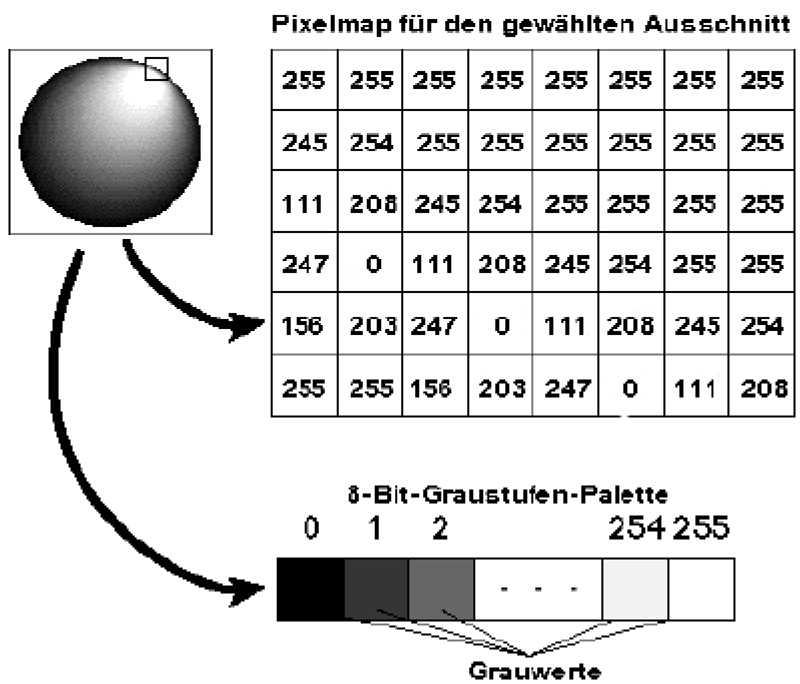
- Texturorientierte Verfahren:
Suchen nach Objekte durch ihre Textur (z.B. Rillen), z.B. Texture-Energy-Measure

Pixelorientierte Verfahren:

In unserem Programm benutzen wir zur Bildsegmentierung ein pixelorientiertes Verfahren.

Pixelorientierte Verfahren treffen für jeden einzelnen Bildpunkt die Entscheidung, ob er zu einem bestimmten Segment gehört oder nicht. Ob ein Pixel dem Objekt oder dem Hintergrund angehört, hängt es von der Eigenschaft des Pixels selbst ab. Diese Eigenschaft kann z.B. der Grauwert sein.

Grauwert eines Bildes:

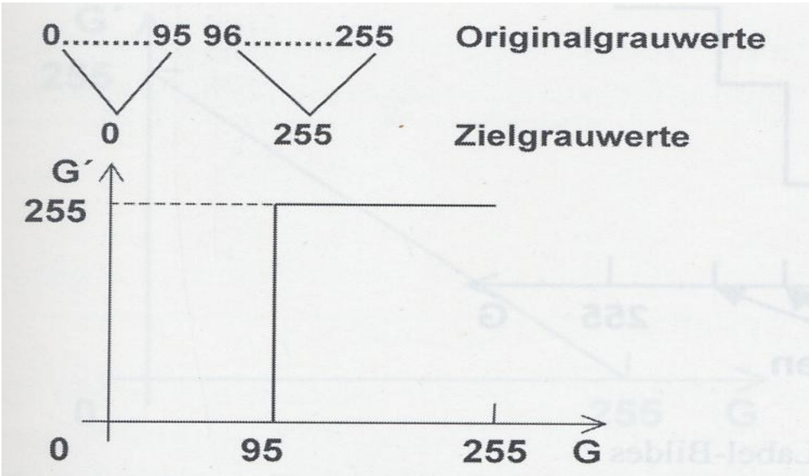


255	255	255	255	255	255	255	255
245	254	255	255	255	255	255	255
111	208	245	254	255	255	255	255
247	0	111	208	245	254	255	255
156	203	247	0	111	208	245	254
255	255	156	203	247	0	111	208

Schwellwertverfahren:

Aus den verschiedenen pixelorientierten Verfahren haben wir uns für das Schwellwertverfahren entschieden. Mit Schwellwert werden die vorgegebenen Zahlen bezeichnet, bei deren Überschreitung sich der Grauwert von schwarz auf weiß ändern.

Schwellwert 95:

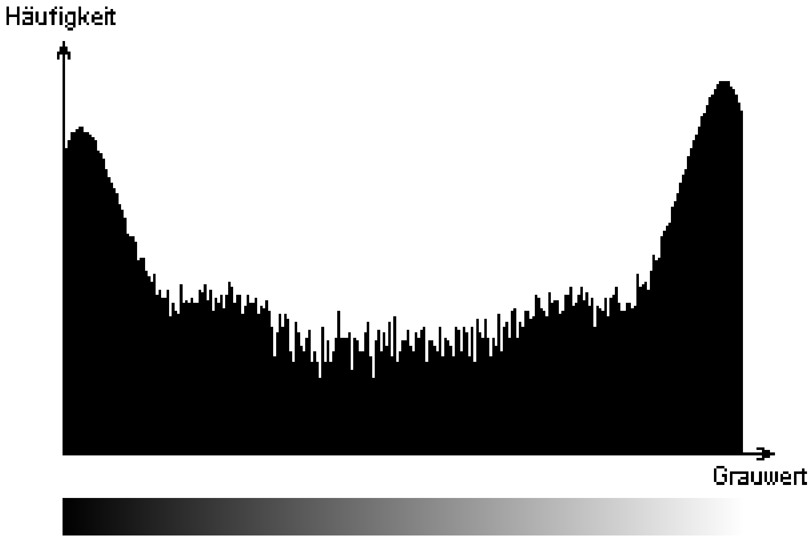


Bei dem Schwellwertverfahren ist das zu segmentierende Bild in Form von Zahlenwerten (Farbwerten pro Bildpunkt) gegeben. Die Zugehörigkeit eines Bildpunktes zu einem Segment wird durch den Vergleich des Grauwertes mit einem Schwellwert entschieden. Der Grauwert eines Pixels ist sein reiner Helligkeitswert, ohne Berücksichtigung weiterer Farbinformationen. Da diese Operation für jeden Pixel einzeln angewendet wird, ist das Schwellwertverfahren ein pixelorientiertes Segmentierungsverfahren.

Das grundlegende Prinzip des Schwellwertverfahrens kann auf verschiedene Art angewendet werden. Beim globalen Schwellwertverfahren wird der Schwellwert (oder die Schwellwerte) global für das gesamte Bild gewählt. Dieses Verfahren ist am einfachsten, ist jedoch auch sehr anfällig für Helligkeitsveränderungen im Bild. Beim lokalen Schwellwertverfahren wird das Ursprungsbild in Regionen eingeteilt und der Schwellwert für jede Region neu gewählt. Als Weiterentwicklung lässt sich das dynamische Schwellwertverfahren ansehen, das für jeden Bildpunkt eine Umgebung festlegt, und auf dieser Umgebung den passenden Schwellwert aussucht.

Bei der statistischen Analyse eines Bildes hilft das zugehörige Histogramm. Das Histogramm wird bestimmt, indem man die Häufigkeit jedes einzelnen Farbwertes bestimmt und durch eine entsprechend hohe Linie angibt.

Histogramm:



Als Legende ist daher unten ein Balken mit den verschiedenen Grauwerten zu sehen, darüber ist dann jeweils durch die Höhe der Linie die Anzahl der Auftritte des jeweiligen Grauwertes angegeben.

Verfahren von Otsu:

Primäre Literatur:
[Ots79] Otsu N., "A threshold selection method from gray-level histograms," IEEE Transactions Systems, Man, and Cybernetics, Vol. SMC-9, 1979.

Dieses Verfahren dient zu einer optimalen Ermittlung eines Schwellwerts. Die Einträge des Histogramms werden als Wahrscheinlichkeiten einer diskreten Zufallsvariablen genommen und derjenige Grauwert wird bestimmt, für den sich, nach einer Varianzanalyse, das Histogramm am besten in zwei zusammenhängende Teile separieren lässt.

Es sei *p(g)* die Auftrittswahrscheinlichkeit des Wertes *g*. Für *g* gilt : $0 <= g < G$.(G ist der maximale Grauwert).

K_0 und K_1 seien zwei Klassen (Hintergrund und Objekt) von Punkten, getrennt durch den Schwelltwert *t*. K_0 hat $g = 0,..., t$ und K_1 hat $g = t+1,..., G$. Die Auftrittswahrscheinlichkeit der Klassen bestimmt sich durch:

K_0:

$$P_0(t) = \sum_{g=0}^{t} p(g)$$

K1:

$$P_1(t) = \sum_{g=t+1}^{G} p(g) = 1 - P_0(t)$$

Der mittlere Grauwert innerhalb des gesamten Bildes sei *g*, der Mittelwert der beiden Klassen dementsprechend $\overline{g_0}$ und $\overline{g_1}$

Die Varianzen (Abweichung von jedem Einzelelement von Mittelwert) innerhalb der Klassen ergeben sich durch:

$$\sigma_0^2(t) = \sum_{g=0}^{t} (g - \overline{g_0})^2 p(g)$$

$$\sigma_1^2(t) = \sum_{g=t+1}^{G} (g - \overline{g_1})^2 p(g)$$

Um einen optimalen Schwelltwert zu bekommen , der die beiden Klassen möglichst gut trennt, soll die Varianz zwischen den beiden Klassen maximiert werden.

$$\sigma_{zw}^2(t) = P_0(t) \cdot (\overline{g_0} - \overline{g})^2 + P_1(t) \cdot (\overline{g_1} - \overline{g})^2$$

Und die Varianz innerhalb der Klassen soll minimiert werden:

$$\sigma_{in}^2(t) = P_0(t) \cdot \sigma_0^2(t) + P_1(t) \cdot \sigma_1^2(t)$$

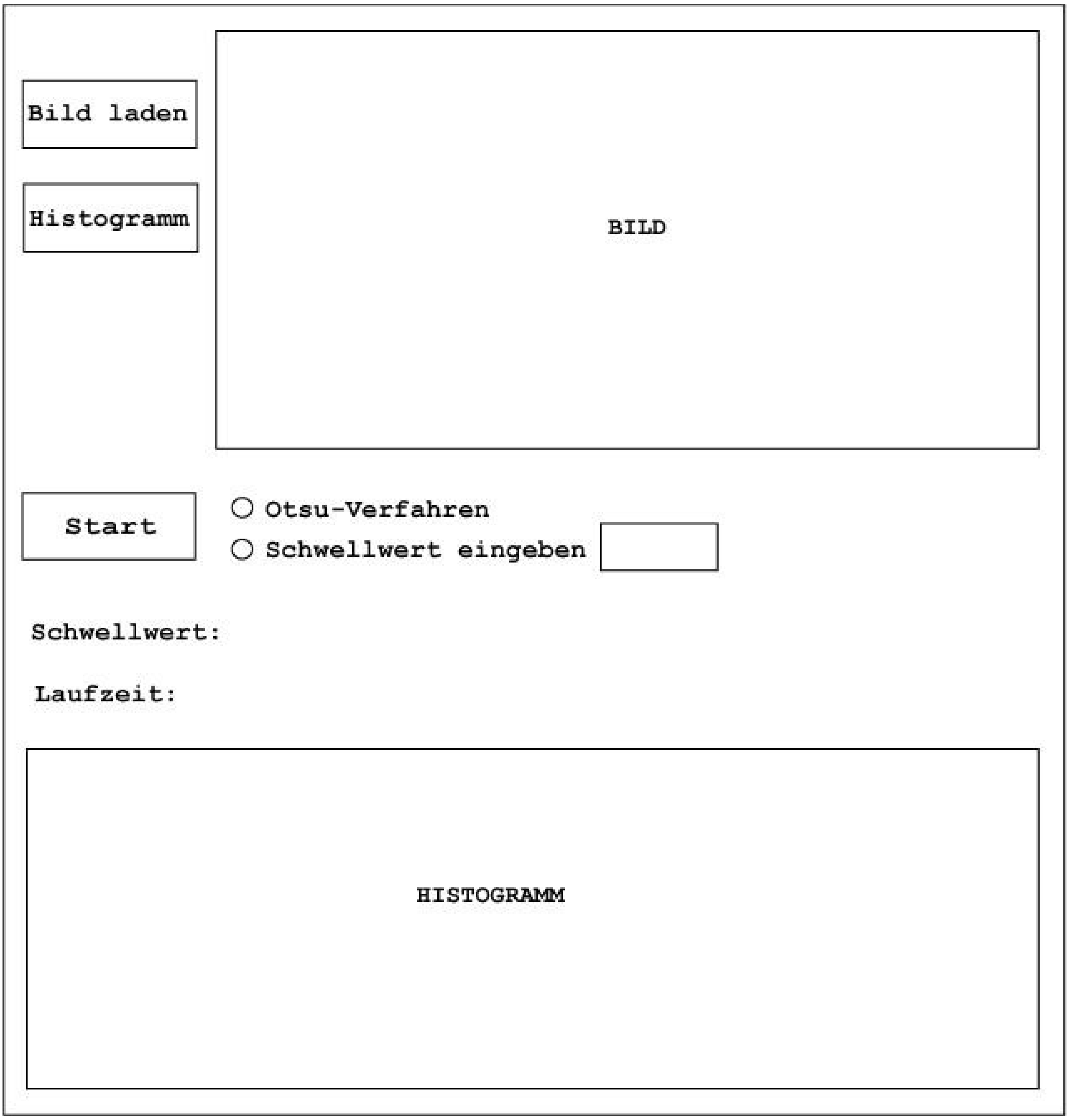
D.h. der Quotient *Q(t)* soll maximal sein.

$$Q(t) = \frac{\sigma_{zw}^2(t)}{\sigma_{in}^2(t)}$$

Programmoberfläche:

Damit wir uns besser vorstellen können, wie das Programm aussehen soll, haben wir eine Programmoberfläche für unser Projekt gezeichnet.

Programmoberfläche:



ImageJ:

ImageJ ist ein sehr weit verbreitetes Java-basiertes Bildverarbeitungsprogramm. Es ist ein Open-Source-Programm, dessen Code auf dem Site http://rsb.info.nih.gov/ij/developer/source/ zu finden ist. Daraus kommt, dass viele Leute Plugins für ImageJ entwickeln.
Es gibt ungefähr 300 Plugins mit verschiedenen Funktionen, die im Programm installiert werden

können. Das macht ImageJ sehr attraktiv. Jeder kann sein eigenes Plugin entwickeln und selber installieren.

Das Programm läuft unter verschiedenen Betriebssystemen, wie Linux, Mac OS 9, Mac OS X, und Windows und es kann auch in unterschiedlichen Dateienformate (TIFF-, GIF-, JPEG-, BMP-, PNG- und PGM-Dateien) speichern und öffnen.

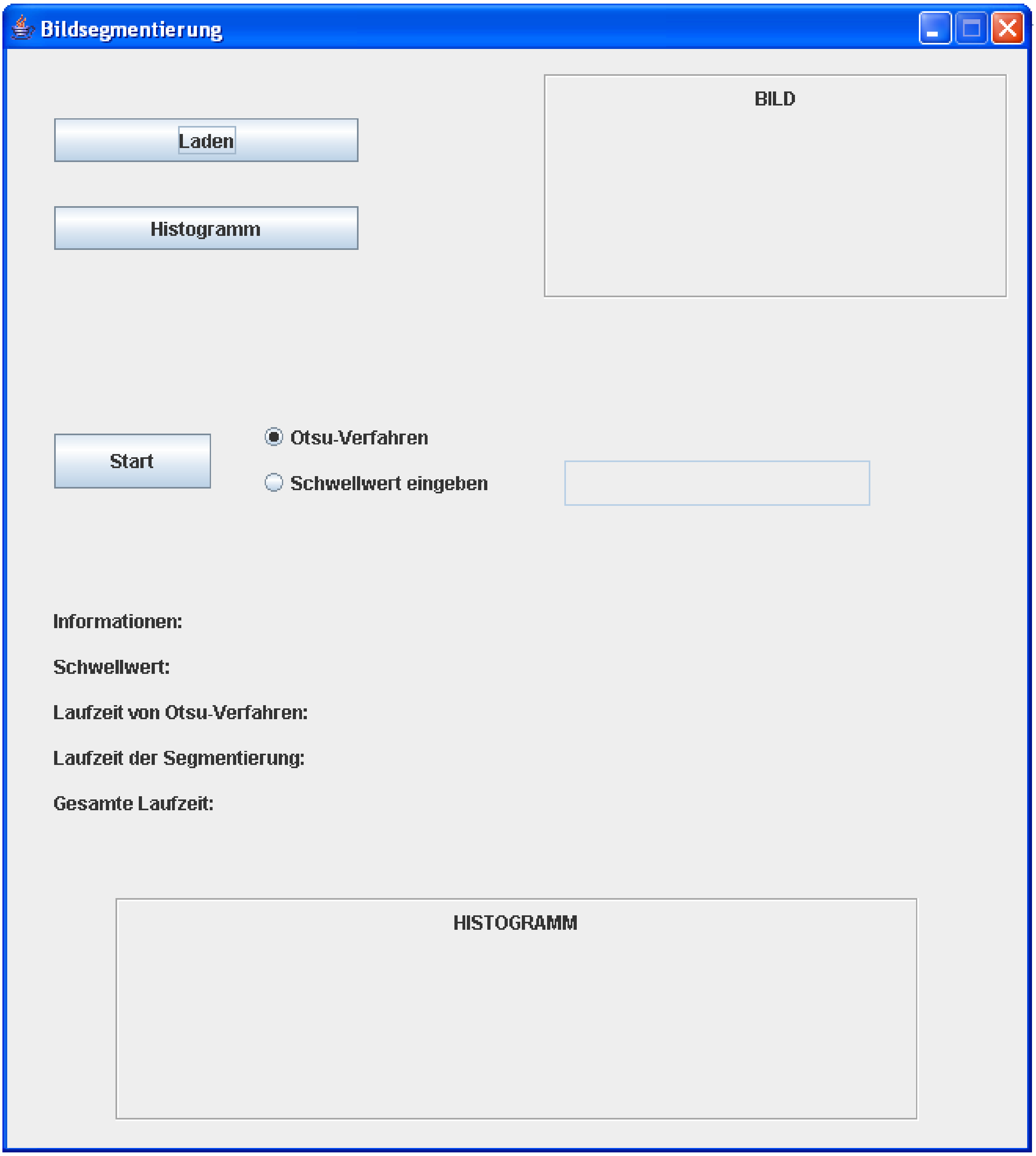
ImageJ ist das schnellste Java-Programm für Image Processing. Es kann ein 2048x2048 Bild in 0.1 Sekunde filtern, es sind 40 Millionen Pixels pro Sekunde.

Aus dem Programm möchten wir für unser Projekt einige Klassen übernehmen:

- class ImageStatistics, aus dem Package ij.process:
  Sie dient zur Erstellung von Statistiken, wie zum Beispiel das Histogramm eines Bildes.

- class FileOpener, aus dem Package ij.io:
  Diese Klasse lädt ein Bild aus einem Verzeichnis und erzeugt daraus ein FileInfo-Objekt.

Betriebsmodell der Oberfläche:



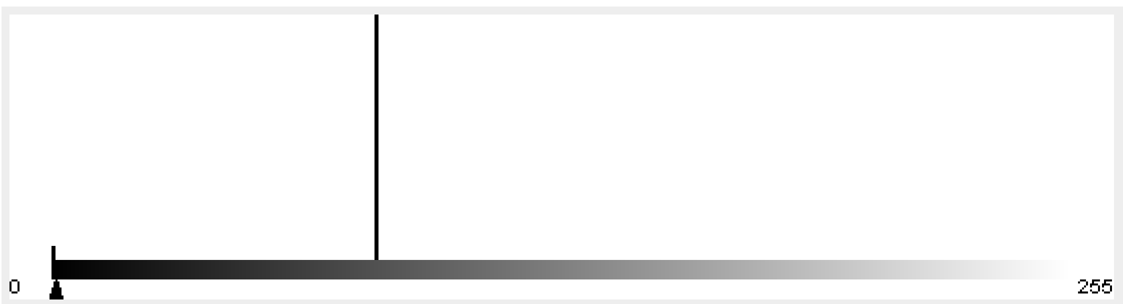
Wir haben bei der Ausgabe von Informationen über die Laufzeit außer Schwellwert und Laufzeit von Otsu-Verfahren, noch dazu Laufzeit der Segmentierung und Gesamtlaufzeit eingefügt. Laufzeit der Segmentierung ist damit gemeint, das pixelorientierte Verfahren, das bei der gesamten Laufzeit mit Laufzeit von Otsu-Verfahren addiert wird.

Durch die Eingabe des Schwellwertes vom Benutzer soll es deutlich werden, dass man keine gute Segmentierung von Bildern bekommt, wenn der Schwellwert nicht optimal ausgewählt wird. In diesem Fall wird keine Methode aufgerufen, um einen Schwellwert zu berechnen, sondern wird das pixelorientierte Verfahren direkt ausgeführt.

Wir haben einige Beispiele von Bildern und ihren entsprechenden Histogramms analysiert:

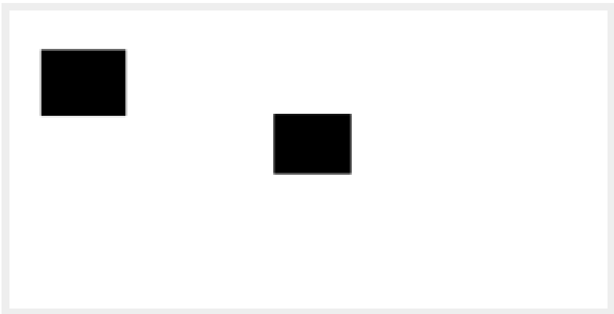
Bild 1:

Histogramm von Bild 1:



Schwellwert von Otsu-Verfahren: 1

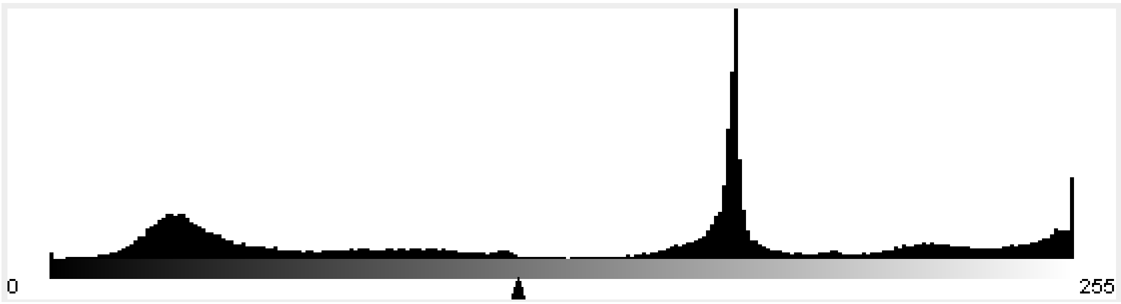
Segmentierung von Bild 1:



Das Bild 1 ist die einfachste Möglichkeit, Objekt auf einem Bild zu erkennen, denn es besteht aus nur zwei verschiedenen Grauwerten. Die beiden schwarzen Quadrate auf dem Bild entsprechen den Grauwert 0 und der Hintergrund den Grauwert 81.

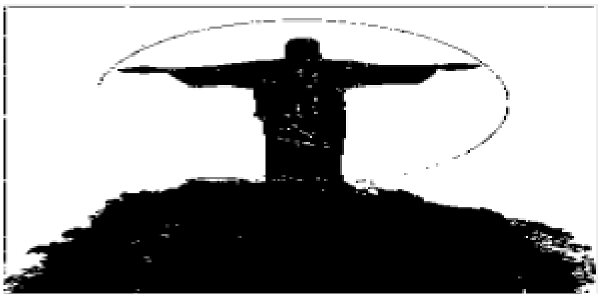
Bild 2:

Histogramm von Bild 2:



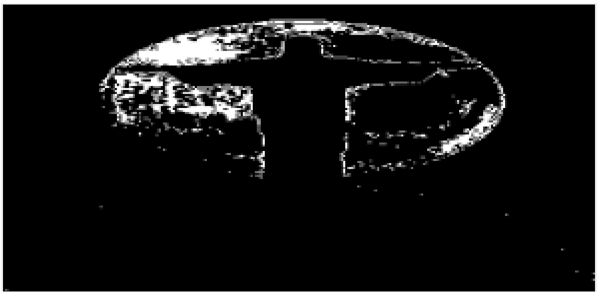
Schwellwert von Otsu-Verfahren: 117

Segmentierung von Bild 2:



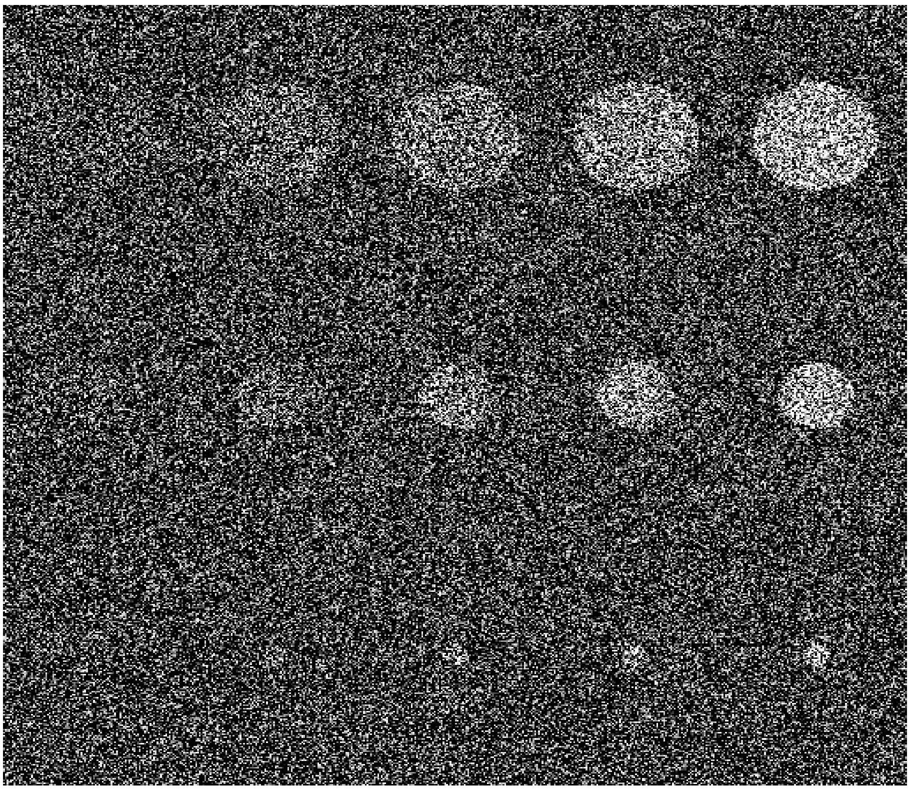
Segmentierung von Bild 2 mit dem Schwellwert 10:

Segmentierung von Bild 2 mit dem Schwellwert 250:



Das Bild 2 ist der normale Fall. Das Bild hat mehrere Grauwerten, deshalb ist ein Erraten des Schwellwertes sehr schwierig.
Mit der Anwendung des Schwellwertes von Otsu-Verfahren haben wir ein gutes Ergebnis der Segmentierung erzielt. Wenn man aber andere Werte als Schwellwert nimmt, bekommt man kein gutes Resultat. Auf den beiden letzten segmentierten Bildern ist kaum etwas zu erkennt.

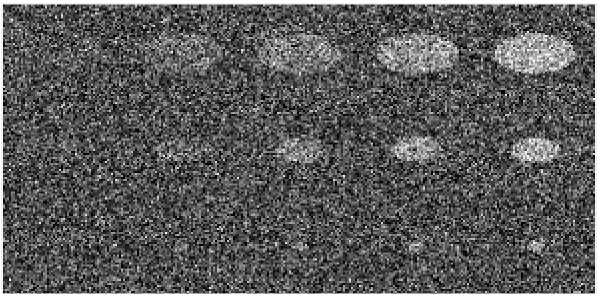
Bild 3:



Histogramm von Bild 3:

Schwellwert von Otsu-Verfahren: 96

Segmentierung von Bild 3:



Das dritte Beispiel ist der schwierigste Fall, ein Bild mit dem Schwellwertverfahren zu segmentieren. Durch den ganzen Rausch ist kaum etwas zu erkennen. Nach der Segmentierung wurde nichts erkannt, deswegen ist zwischen dem segmentierten Bild und dem normalen Bild kein Unterschied zu sehen.

Programmlaufzeit:

Bei der Berechnung der Laufzeit muss es berücksichtigt werden, dass im Programm zwei algorithmische Verfahren benutzt werden. Das Otsu-Verfahren, um den Schwellwert zu berechnen und das pixelorientierte Verfahren, das das Bild segmentiert.

Otsu-Verfahren:

Außer der mathematischen Berechnungen, die bei der Laufzeit keine Rolle spielen, haben wir bei dem Otsu-Verfahren eine for-Schleife, um den Schwellwert zu suchen.

Deswegen hat dieses Verfahren eine Θ-Notation von **$\Theta(n)$**.

Pixelorientiertes Verfahren:

Bei diesem Verfahren muss jeder einzelne Pixel mit dem Schwellwert verglichen werden. Und weil ein Bild Breite und Länge (wie eine Matrix) hat, braucht man zwei for-Schleife, um einmal über alle Pixels durchzulaufen.

Dieses Verfahren hat eine Θ-Notation von **$\Theta(n^2)$**.

Literatur:

1. B. Neumann, „Bildverarbeitung für Einsteiger“, Springer, 2004
2. W. Burger, M. J. Burge, "Digitale Bildverarbeitung", Springer, 2005.
3. A. Janser, W. Luther, W. Otten, "Computergraphik und Bildverarbeitung", Vieweg, 1996.
4. S. Osher, R. Fedkiw, "Level set methods and dynamic implicit surfaces", Springer, 2003.
5. S. Osher, N. Paragios, "Geometric level set methods in imaging, vision, and graphics ", Springer, 2003.
6. P. Soille "Morphological image analysis", Springer, 2003.

Internetquellen:

http://mbi.dkfz-heidelberg.de/
http://www.ruhr-uni-bochum.de/
http://www.on-design.de/

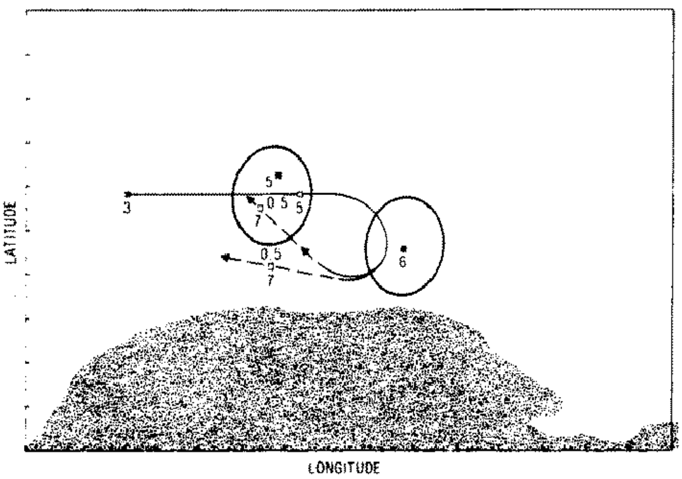


Fig. 6. ID plot of ship 10001 after the second round of operator-imposed assignment constraints.

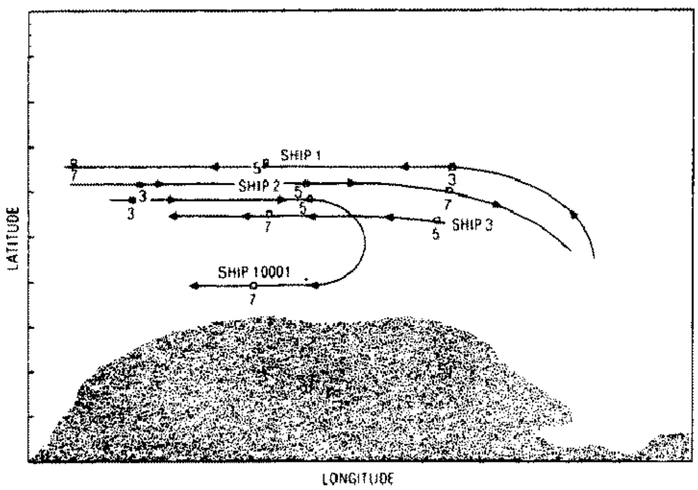


Fig. 7. Actual ship movements.

of the two last sighted locations. The true trajectories are shown in Fig. 7 where it can be seen that ship 10001 did, in fact, turn toward the coast.

IV. Concluding Remarks

The procedure of ship identification from DF sightings has been oversimplified in this discussion. Often DF sightings are not completely identified but, instead, contain only ship class information. The interactive technique still applies, but additional identification and display flexibility must be provided.

Any additional information contained in the sightings can be used to discriminate among radar and DF sightings. Factors such as measured heading and visual ID will permit further automatic reduction of the P and Q matrices.

It is also possible to automate some of the more routine manual functions. However, experience has shown that better results are obtained by having a human operator resolve ambiguous situations arising from sparse data.

References

[1] R. W. Sittler, "An optimal data association problem in surveillance theory," *IEEE Trans. Mil. Elect.*, vol. MIL-8, pp. 125–139, 1964.
[2] M. S. White, "Finding events in a sea of bubbles," *IEEE Trans. Comput.*, vol. C-20 (9), pp. 988–1006, 1971.
[3] A. G. Jaffer and Y. Bar-Shalom, "On optimal tracking in multiple-target environments," *Proc. of the 3rd Sym. on Nonlinear Estimation Theory and its Applications*, San Diego, CA, Sept. 1972.
[4] P. Smith and G. Buechler, "A branching algorithm for discrimination and tracking multiple objects," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 101–104, 1975.
[5] D. L. Alspach, "A Gaussian sum approach to the multitarget-tracking problem," *Automatica*, vol. 11, pp. 285–296, 1975.
[6] C. L. Morefield, *Application of 0-1 Integer Programming to a Track Assembly Problem*, TR-0075(5085-10)-1, Aerospace Corp. El Segundo, CA, Apr. 1975.
[7] D. B. Reid, *A Multiple Hypothesis Filter for Tracking Multiple Targets in a Cluttered Environment*, LMSC-D560254, Lockheed Palo Alto Research Laboratories, Palo Alto, CA, Sept. 1977.
[8] P. L. Smith, "Reduction of sea surveillance data using binary matrices," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6 (8), pp. 531–538, Aug. 1976.

A Threshold Selection Method from Gray-Level Histograms

NOBUYUKI OTSU

***Abstract*—A nonparametric and unsupervised method of automatic threshold selection for picture segmentation is presented. An optimal threshold is selected by the discriminant criterion, namely, so as to maximize the separability of the resultant classes in gray levels. The procedure is very simple, utilizing only the zeroth- and the first-order cumulative moments of the gray-level histogram. It is straightforward to extend the method to multithreshold problems. Several experimental results are also presented to support the validity of the method.**

I. Introduction

It is important in picture processing to select an adequate threshold of gray level for extracting objects from their background. A variety of techniques have been proposed in this regard. In an ideal case, the histogram has a deep and sharp valley between two peaks representing objects and background, respectively, so that the threshold can be chosen at the bottom of this valley [1]. However, for most real pictures, it is often difficult to detect the valley bottom precisely, especially in such cases as when the valley is flat and broad, imbued with noise, or when the two peaks are extremely unequal in height, often producing no traceable valley. There have been some techniques proposed in order to overcome these difficulties. They are, for example, the valley sharpening technique [2], which restricts the histogram to the pixels with large absolute values of derivative (Laplacian or gradient), and the difference histogram method [3], which selects the threshold at the gray level with the maximal amount of difference. These utilize information concerning neighboring pixels (or edges) in the original picture to modify the histogram so as to make it useful for thresholding. Another class of methods deals directly with the gray-level histogram by parametric techniques. For example, the histogram is approximated in the least square sense by a sum of Gaussian distributions, and statistical decision procedures are applied [4]. However, such a method requires considerably tedious and sometimes unstable calculations. Moreover, in many cases, the Gaussian distributions turn out to be a meager approximation of the real modes.

In any event, no "goodness" of threshold has been evaluated in

Manuscript received October 13, 1977; revised April 17, 1978 and August 31, 1978.
The author is with the Mathematical Engineering Section, Information Science Division, Electrotechnical Laboratory, Chiyoda-ku, Tokyo 100, Japan.

most of the methods so far proposed. This would imply that it could be the right way of deriving an optimal thresholding method to establish an appropriate criterion for evaluating the "goodness" of threshold from a more general standpoint.

In this correspondence, our discussion will be confined to the elementary case of threshold selection where only the gray-level histogram suffices without other *a priori* knowledge. It is not only important as a standard technique in picture processing, but also essential for unsupervised decision problems in pattern recognition. A new method is proposed from the viewpoint of discriminant analysis; it directly approaches the feasibility of evaluating the "goodness" of threshold and automatically selecting an optimal threshold.

II. Formulation

Let the pixels of a given picture be represented in L gray levels $[1, 2, \cdots, L]$. The number of pixels at level i is denoted by n_i and the total number of pixels by $N = n_1 + n_2 + \cdots + n_L$. In order to simplify the discussion, the gray-level histogram is normalized and regarded as a probability distribution:

$$p_i = n_i/N, \qquad p_i \geq 0, \ \sum_{i=1}^{L} p_i = 1. \tag{1}$$

Now suppose that we dichotomize the pixels into two classes C_0 and C_1 (background and objects, or vice versa) by a threshold at level k; C_0 denotes pixels with levels $[1, \cdots, k]$, and C_1 denotes pixels with levels $[k+1, \cdots, L]$. Then the probabilities of class occurrence and the class mean levels, respectively, are given by

$$\omega_0 = \Pr(C_0) = \sum_{i=1}^{k} p_i = \omega(k) \tag{2}$$

$$\omega_1 = \Pr(C_1) = \sum_{i=k+1}^{L} p_i = 1 - \omega(k) \tag{3}$$

and

$$\mu_0 = \sum_{i=1}^{k} i \Pr(i \mid C_0) = \sum_{i=1}^{k} i p_i/\omega_0 = \mu(k)/\omega(k) \tag{4}$$

$$\mu_1 = \sum_{i=k+1}^{L} i \Pr(i \mid C_1) = \sum_{i=k+1}^{L} i p_i/\omega_1 = \frac{\mu_T - \mu(k)}{1 - \omega(k)}, \tag{5}$$

where

$$\omega(k) = \sum_{i=1}^{k} p_i \tag{6}$$

and

$$\mu(k) = \sum_{i=1}^{k} i p_i \tag{7}$$

are the zeroth- and the first-order cumulative moments of the histogram up to the kth level, respectively, and

$$\mu_T = \mu(L) = \sum_{i=1}^{L} i p_i \tag{8}$$

is the total mean level of the original picture. We can easily verify the following relation for any choice of k:

$$\omega_0 \mu_0 + \omega_1 \mu_1 = \mu_T, \qquad \omega_0 + \omega_1 = 1. \tag{9}$$

The class variances are given by

$$\sigma_0^2 = \sum_{i=1}^{k} (i - \mu_0)^2 \Pr(i \mid C_0) = \sum_{i=1}^{k} (i - \mu_0)^2 p_i/\omega_0 \tag{10}$$

$$\sigma_1^2 = \sum_{i=k+1}^{L} (i - \mu_1)^2 \Pr(i \mid C_1) = \sum_{i=k+1}^{L} (i - \mu_1)^2 p_i/\omega_1. \tag{11}$$

These require second-order cumulative moments (statistics).

In order to evaluate the "goodness" of the threshold (at level k), we shall introduce the following discriminant criterion measures (or measures of class separability) used in the discriminant analysis [5]:

$$\lambda = \sigma_B^2/\sigma_W^2, \qquad \kappa = \sigma_T^2/\sigma_W^2, \qquad \eta = \sigma_B^2/\sigma_T^2, \tag{12}$$

where

$$\sigma_W^2 = \omega_0 \sigma_0^2 + \omega_1 \sigma_1^2 \tag{13}$$

$$\sigma_B^2 = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2 = \omega_0 \omega_1 (\mu_1 - \mu_0)^2 \tag{14}$$

(due to (9)) and

$$\sigma_T^2 = \sum_{i=1}^{L} (i - \mu_T)^2 p_i \tag{15}$$

are the within-class variance, the between-class variance, and the total variance of levels, respectively. Then our problem is reduced to an optimization problem to search for a threshold k that maximizes one of the object functions (the criterion measures) in (12).

This standpoint is motivated by a conjecture that well-thresholded classes would be separated in gray levels, and conversely, a threshold giving the best separation of classes in gray levels would be the best threshold.

The discriminant criteria maximizing λ, κ, and η, respectively, for k are, however, equivalent to one another; e.g., $\kappa = \lambda + 1$ and $\eta = \lambda/(\lambda + 1)$ in terms of λ, because the following basic relation always holds:

$$\sigma_W^2 + \sigma_B^2 = \sigma_T^2. \tag{16}$$

It is noticed that σ_W^2 and σ_B^2 are functions of threshold level k, but σ_T^2 is independent of k. It is also noted that σ_W^2 is based on the second-order statistics (class variances), while σ_B^2 is based on the first-order statistics (class means). Therefore, η is the simplest measure with respect to k. Thus we adopt η as the criterion measure to evaluate the "goodness" (or separability) of the threshold at level k.

The optimal threshold k^* that maximizes η, or equivalently maximizes σ_B^2, is selected in the following sequential search by using the simple cumulative quantities (6) and (7), or explicitly using (2)–(5):

$$\eta(k) = \sigma_B^2(k)/\sigma_T^2 \tag{17}$$

$$\sigma_B^2(k) = \frac{[\mu_T \omega(k) - \mu(k)]^2}{\omega(k)[1 - \omega(k)]} \tag{18}$$

and the optimal threshold k^* is

$$\sigma_B^2(k^*) = \max_{1 \leq k < L} \sigma_B^2(k). \tag{19}$$

From the problem, the range of k over which the maximum is sought can be restricted to

$$S^* = \{k;\ \omega_0 \omega_1 = \omega(k)[1 - \omega(k)] > 0, \qquad \text{or } 0 < \omega(k) < 1\}.$$

We shall call it the effective range of the gray-level histogram. From the definition in (14), the criterion measure σ_B^2 (or η) takes a minimum value of zero for such k as $k \in S - S^* = \{k; \omega(k) = 0$ or $1\}$ (i.e., making all pixels either C_1 or C_0, which is, of course, not our concern) and takes a positive and bounded value for $k \in S^*$. It is, therefore, obvious that the maximum always exists.

III. Discussion and Remarks

A. Analysis of further important aspects

The method proposed in the foregoing affords further means to analyze important aspects other than selecting optimal thresholds.

For the selected threshold k^*, the class probabilities (2) and (3), respectively, indicate the portions of the areas occupied by the classes in the picture so thresholded. The class means (4) and (5) serve as estimates of the mean levels of the classes in the original gray-level picture.

The maximum value $\eta(k^*)$, denoted simply by η^*, can be used as a measure to evaluate the separability of classes (or ease of thresholding) for the original picture or the bimodality of the histogram. This is a significant measure, for it is invariant under affine transformations of the gray-level scale (i.e., for any shift and dilatation, $g_i' = ag_i + b$). It is uniquely determined within the range

$$0 \le \eta^* \le 1.$$

The lower bound (zero) is attainable by, and only by, pictures having a single constant gray level, and the upper bound (unity) is attainable by, and only by, two-valued pictures.

B. Extension to Multithresholding

The extension of the method to multihresholding problems is straightforward by virtue of the discriminant criterion. For example, in the case of three-thresholding, we assume two thresholds: $1 \le k_1 < k_2 < L$, for separating three classes, C_0 for $[1, \cdots, k_1]$, C_1 for $[k_1 + 1, \cdots, k_2]$, and C_2 for $[k_2 + 1, \cdots, L]$. The criterion measure σ_B^2 (also η) is then a function of two variables k_1 and k_2, and an optimal set of thresholds k_1^* and k_2^* is selected by maximizing σ_B^2:

$$\sigma_B^2(k_1^*, k_2^*) = \max_{1 \le k_1 < k_2 < L} \sigma_B^2(k_1, k_2).$$

It should be noticed that the selected thresholds generally become less credible as the number of classes to be separated increases. This is because the criterion measure (σ_B^2), defined in one-dimensional (gray-level) scale, may gradually lose its meaning as the number of classes increases. The expression of σ_B^2 and the maximization procedure also become more and more complicated. However, they are very simple for $M = 2$ and 3, which cover almost all practical applications, so that a special method to reduce the search processes is hardly needed. It should be remarked that the parameters required in the present method for M-thresholding are $M - 1$ discrete thresholds themselves, while the parametric method, where the gray-level histogram is approximated by the sum of Gaussian distributions, requires $3M - 1$ continuous parameters.

C. Experimental Results

Several examples of experimental results are shown in Figs. 1–3. Throughout these figures, (a) (as also (e)) is an original gray-level picture; (b) (and (f)) is the result of thresholding; (c) (and (g)) is a set of the gray-level histogram (marked at the selected threshold) and the criterion measure $\eta(k)$ related thereto; and (d) (and (h)) is the result obtained by the analysis. The original gray-level pictures are all 64 × 64 in size, and the numbers of gray levels are 16 in Fig. 1, 64 in Fig. 2, 32 in Fig. 3(a), and 256 in Fig. 3(e). (They all had equal outputs in 16 gray levels by superposition of symbols by reason of representation, so that they may be slightly lacking in precise detail in the gray levels.)

Fig. 1 shows the results of the application to an identical character "A" typewritten in different ways, one with a new ribbon (a)

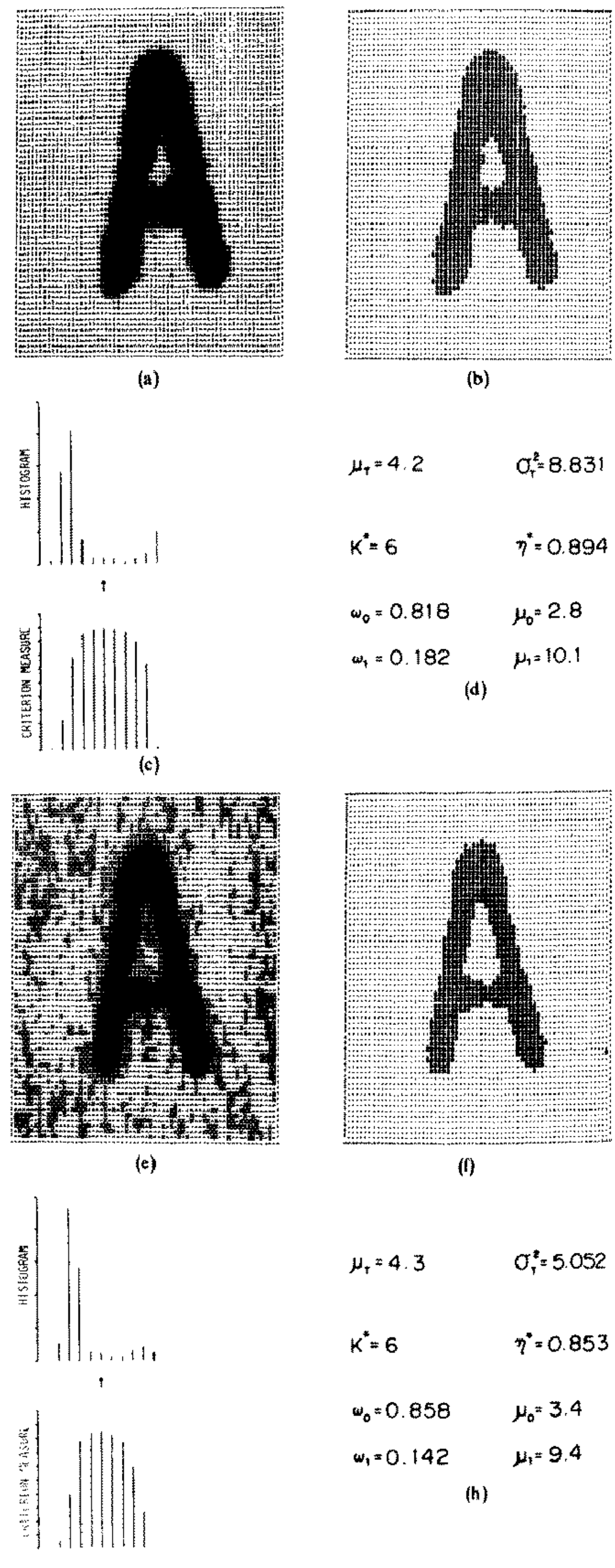


Fig. 1. Application to characters.

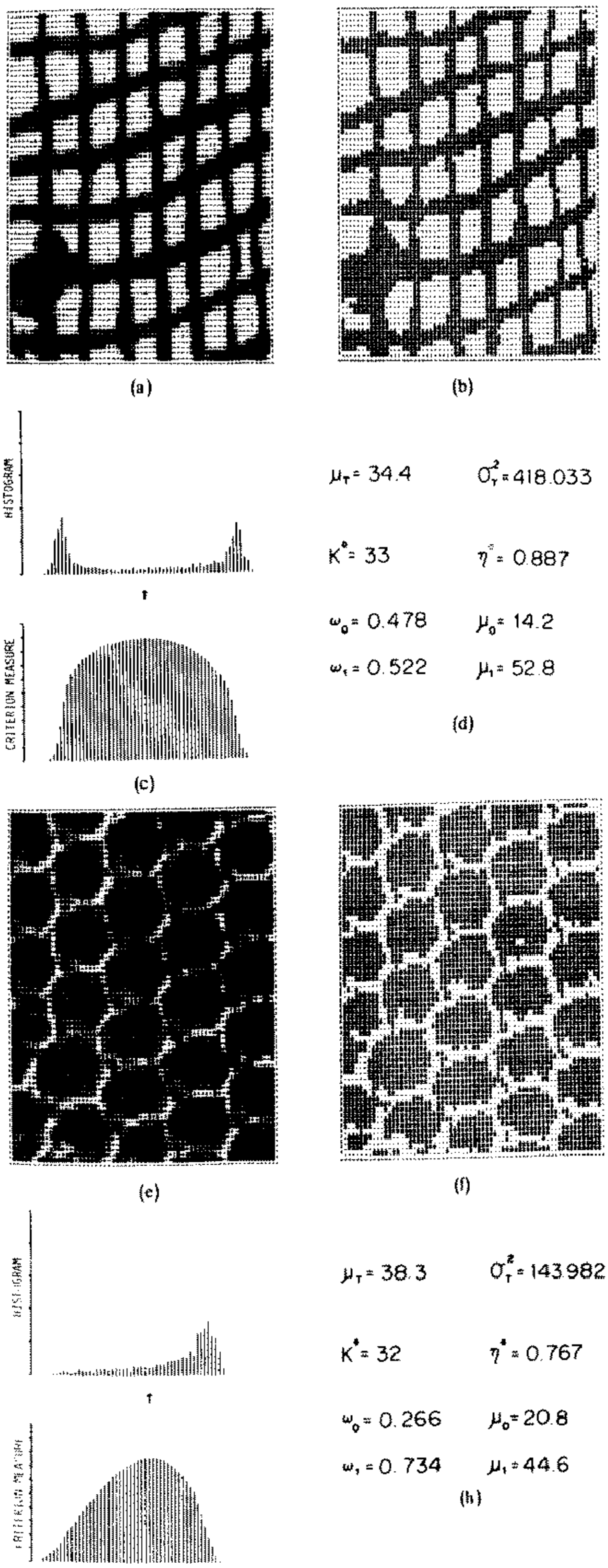


Fig. 2. Application to textures.

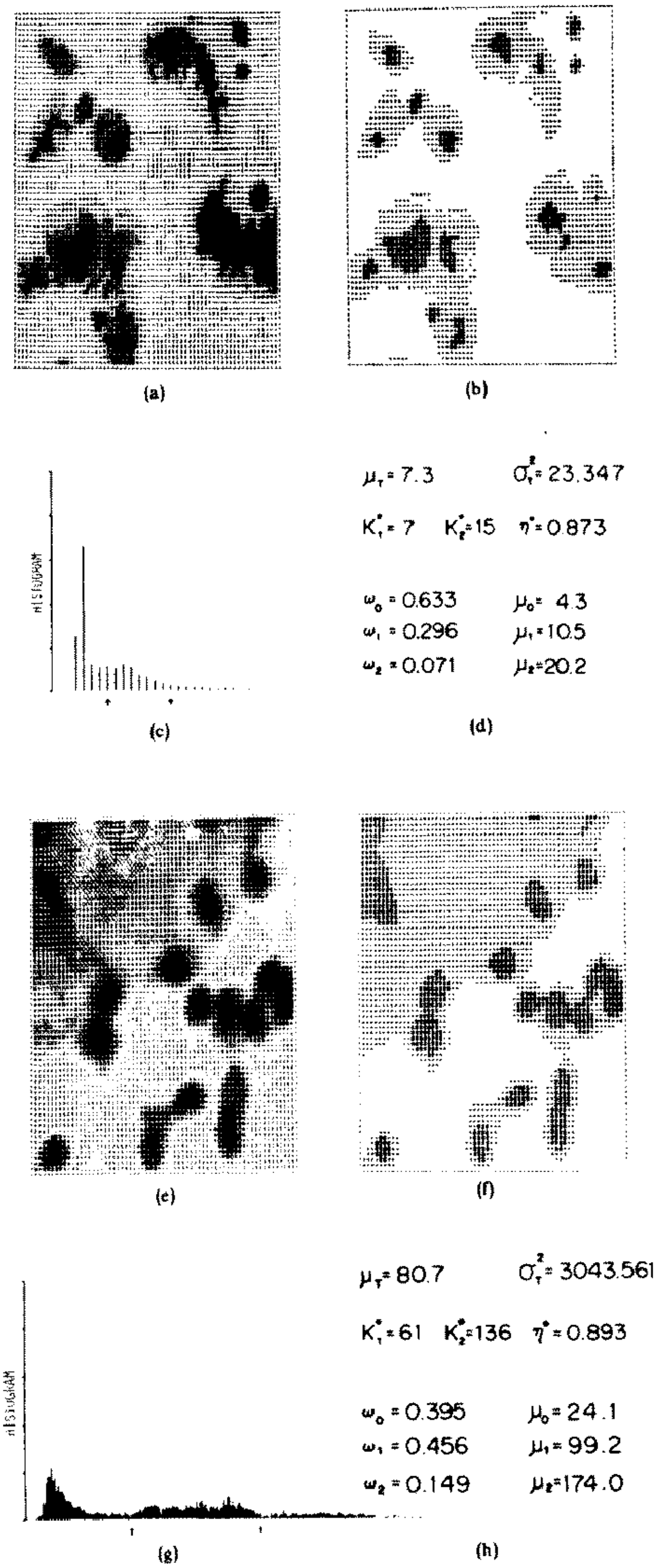


Fig. 3. Application to cells. Criterion measures $\eta(k_1, k_2)$ are omitted in (c) and (g) by reason of illustration.

and another with an old one (e), respectively. In Fig. 2, the results are shown for textures, where the histograms typically show the difficult cases of a broad and flat valley (c) and a unimodal peak (g). In order to appropriately illustrate the case of three-thresholding, the method has also been applied to cell images with successful results, shown in Fig. 3, where C_0 stands for the background, C_1 for the cytoplasm, and C_2 for the nucleus. They are indicated in (b) and (f) by ( ), (=), and (*), respectively.

A number of experimental results so far obtained for various examples indicate that the present method derived theoretically is of satisfactory practical use.

D. Unimodality of the object function

The object function $\sigma_B^2(k)$, or equivalently, the criterion measure $\eta(k)$, is always smooth and unimodal, as can be seen in the experimental results in Figs. 1–2. It may attest to an advantage of the suggested criterion and may also imply the stability of the method. The rigorous proof of the unimodality has not yet been obtained. However, it can be dispensed with from our standpoint concerning only the maximum.

IV. Conclusion

A method to select a threshold automatically from a gray level histogram has been derived from the viewpoint of discriminant analysis. This directly deals with the problem of evaluating the goodness of thresholds. An optimal threshold (or set of thresholds) is selected by the discriminant criterion; namely, by maximizing the discriminant measure η (or the measure of separability of the resultant classes in gray levels).

The proposed method is characterized by its nonparametric and unsupervised nature of threshold selection and has the following desirable advantages.

1) The procedure is very simple; only the zeroth and the first order cumulative moments of the gray-level histogram are utilized.

2) A straightforward extension to multithresholding problems is feasible by virtue of the criterion on which the method is based.

3) An optimal threshold (or set of thresholds) is selected automatically and stably, not based on the differentiation (i.e., a local property such as valley), but on the integration (i.e., a global property) of the histogram.

4) Further important aspects can also be analyzed (e.g., estimation of class mean levels, evaluation of class separability, etc.).

5) The method is quite general; it covers a wide scope of unsupervised decision procedure.

The range of its applications is not restricted only to the thresholding of the gray-level picture, such as specifically described in the foregoing, but it may also cover other cases of unsupervised classification in which a histogram of some characteristic (or feature) discriminative for classifying the objects is available.

Taking into account these points, the method suggested in this correspondence may be recommended as the most simple and standard one for automatic threshold selection that can be applied to various practical problems.

Acknowledgment

The author wishes to thank Dr. H. Nishino, Head of the Information Science Division, for his hospitality and encouragement. Thanks are also due to Dr. S. Mori, Chief of the Picture Processing Section, for the data of characters and textures and valuable discussions, and to Dr. Y. Noguchi for cell data. The author is also very grateful to Professor S. Amari of the University of Tokyo for his cordial and helpful suggestions for revising the presentation of the manuscript.

References

[1] J. M. S. Prewitt and M. L. Mendelsohn, "The analysis of cell images," *Ann. N. Y. Acad. Sci.*, vol. 128, pp. 1035–1053, 1966.
[2] J. S. Weszka, R. N. Nagel, and A. Rosenfeld, "A threshold selection technique," *IEEE Trans. Comput.*, vol. C-23, pp. 1322–1326, 1974.
[3] S. Watanabe and CYBEST Group, "An automated apparatus for cancer prescreening: CYBEST," *Comp. Graph. Image Process.*, vol. 3, pp. 350–358, 1974.
[4] C. K. Chow and T. Kaneko, "Automatic boundary detection of the left ventricle from cineangiograms," Comput. Biomed. Res., vol. 5, pp. 388–410, 1972.
[5] K. Fukunage, *Introduction to Statistical Pattern Recognition*. New York: Academic, 1972, pp. 260–267.

Book Reviews

Orthogonal Transforms for Digital Signal Processing—N. Ahmed and K. R. Rao (New York: Springer-Verlag, 1975, 263 pp.). *Reviewed by Lokenath Debnath, Departments of Mathematics and Physics, East Carolina University, Greenville, NC 27834.*

With the advent of high-speed digital computers and the rapid advances in digital technology, orthogonal transforms have received considerable attention in recent years, especially in the area of digital signal processing. This book presents the theory and applications of discrete orthogonal transforms. With some elementary knowledge of Fourier series transforms, differential equations, and matrix algebra as prerequisites, this book is written as a graduate level text for electrical and computer engineering students.

The first two chapters are essentially tutorial and cover signal representation using orthogonal functions, Fourier methods of representating signals, relation between the Fourier series and the Fourier transform, and some aspects of cross correlation, autocorrelation, and convolution. These chapters provide a systematic transition from the Fourier representation of analog signals to that of digital signals.

The third chapter is concerned with the Fourier representation of discrete and digital signals through the discrete Fourier transform (DFT). Some important properties of the DFT including the convolution and correlation theorems are discussed in some detail. The concept of amplitude, power, and phase spectra is introduced. It is shown that the DFT is directly related to the Fourier transform/series representation of data sequences $\{X(m)\}$. The two-dimensional DFT and its possible extension to higher dimensions are investigated, and the chapter closes with some discussion on time-varying power and phase spectra.