

(1977. 11. 22)

認識問題としての二値化と各種方法の検討

森 俊二 大津展之
(電子技術総合研究所)

はじめに 対象とするパターンの持つ情報がその形にあり、濃度の微妙な変化にはないような比較的単純な場合には、画面全体をまず二値化することが、それに続く処理の簡略化ばかりでなく、対象を切り出す又は強調するという意味においても、非常に重要である。実際、文字読取装置では、二値化が必ずといってよいほど行なわれている。しかし、この問題は最っとも良く利用している文字認識技術の分野では、Bartzの論文以外ほとんど正面から取り上げて議論されていない。一方画像の方では、この問題は比較的良く取り上げられて、今も各種の方法が発表されてきている。多分この様な状況は、文字の場合には、濃度がはっきりしている対象を選べる事と、単一的に認識アルゴリズムに研究の興味があるということ、それに実用装置を作る立場からすれば二値化の前処理はできるだけ単純にすませる必要のあった事が原因になっていると思われる。勿論上述の事は一般的傾向を示すもので、郵便番号読取機などは例外である。一方画像の方は自然に与えられるものが多く、しかも、まず認識のサーステップとして対象を切り出すということが重要課題となるので、二値化問題は技術的にも学問的にも取り上げられてきたと思われる。しかし文字認識技術の分野でも、認識アルゴリズムがかなりの水準に達してきて、それと共に必然的にかなり自由な対象を処理する事が要求されるようになった。また使用用紙にしろ必ずしもOCR用紙でなく上質紙の使用がユーザーから要求されるようになってきた。この様な背景から、一番最初に問題になった二値化の問題に回帰する必要に迫られたのである。しかし再びこの問題に帰って見ると、それは認識の本質的問題と深いかわり合いを持ち、困難な問題である事があらためて認識させられる。

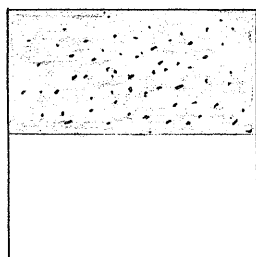
以下、文字、画像データ(線図形)を対象として、二値化問題を認識問題として見直し、今まで開発された手法を含めて、二値化手法の総括的な検討を試みる。

§1 認識問題としての二値化

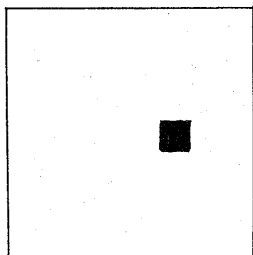
対象画像を二値化する問題は基本的には認識の問題になる。具体的には画像全体から対象を黒地として認識し、これを背景から分離する事、各点の濃度により画像全体を黒地(1)、と背景又は白地(0)の二つのクラスに分離する事である。通常各点の濃度をある一定の閾値で切って1か0に分ける事が行なわれているが、必ずしも各点の濃度のみによって、1、0に分ける必要はない。要は各クラスを分離する特徴が適当に選らばれば良い。この特徴の選択は本質的には主体者としての評価の立場から選らばれるものである。この二つのクラスを分離する特徴として最っとも明らかなものは、その濃度である。最高濃度の点は黒なのであり、最低は白であることは自明である。したがって濃度分布を取り、その谷を見出してクラス分けすることがよく行なわれている。しかし同一の濃度分布を持つ異なる閾値を必要とする二つの画像を容易に構成できる。

この例がサノ四に示されている。単純のため、3レベル画像を考えているがA図の閾値とB図の閾値は明らかに異なるのである。この場合には、黒地の特徴として「平坦さ」というものがあるということを経験として知っている。A図の場合、白地の

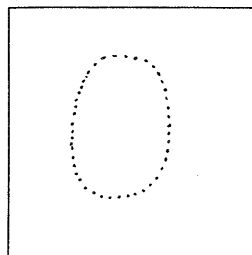
特徴としても'平坦さ'がある。この場合は、両方がその特徴として持つ'平坦さ'を充分持つように閾値は取られる。



(A)



(B)



(C)

オ1図 同一の濃度ヒストグラムを持つ閾値の異なる三つの画像。しかしBとCの閾値は等しい。

しかし、Cの点とAのそれと異なるのは、Cのそれは規則的であり、Aのそれはランダムであるということである。この場合には、黒地の特徴は単なる local feature だけでは検出されないということであり、この例は閾値選択の問題がいかにか高次の認識と関係があるかを良く示している。この様に、対象を二値化して、主体者が求めるものを切り出すには、主体者の切り出されるものに対する評価が必要である。実際もしその様な評価法が存在するとしたならば、その評価が最大になる様に、閾値を決定する事により、最良の二値化法が得られることになる。したがって正しい評価法を得る事と正しい二値化法を得ることは同一の問題となる。しかしこの評価の問題は本質的に重要であるが、一般的に取り扱うのは極めて困難な問題である。一般的に取り扱うとすれば、必然的に単純化を要求される。具体的に扱うとすれば、問題の限定が必要となる。ここでは前者の例としては判別分析の立場による評価法を、また後者の例としては、印刷文字の面積の分散評価、それに線図形に対するものとして連結性を評価する手法を具体的に述べる。しかし印刷文字以外については、それらの評価法を用いたの各方法の比較は行なっていない。それは自分の評価法を使用して比較評価すれば必然的に自分が一番良くなって客観的な評価とはならなくなってしまうというジレンマがあるからである。それ故、それぞれの評価法はそれぞれの評価による閾値の選択法ということに留めている。

§2 特徴抽出とヒストグラム

先にも述べた如く閾値選択の最っとも自然な手法は濃度ヒストグラム法である。これは一次元の特徴軸として濃度を用いたものである。より一般的には、特徴軸を濃度ばかりでなく、他の特徴軸、例えば河田等によって試みられた微分値を取ることが考えられ、多次元特徴空間上での2クラスの分離問題として考察することが出来る。従来のヒストグラムは、この特徴空間の濃度軸への周辺分布(射影)と考えられる。この濃度軸の分布が理想的にクラスター化されていれば、閾値は容易に求めることが出来る。しかしいかなる否が明瞭に出るとは限らず、この軸へのある他の特徴の射影分布を取ることににより、この軸上のクラスター化を顕著にすることが考えられる。ここで具体的に考えるのはこの問題である。

さてここに注意しなければならないのは、これらの特徴は必ずしも2つのクラスを代表するものである必要はない。2つのクラスの間を抽出するものであ

て良い。この様に濃度軸へ射影する特徴は通常の特徴抽出の如く、必ずしも片方のクラスになく、片方のクラスにあるといったものでなくとも良い。注目する特徴が各クラスの間にはない、又はこの逆の条件が必要である。我々は二つのクラス分けの問題を考えているのであるが、しかし具体的には、二つのクラスとその中間クラスという三つのクラスを頭におかなければならない。具体的には、ここでは例えば次のような *local feature* を取る。即ち 3×3 の真中の濃度を P とした時 8 方向の濃度 $N_i (i=1 \sim 8)$ との関係による特徴である。この特徴は大きく分けて次の如くなる。

極大；極小	$P > N_i ; P < N_i \quad i=1 \sim 8$
傾斜	$N_i > P > N_j, N_i = P > N_j, N_i > P = N_j \quad i \neq j$
平坦	$P = N_i \quad i=1 \sim 8$

ここで傾斜に注目したのが、坂井；長尾；渡辺；らによるものである。また Weska-Rosenfeld の方法は Laplasian の大なる点での濃度分布を求めるものであり、基本的には 3×3 の *local feature* を用いて、濃度分布を改善する一つの方法である。

§3 閾値の決定とヒストグラム

上述の如く、濃度軸上での各種特徴にヒストグラムが得られた時に、最後に閾値を求める方法が問題となる。いわば認識機構をいえば決定論理の所に相当する。もしここで正しい評価関数が上述の特徴の関数として、ということは未知数である閾値の関数として分っているならば、その最大化をばかれば良い。しかしこの様な一般的スキームによる方法とは別に、通常より単純な方法が使用されてきた。その中でも最もとも簡単なのは分布の range から median を求める方法、それから濃度分布の平均値を求める方法、特にある固定の白地 level 以上の点について平均値を求め、これを利用する方法が Bartz によって提案されている。これらの方法は比較的簡便で、文字読取装置について使用されている。一方、濃度分布での又は改善された濃度分布での、二つのクラスターの間接点としての谷を求める方法が、画像処理などで広く使用されてきた。一方特徴として差分をとり、濃度軸での差分値の総体的大きさを取り、丁度クラスターの間接点で最大値を与える事を期待しての方法が渡辺の方法であるが、要するに閾値点下ヒストグラムが最大になるようにヒストグラムを構成するのである。

先に述べた一般的スキームとしては、濃度分布を二つの正規分布の和と仮定する Kaneko による統計的決定理論による方法があるが、考え方は単純でもかなり計算が大変である。一方、これに対して大津による判別分析を用いる方法があるが、これは計算がかなり単純であり、分布の形によらず、ほとんど一意的に閾値が決定される。

これらの決定法の比較で注意しなければならないのは、それらの方法の手續さ上の簡単さもさることながら、それらの方法の一般性である。たとえば濃度レベルの全体的変動に対し、また黒地の面積変動、濃度勾配の変動、雑音の変動等によらず閾値が求められることが望ましい。この意味では分布の最大点、または一般には評価関数の最大値を求める方法は、データに対して特にあるパラメータを調整する必要がなく望ましいものである。しかしこの場合、最大点が一意的である事が必要である。これに対し、たとえば Bartz の如く $T = \alpha (\text{平均濃度}) + \beta$ として、 α, β を調整する方法は、データオリエントであり、特定のデータに対しては有効であるが、一般性という意味では劣る。しかし使用者が問題とするデータ

に対して、非常に有効に α, β が決定できれば、その範囲では十分意味のあるものである。

§4 各種の方法

上述の如く、われわれは、いわば特徴抽出に相当する濃度を軸としたヒストグラムと、それに基づいて閾値を決定する論理とを分けて考える。したがってヒストグラムとして濃度分布を取るならば、その *median* を使って閾値を決める方法、適当な平均値を求めてその一次式として閾値を決める方法、そして分散をも考慮した判別分析による方法と統計的な手法の簡略なものからオーソドックスな手法まである。一方、分布の谷を取るという上述の統計的手法とは直接関係のない直観的方法がある。

我々が行なった各種の方法について、上述の立場から作った手法マトリックスは次の如くなる。

(*median*法, 平均一次式法, 平均高次式法, 平均化法, 判別分析法, 谷検出法, 変曲点検出法, 最大点法)(濃度分布, 平滑濃度分布, ラフン シェン濃度分布, 差分分布, 傾斜分布, 極平坦分布, 連結性評価分布)^T

以下、これらの手法を簡単に説明する。まずヒストグラムより始める。濃度分布は今までに説明した如くである。平滑濃度分布は河田らの論文によるもので、その論文で良くと述べられている濃度一次微分平面での閾値線 $L2$ を採用したものであるが、実効的には平滑化して濃度分布を取るに等しいので、こう名づけた。ラフン シェン濃度分布法というのは、*Weska, Rosenfeld* によるものであり、すでに説明した。差分分布は渡辺によるもので、すでに説明した如くである。傾斜分布法は、上述の渡辺の方法のバリエーションであり、上述の計算の際、極大および極小点を除いて総和を求めるものであり、実効的に濃度勾配の傾斜部に注目するものである。極平坦分布は極大(小)点と平坦点の分布である。正確には、

$$N_i \leq P \quad \text{or} \quad N_i \geq P \quad i = 1 \sim 8$$

を満足する点の濃度を軸とした分布である。

連結評価分布は 3×3 のウィンドウで閾値で切ったとした場合の 1, 0 の連結性の度数と、非連結性の度数の比を求めるものである。ここで連結性とは 8 方向のいずれかに 1 が存在すること、非連結性とは自分自身は 1 であるがまわりに 8 の白がある場合、4 のウエイトを掛けて数える。また自分自身が白で、周囲に 1 が 8 ケあれば同様に 4 のウエイトを掛けて数える。

閾値決定法については、すでに *median* 法, 平均一次式法 ($\alpha\bar{v} + \beta$) は説明した。平均高次式とは具体的には、

$$T = \alpha(\bar{v} - T_w)^4 + T_w \quad ; \quad T_w \text{ は白地の閾値} \\ \text{(濃度レベル) } \leq T_w \text{ は白地である}$$

Bart の論文ではこの T_w は固定しているが、極平坦分布法では、これを以下に述べる変曲点検出法で自動的に決定している。平均化法というのは、 T_w により白地部分に分けられるので、それぞれのクラスターの中心、すなわち平均値 μ_1, μ_2 , を求め、 $T = \frac{\mu_1 + \mu_2}{2}$ により閾値を決定するものである。

それ故この方法は暗に、白地の閾値を決定することが含まれている。具体的問題ではたとえば白地を前もって走査しておくといった方法で知ることが可能である。しかし、我々の使用したデータではこの様な事はしていないし、また後述べる如く文字濃度より白地のレベルが変化しているので、白地の走査という方法で白地の閾値は決定できない。

判別分析法は前述した如く大津によるものでクラスターの級間分散と級内分散の比を、最大ならしめるように閾値を設定する方法で、分布の全体的形により安定に閾値を決定するものである。

谷検出法は分布の谷を閾値とするものである。変曲点検出法とは分布において、凸から凹への、また凹から凸への変化点を閾値とするものであり、これは極平坦分布法で使用された。最大点法とは、単純に分布の最大点を閾値とするものである。

§5 実験データ

実験データは線図形的なものを対象とした。それは大きく分けて文字データと画像データである。文字データは印刷文字と手書文字に分けられる。

文字データ：印刷文字

使用回数 0, 10万, 20万, 30万, 40万, 50万回のリボンを使用した時の ISO B フォントの英数字印刷文字。

観測装置はビデコンスキャナーで実効レベル数は 11, 64×64 。

白地は 3 レベルの変動あり。

データ数, $37 \times 6 \times 10 = 2220$ ケ

：手書き文字

古いボールペンによる手書き数字。実効レベル数、7 レベル, 64×64 。

データ数, 200 ケ

画像データ：Phil Bradatz の Texture の写真集

D3 ; Reptil skin

D6 ; Woven aluminum wire

D22 ; Reptil skin

64 level, 64×64 を切出して使用。ドラムスキャナー。

§6 実験結果

まず注意しなければならないのは閾値の実際のとり方である。level $> T$ と採るか level $\geq T$ と採るかの二つの取り方がある。この場合は習慣に従って ' $>$ ' を採用した。したがって $T - \alpha$ という調整を導入するとき、 $\alpha = 1$ は ' $\geq$ ' を採用した場合と考えることができる。

§6.1 印刷文字：

面積の分散平均(カテゴリー毎の平均の平均)を表 1 に表わす。この評価によれば、平均一次式法が一番良い。これは白地の閾値をぎりぎり高めに一定に取ったためである。これは未定とした他のものでは平均高次式法が一番良い。これは線巾を一定にしようとした非線形効果が利いている。なお、この面積を一定にしようと閾値を制御した他の方法では、この値は 74 となり当然ながら一番良い。特に問題となる M のような複雑な文字の面積を恒って制御しているので、たとえば、'M' 文字の面積分散は 27 で、他のものと比べて 1 けた以上も良く、この基準の直接

利用の効果が利いている。なお()つきのものは閾値が決定不能 (UDのケ数を示す) であつたり、2ビット以上の切れ、10ビット以上のしみなどが、50万回以上の文字に若干見られるもので、実際は分散も大きい、定性的評価もこれ以外のものに比べて悪い。この点、ひん弱な濃度分布より閾値を決定する判別分析法はこの場合効果的に利いている。差分分布法は印刷文字では多くのデータで閾値が低くできるので、ここに載せることができなかった。

<u>median</u>	<u>平均一次式</u>	<u>平均一次式</u>	<u>平均高次式</u>	<u>判別分析</u>	<u>谷検出</u>	<u>谷検出</u>	<u>最大点</u>	<u>変曲点検出</u>
濃度分布	濃度分布	極平坦分布	極平坦分布	濃度分布	濃度分布	平滑濃度分布	傾斜分布	極平坦分布
290	120	236	153	296	UD2 (346)	UD4 (369)	(387)	240

表 1
各方法の面積分散平均による評価表

3.6.2 ホールヘンによる手書き文字

このデータは非常に厳しいもので、いずれの閾値を採っても、しみや切れは避け難い悪いデータが含まれている。しかし実用的な閾値法としてはそれでも、なお、できるだけ人間の選ぶものに近いものを切り出すことが要求される。この評価は人間によつた。人間が選んだ閾値範囲からはずれた閾値を与えるデータ数が表2に載っている。なおパラメータの調節を必要とするものは、印刷文字で学習したものをそのまま使用している。

この場合、方法の割には良い結果を与えていた median 法は悪くなる。

median のもの ($T = median - 1 \Rightarrow T = median$) を使つても改善されない。印刷文字で悪かった差分分布法は今度はかなり良くなる。しかし傾斜分布法はあまり良くない。差分分布法を批判して、その改良として提案されたラフウシアン法は、この種のデータではかえつて比較的が悪くなる。(平滑)濃度分布法で谷を見る方法は良くないが、判別分析を適用するとほぼ $\frac{1}{2}$ に減少する。またこれをラフウシアン濃度分布に適用するとほぼ $\frac{1}{2}$ に減少する。この様に判別分析法が谷の不明瞭なものに対し効果あることと、分布がよくなると同じ判別分析法でもやはり分布の良いものが良い事が示されている。

全体的に見て極平坦分布による平均的な閾値決定法と連結評価法が非常に良い結果を与えている。しかし連結評価法はこのデータを学習データとしている。

<u>median</u>	<u>平均一次式</u>	<u>平均高次式</u>	<u>平均化</u>	<u>判別分析</u>	<u>判別分析</u>
濃度分布	極平坦分布	極平坦分布	濃度分布	濃度分布	ラフウシアン濃度分布
55	0	0	55	67	32

<u>谷検出</u>	<u>谷検出</u>	<u>谷検出</u>	<u>最大点</u>	<u>最大点</u>	<u>変曲点検出</u>	<u>最大点</u>
濃度分布	平滑濃度分布	ラフウシアン濃度分布	差分分布	傾斜分布	極平坦分布	連結評価法
108	91	58	10	45	10	0

表 2
各方法の人間による評価表

3.6.3 画像

そのむづかしさから3種類のものを使った。D22はやさしく、D3は中くらい、D6はかなりむづかしいデータである。D3の写真とその濃度分布が図a, bに示されている。各種の方法が選んだ閾値と、人間の選んだ閾値範囲が表3に示されている。斜線はこの方法では決め難いということを示し、×印は二値化結果がかなり悪いという事を示している。図2に示す如く、64 levelでは、分布がたつき、谷や変曲点など決め難い。それ故レベルを16に縮約した場合を試みた。それが64 level換算で同じ表3に示されている。(平滑)濃度分布で谷を用いるものは良い結果を与えない。差分法は文字の場合と同様、復の悪いデータに対するよりも復の悪いものに対する方が良い。また判別分析は悪い濃度分布に対しても有効に働くことが示されているが、D6に対しては悪い。これは一つには、もともとの画像は4つのレベルを持つので、これが原因であるのかも知れない。比較的良い結果を与えているのは傾斜分布法と連結評価法である。しかし、連結評価法は、端点も少し悪いという評価になっている。その他は印刷文字でセットしたパラメータを使用している。最後に図2cに白黒を反転させたD3の二値化したものが示されている。これは人間と傾斜分布、連結評価法が選んだ閾値を使っている

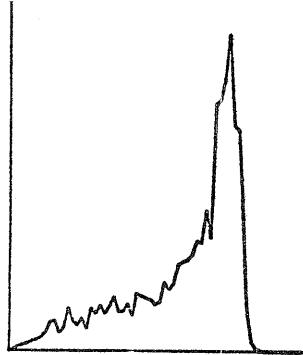


図2(b) 濃度分布



図2(a) D3の写真

	人間	median	谷検出	谷検出	谷検出	判別分析
		濃度分布	濃度分布	平滑濃度分布	ラランシェ濃度分布	濃度分布
64	D3	32~35	26			29
	D6	12~14	22			29
	D22	28~32	24		24	25
16	D3		20		28	28
	D6		16	1	16	28
	D22		20	16	20	24

	判別分析	最大点	最大点	変曲点検出	平均化	最大点
	ラランシェ濃度分布	差分分布	傾斜分布	極平坦分布	極平坦分布	連結評価
64	D3	28	22	28		32
	D6	28	11	10	13	12
	D22	25	22	24		20
16	D3	24	22	32	20	28
	D6	24	12	12	13	12
	D22	20	22	20	16	16

表3 画像に対する種々の手法の比較

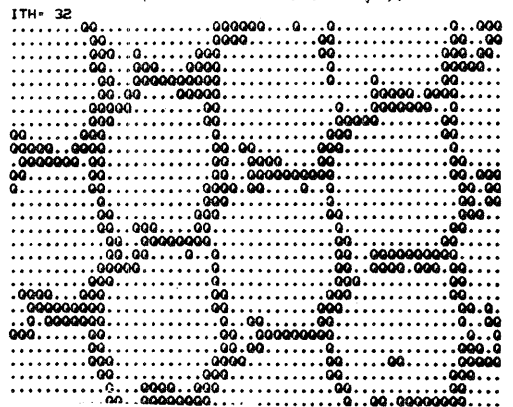
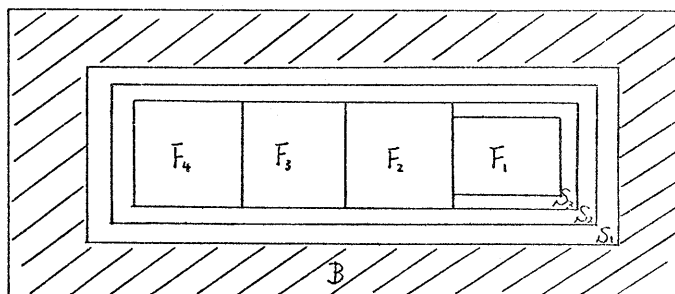


図2(c) D3の二値化パターン

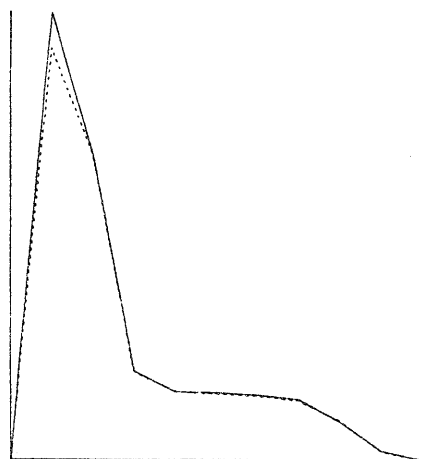
問題が単純であれば、どんな方法でも良い結果を与える。そこで我々が興味があるのは、困難な問題に対してどうなのかということである。しかもこの困難な問題というのは、実際によく起こり得る場合でその解決が実用的意味があるのかもしれない。この様な意味での困難な問題として、上述のデータから、そして種々方法による *Pilot experiment* から得る図の様なモデルを作った。直観的にこのモデルは、"線のむらによる困難"な問題のモデル化になっている。実際、体的データである図4(a)を見れば、このモデルがこの様な具体的問題を映していることがわがらう。またこのモデルによる濃度分布も確率パラメータを調整することにより現実データのものに似たものが得られる。これが図4(b)(c)示されている。

さて我々の目的は、ある方法が与えられた時、それがどのようなモデル集合で悪い結果を与えるかを知る事、それによってその方法の性質を明らかにすることである。しかし、方法といっても組合せになっていて、全ての組合せ方法について調べる事は、大変な事で、ここではその典型的な例を拾って述べる事にする。まず濃度分布から取り上げる。



B: Background, 例としては $0, 1, 2, 3$ を P_0, P_1, P_2, P_3 でランダム化する。
 S: Slope, 例としては S_1 で $3, 4$ を P_3, P_4 でランダム化する。---
 F: Flat, 一般には F_1, F_n , n のブロックに分割される。

これらのブロック内では、レバレッジは一定の値を取る。



※4図 (C) 観測パターン(a)実線とモデル(虚線)の濃度分布

21112222121112121211111221110111181121111121212111212222222
21212112121222211112113111111111222121122223121
2323133221222112112122332221111112222223441122122222334
343233331221112121223355544211111212463444762221222222334
45554443222121221224578887764211112223576556577851221021223246
667765533212221222357888776432111112467777777851221021212357
7777764321122122378897778532121235887788777521121212357
8667865331221212234788977776574321212478877777752110221121347
6557764322112121223478867655743222247887667775221022121347
6667764321211212234678855554322212136655456664221022121347
6677876432121121223577655456642212377865456663211212212136
666776432121121223577655456642212377865456663211212212136
6668764321112222353765545664225787755676421111212134
667764322111112123577664365554326776445677642111112136
5667753321111135887654456654322557643436664321211112135
5455654321111123478876443666443266653265554521111121234
44654533111111235776444425543245543223555431111121224
4465453311111235776444425543245543223555431111121224
44664432121110122367764321467654321467654321211112134
445775431211112213776421234545457874422256754312211112134
445774332111111236765321334555576422247766632111112223
4466433211111124687432113345555774321247556421211122235
33333331111112368753212366666542123665532112112235
34666632111111247875321225555532122455532112112235
346666321211111247875321225555532122455532112112235
445777532111111257886432122567655431223666543211122234
455777532111111258997632122367885431225665432112121235
4457864321111113689986421134578864311225675432112121235
445777532111111368999642113368753211123576443211111135
44477532111111136899964211336875321112357644321111135
3346664321121011113688886432123354221223664321121213
3346664321121011113688886432123354221223664321121213

[illegible]

才4回 (a) 文字'M'の観測値

(b) モデル

§§7.1 濃度分布

§§7.1 median 法: この場合は考察が比較的簡単である。平坦部の最大レベルを $\max F$, 白地の最小レベルを $\min B$, 傾斜部の最大を $\max S$ とすれば, $(\max F + \min B) / 2 > \max S$, $< \max B$ の場合にそれぞれ, 欠け, 溢りの雑音が見れる。

§§7.2 平均化法; この場合は容易にわかる如く, median 法を緩和した方法になっている。直観的に言って少数の点が高くても平均としては利いてこないからである。したがって median 法でだめな場合がこれで克服されることがわかる。しかし平均化法が全ての点で median 法に勝るとは必ずしも言えない。これについては後で述べる。この方法で悪い場合を考える。まず閾値が低すぎる事が起こるかどうを見る。そのため白地では $\max B$ と $\min B$ の二つのレベルを考え平坦部の面積がレベル 5 の中 W の線の場合を考える。 S_1 はレベル 3, S_2 はレベル 4 とし十分細長いとすれば黒地の平均値 $(14+5W)/(4+W) \equiv \bar{B}(W)$ で与えられる。このとき,

$$\frac{\alpha \min B + \beta \max B + \bar{B}(W)}{2} < \max B; \quad \alpha + \beta = 1, \quad \alpha, \beta \leq 0$$

が成立立つとき, 閾値が低すぎることになる。 $\min B = 0$, $\max = 2$, $W = 1$ とすれば $\beta < 0.1$ のとき上式は成立する。すなわち白地の 0 が 90% 以上のとき白地に 10% 以下のノイズが見れることになる。

一方高すぎる場合のモデルとしては, まずすぐ分かる如く, 黒地の平均値は傾斜部分が低める様に働く, そこでこの傾斜部分がなくなった場合どうなるかを見る。しかも平坦部分は $\max F$ と $\min F$ のみとする。また白地は $\max B$ のみとする。このとき

$$\frac{\alpha \min F + \beta \max F + \max B}{2} > \min F$$

の条件が成立するとき, 閾値が高すぎることになる。たとえば $\max F = 0$, $\min F = 5$, $\max B = 2$, $\beta = 0.9$, $\alpha = 0.1$ としたとき, $T = 5.7 \sim 6$ となり 10% の欠けが生ずる。

上述の場合はかなり極端な例で, 実際のほとんどの起り得る場合は黒地の中間レベルがあり, より低く黒地の平均値は決る。しかし文字の場合など, 片うちの様な場合, 上述の様な事が十分起り得る。ともかく, 白地閾値が既知という条件はあるが, この平均化法はかなり有効な方法である。

§§7.3 谷の検出; この場合は明らかに, 悪い場合は谷が消えることである。この谷の消滅は上述のモデルから容易に構成できる。すなわち傾斜部分の面積が平坦部分の面積より大きくなれば起り得る。

上述のモデルで背景レベルが 2 で S_1 のレベルが 3, S_2 が 4 とし F が全てレベル 5 とする。この時 F の矩形を $a \times b$ とすれば近似的に $x + y > \frac{1}{2} xy$ のとき谷は消滅する。さて F が $(\alpha_1 + \alpha_2 + \alpha_3 + \dots + \alpha_n) F = F$, $\alpha_1 > \alpha_2 > \dots > \alpha_n$, (α_n に相当する領域が最高レベルを持つとする) を満足するように連続する n 個のレベル領域に分割されると, 最大の面積は $\alpha_1 F$ である。このとき, $x + y > \frac{1}{2} \alpha_1 F$, のとき谷は消滅し, $F = x \times y$ で α_1 は 1 より小さいのでこの様な分割が起こると, より容易に谷が消えやすくなることが分る。一般には線状の黒地であればこの現象は容易に起りやすいということが分る。このモデルは現実のものによく合っている。

§§7.2 差分分布

この方法は上述の濃度分布の谷が消滅するような場合、すなわち平坦部分の面積が小さくなる場合でも edge の値を取っているので、無関係に最大値が決まる点であぐれている。しかし一つの問題点は、白地の領域が広く、しかも雑音が多い場合、白地に peak が出る可能性がある。たとえば白地レベルは 0,1 のみとし、それが等確率とする。白地の面積を W とすれば、自分より低い場合を取る D^+ の場合の D^+ の期待値は $4 \times 0.5 W$ である。一方黒地の方の面積を B とし、矩形の縦横の比を $10:1$ (印刷文字の場合この程度) とすると、周長は $2\sqrt{10}\sqrt{B}(1+0.1) \cong 2\sqrt{10}\sqrt{B}$ となる。今こゝで差分法に有利に考えて、黒地部は S 領域はなしとして、平坦部のみとし、しかもそれが一様な高さ m を持ち急峻に下に落ちていくとする。そのとき、この端 m レベル点で白地への落差は最大 m で最小 $m-1$ である。そこでこれを差分法に有利に m と置く。差分に利くのは端の片方の 3 方向のみである。総面積を A とし、白地との比を γ とすれば $W = \gamma A$, $B = (1-\gamma)A$ となる。この時、白地の peak が大なる条件は、 $\frac{\sqrt{A}}{m} \cdot \frac{\gamma}{1-\gamma} > 3\sqrt{10}$ で与えられる。 $\gamma = 0.5$ とし、 $A = 60 \times 60$ とすれば、 $m \leq 4$ のとき上式は満足する。それ故に上述の現象は実際のデータで十分起り得ることがわかる。

さて一方、白地にノイズがないとしても、上述のモデルで平坦部が $m, m-1$ の 2 レベルでそれぞれ α, β の比でランダムに存在していて、 $\alpha > \beta$ のとき m レベルが閾値になる可能性がある。すなわち D^+ 分布では最高レベルのところで Peak になる。 D^- では白地レベルのところで最高値を示す。それ故渡辺が彼の論文で述べている如く、この平均をとることが考えられ、実際この悪い場合は容易に回避できる。

次にこの方法が問題になるのは Weska-Rosenfeld の論文で指摘されている如く、傾斜が長く直線的である場合である。この場合には、最高点近傍が平坦になり最高点を定めにくい。しかしもしこの傾斜部分だけをこの分布全体から切り出すことができればこの平均値を求めることによりこの困難を回避できる。

最後に、考えられ、また実際にあったこの分布の悪い場合は、連続的ではなく、かけはなれた二つのレベルでの極大点がほぼ同一の値を持つ場合が起り、不安定になるということである。この原因は単純で上述のモデルで言えば平坦部の値は一定 m として正常な場合を想定し、 SZ の分布が level d_1 と d_2 ($d_1 > d_2$) が $r, 1-r$ の割合で発生するものとする。 S_1 は白地レベルの 0 とすると、 d_1 での D^+ と m での D^+ の値が等しくなる条件は次の如く $m < 2d$ で与えられる。要するに傾斜部にむらがあれば、この現象は起り得るのであり、実際にボールペンのデータで一それが発見された。この実際のデータは上述のモデルと良く合っている。

§7.3 ラフロシアン濃度分布

この悪い場合はまず次の如く構成される。すなわちステップ毎に高さの異なる階段を考える。各段のレベルを m_i とすると $(m_{i+1} - m_i) = (m_i - m_{i-1}) - C$ のとき、 m_i レベル点でラフロシアンが等しくなる。こゝで $C > 0$ のとき、項上がなめらか、 $C < 0$ のとき急峻となる。ともかくこの様な場合には分布はガウスがたとなる。

次に最初のモデルで平坦部が L 々に分割され、それぞれを m_i とする。それぞれが対応する SZ 領域で値 d_i を持つとする。 S_3 領域は無し。このとき、 $m_i - d_i = C$ とすれば各 m_i レベル点でラフロシアンの価は等しくなる。一方 S 領域では上述

の漸化式を適応すれば、各レベル点でもラフレーションが等しくなる様に構成できる。この様に各濃度レベル点でのラフレーションが等しいというのはさつぱい条件で、これによりラフレーションの調整を不能にしている。実際の例ではボールペンのデータで5ヶ見られたが、これはラフレーション最大値の30%以上を持つ点の分布で谷がないということで、これを50%または60%にあると谷が出る。しかし点の数が少なく不安定であり、かつデータにより谷の出る%が異なるのも問題である。また画像データでは64 levelでは、分布はかなりのがたつきが見られるが、これは上述の議論から予想されるものである。

§7.4 極平坦分布

いままで述べてきた各分布の悪い場合について、この分布は谷を与える分布を持つが、しかし一方において次の様な場合に谷が消滅する。すなわち白地に傾斜部と同じ濃度を持つノイズが多く存在する場合である。傾斜部は極大でもなければ平坦でもないのだから、この部分は分布から抜けることになりここに谷が現れるということであるが、しかし白地の濃度分散が大きく、傾斜部にある中間レベルが白地に多く存在すれば、これが極大、平坦として数えられて谷が消滅する。実際、ボールペンのデータがこの様なデータが3ヶ見られた。

一オ、傾斜分布はこの様なノイズには一切影響を受けないが、単なる濃度分布、微分分布、ラフレーション分布ではこの種のノイズには不利である。次に述べる連結評価法では、この様なノイズがランダムであれば影響を受けない。しかし単純な滑らかな形を持つば影響を受ける。

§7.5 連結評価法

この方法の問題点は閾値 '0' に *singularity* が存在することである。一般にこの方は1の数を A とすると適当な δ, δ' で

$$\rho = \frac{A - \delta}{A + \delta'} \quad \delta, \delta' > 0$$

の如く表わせる。明らかに閾値 = 0 では $A \rightarrow R \times R, \delta' \rightarrow 0, \delta \rightarrow \epsilon$ であるほとんど1になって最大の評価値を得ることになる。閾値 '0' はしかし、ある意味では *trivial case* として除外して考えられるが、'0' の近くで、上述のようなことが起り得る。たとえば、輪郭が滑らかで大きな平盤は汚みノイズが白地にあった場合、これを '1' とした方が評価値は高い。しかしながら、この汚みは形として評価され得るものであり、物理的なものでなく、形による評価によるこの方法では、ある意味では当然の結果である。

あとがき 一般にパターン認識の研究というのは、研究者それぞれの手法、アイデアは非常に良く発表されているが、それが他の手法とどの様に関係があり、それと比較してどうなのかという非常に重要な点が欠けるらしいがあった。しかも使っているデータも異なっているのが普通なので、両者が共通の土俵で議論される事が少ないのが現状である。この様な現状に対し、この論文がいくらかでも役立てば筆者の幸いとするところであり、これを契機として多くの議論の起ることを期待するものである。

謝辞 この研究は通産省工業技術院のパターン大型プロジェクトによるものであり、西野部長をはじめとして関係諸氏に深く感謝するしだいです。また東芝総研の森健一氏、渡辺貞一氏に有益な御討論をいただいたこと、また同様に図形

