

多階調画像と2値画像の相互変換

松岡 輝彦 湯田 美和子 小菅 佳克 千種 康民 伊吹 公夫

東京工科大学 工学部 情報工学科

多階調画像を2値画像に変換する方式は数多くあるが、その方式は、どれも一長一短であり、用途に応じて使い分けているのが現状である。またその逆問題である2値画像から多階調画像への変換はいまだ確立された手法がない。同時にそれらを評価する手法についても、ほとんど全ての変換手法が元の画像の質に影響されるため、定性的な評価法はもちろん、定量的な評価を行うことも難しい。しかし、それらは極めて重要性が高いものである。

そこで、著者らは、2値化法と多階調画像への復元法を各一つ提案し、文字が図中に含まれる画像について、アンケート調査を使用して、従来法を含んだ比較調査を行い、多階調画像→2値画像→多階調画像の一連の相互変換法についての定量的な評価を試みた。

本稿では、第1に、2値化法について、従来の典型的な手法と新たに提案する神経回路網を用いた手法について論述するとともに、復元法について、平均化による手法とそれを改良したエッジを強調する手法について述べる。第2に、それぞれについてシミュレーションを行い、定量的評価を行った。

Mutual Conversion between Gray Scale Images and Binary Images

Teruhiko Matsuoka, Miwako Yuda, Yoshikatsu Kosuge,
Yasutami Chigusa and Kimio Ibuki

Faculty of Engineering, Tokyo Engineering University

1404-1 Katakura, Hachioji, Tokyo 192, Japan

There are many methods to convert from binary images to multilevel gray scale images. One of the methods employed depends on its aspect case by case now. There is no established way to regenerate the gray scale images from the binary images. Moreover it is difficult to evaluate it quantitatively as well as qualitatively because every method employed depends on the quality of the original images. It is, however, very important to evaluate the methodology.

We proposed a novel binarization method and a regeneration method. By using questionnaires we compared the methods including the conventional methods about a picture with characters. Then we tried to make quantitative evaluation about the mutual conversion.

1 序論

現在ファクシミリが普及しているが、多階調画像の2値化に利用されている手法では、絵柄領域は比較的よい2値画像を得られるが、文字領域は十分な画像を得られない。また、文字領域の2値化に利用されている手法では、文字領域は十分な画像を得られるが、絵柄領域は十分な画像を得られない。このため、文字／絵柄領域の混ざった画像を2値化し良質の2値画像を得るためには、文字／絵柄領域の領域分割が必要であった。

そこで、今回著者らは、新しくホップフィールド型の神経回路網[2]に着目し、領域分割をしないで文字／絵柄領域の混在した画像から良質の2値画像を得ることを試み、アンケート調査により従来法と比較評価した。

さらに、その2値画像を復元するという逆問題にも取り組み、平均化による復元法とその手法を改良しエッジを強調する手法を試みた。これらに関しては今後の研究がさらに必要である。

2 2値化法

2.1 2値化とは

各画素が白または黒のみで表示された画像を2値画像(binary image)といい、それに対して白と黒の間に何段階かの明るさをもつ画像を多階調画像(multilevel gray scale image)という。白黒の多階調画像を白黒の2値画像に変換する処理を2値化という[1]。本研究では、256階調で表現した縦513×横512画素の画像を2値の画像に変換する。各画素が0(黒)から255(白)までの値をもつ56階調では8ビット使用するが、0(黒)か1(白)かの2値にすることにより1ビットになり、データ量が1/8に低減され、またこれを処理する場合、CPU時間や記憶容量において有利である。

本稿では、座標 (i, j) のある画素に注目するとき、原画像の濃度値を $F(i, j)$ とし、出力画像の濃度値を $f(i, j)$ とする。

2.2 従来法による2値化

2.2.1 単純2値化法

2値化の方法の中で、もっとも簡単な方法といわれるのがこの単純2値化法である。これは入力画像の各画素の濃度とあるしきい値の大小関係により、出力画像の濃度値を1か0に決定する方法である。注目画素 $F(i, j)$ に対し、しきい値を t とすると式(1)のように2値化される。

$$f(i, j) = \begin{cases} 1 & (F(i, j) \geq t) \\ 0 & (F(i, j) < t) \end{cases} \quad (1)$$

2.2.2 組織的ディザ法

単純2値化法の画像は、原画にはない白と黒のはっきりした輪郭を表示しているため自然性を損ねている。雑音を加えることにより白と黒のはっきりした輪郭を目立たなくする方法が、ディザ法(dither method)である[1]。組織的ディザ法とは周辺画素には関係なく決定する独立決定法である。組織的ディザ法は入力画像にディザマトリクスをあわせて、各画素の入力画像の値とマトリクスの値を比較して、入力画像の値の方が大きい場合は1、小さい場合は0とするものである。本研究では、 4×4 画素のマトリクスを用いて2値化を行った。マトリクスの中には0からマトリクスの画素数より1小さい値までが入っているので、入力画像の値もその範囲内の値に正規化しなければならない。マトリクスの画素数を mat 、正規化した値を $fd(i, j)$ とすると式(2)のように表される。ここで、 $fd(i, j)$ は整数値であり、 $int(x)$ は x について少数点以下切り捨ての整数化関数である。

$$fd(i, j) = int \left(\frac{F(i, j) \times mat}{256} \right) \quad (2)$$

変換する画素のマトリクスの値を $dit(i, j)$ とすると、組織的ディザ法は式(3)のようになる。

$$f(i, j) = \begin{cases} 1 & (fd(i, j) \geq dit(i, j)) \\ 0 & (fd(i, j) < dit(i, j)) \end{cases} \quad (3)$$

この組織的ディザ法は、入力画像の値とディザマトリクスの値を比較するだけなので処理速度も速く現在も広く利用されている。

2.2.3 平均誤差最小法

原画像の濃度値と出力画像の濃度値との空間的な誤差を平均的に小さくするようにしきい値を決定する方法が、平均誤差最小法である。処理はラスタスキャン方向に行い、また、注目画素のしきい値 $t_{x,y}$ は、周辺の画素の誤差の重み付き平均で決定される。これを式に表すと式(4)のようになる。

$$t_{x,y} = \frac{1}{2} + \frac{1}{\sum_{k,j} \alpha_{k,j}} \sum_{k,j} \alpha_{k,j} \times E_{i+k,j+1} \quad (4)$$

$$E_{i,j} = F(i,j) - f(i,j)$$

$$\alpha = \begin{bmatrix} 1, 3, 5, 3, 1 \\ 3, 5, 7, 5, 3 \\ 5, 7, 9, 7, 5 \end{bmatrix}$$

ここで $E_{i,j}$ は、すでに決定した出力画像の濃度値と入力画像の濃度値の誤差であり、 $\alpha_{k,j}$ は重み付けの係数で、現在注目している画素に近いほど大きな重み付けを行うものである。

2.2.4 濃度パターンによる2値化法

入力画像をある大きさに区切り、その中の濃度情報によりその範囲内の画素の値をパターンに従って決定していく方法がこの濃度パターンによる2値化である。今回、2通りのパターンを作った。

2.2.4.1 濃度パターンによる2値化法①

この方法は、 4×4 画素を1つの画素のように考え、その濃度値の総和に対応する白黒のパターンに割り当て2値化を行うものである。17通りのパターンにより、高い濃度から低い濃度まで17段階の濃度を表現することができる。17段階のパターンは、 4×4 画素の左上の点から右下の点に向かって斜めに増えていくものとする。

ここではまず、入力画像を 4×4 画素に区切りその範囲内の濃度値を0から16までの値に変換し、

17通りの白黒のパターンから濃度値に適するものを用いて2値化する。入力画素の 4×4 画素のそれぞれの値を $i_1 \sim i_{16}$ とし、対応する面積濃度値を fp とすると式(5)のように表される。

$$fp = \text{int} \left(\frac{\sum i_n}{16} \times \frac{17}{256} \right) \approx \text{int} \left(\frac{\sum i_n}{256} \right) \quad (5)$$

2.2.4.2 濃度パターンによる2値化法②

前に述べた濃度パターン①による2値化とこの濃度パターン②による2値化との違いは、前者は濃度に応じて常に決まったパターンで2値化されていたのに対し、後者は2値化後の面積濃度値は同じになるが、どの画素が白または黒になるかは入力画像の濃度値に依存するという点である。

4×4 画素の範囲の濃度値 fp は、式(5)により求められる。この16画素のうち、入力画像の濃度値が高いものから順に fp の数だけ白になり、残りは黒になる。

2.3 Binamics法による2値化

2.2.1 神経回路網と2値化

今日では、パターン認識の問題だけでなく、学習、推論、記憶、言語などの人間の認知機能に関する問題を神経回路網のシステムをベースとして様々な研究が行われている。

本研究では、近年注目されているホップフィールド型の神経回路網により、2値化を行うことにした。そして、この方法をBinamics法と名付ける。ホップフィールド型の神経回路網は、ニューロンの人工的なモデルであるユニットがそれぞれ非同期的に動作し、自分の状態(出力)を変化させていくことによって処理が進むものである。このとき各ユニットが、完全に独立して動作し、並列計算機のモデルとしては非常に魅力的な並列性を有している。また、ホップフィールド型の神経回路網の時間発展規則では、ネットワークの中からランダムに1つのユニットを選び出し、その選ばれたユニットへの入力の総和を求め、その値に従ってユニットの出力を更新し、それ以外のユニット

の出力は変化させないという処理を、繰り返すことにより実現する。

2.3.2 入力画像の濃度値の前処理

本研究では、さらに2値画像の自然性を出し、画像の質を高めるために、2値化の処理に入る前に前処理を行う。この処理を行うことにより、処理結果をよりコントラストの強いものとし、出力画像がより明確になり、細かい部分も表現できるようになった。この処理は図1の実線で示す非線形関数を使用する。図1で破線は前処理を行っていない状態である。今回、何種類かの画像に対して、実験的に定めた非線形関数を使用した。

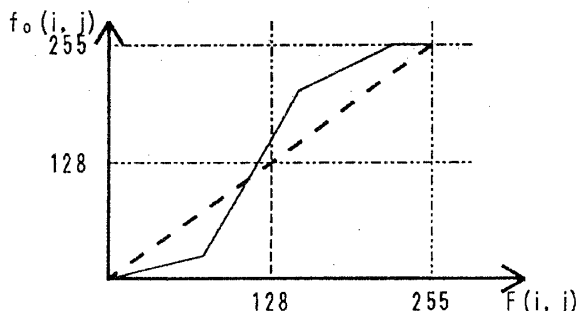


図1 前処理による入力画像の濃度値の変換

この非線形関数を式で表すと、次式(6)になる。ここで、入力画像の濃度値を $F(i, j)$ 、前処理後の濃度値を $f_o(i, j)$ とする。

$f_o(i, j) =$

$$\begin{aligned} & \left\{ \begin{array}{l} \text{int} (F(i, j) \times 0.25) \\ \text{int} (F(i, j) \times 2.5 - 180) \\ \text{int} (F(i, j) \times 2.7 - 204) \\ \text{int} (F(i, j) \times 0.7 - 96) \\ \text{int} (F(i, j) \times 0.14 + 220) \end{array} \right\} \left\{ \begin{array}{l} 80 \leq F(i, j) < 120 \\ 120 \leq F(i, j) < 150 \\ 150 \leq F(i, j) < 220 \\ 220 \leq F(i, j) \end{array} \right\} \end{aligned}$$

2.3.3 Binamics法による2値化法

今回の多階調画像の2値化ではまず、ランダムに1画素を抽出し、その画素の入力画像の濃度値といくつかの周りの画素により、選び出された画素は0か1に2値化される。この処理をランダムに513×512画素について1通り行う。さらに、一連の処理をを繰り返し行い、変化のあるところを更新する。これを何度か繰り返すことにより、しだいに変化するところが減少し、ある程度収束し

たところで処理を打ち切り、2値画像を出力する。

本研究ではランダムに選び出された画素の濃度値 $F(i, j)$ の0か1を決定する方法は、その画素の周りの12画素の総和に、256階調で黒を表す濃度値0を加えたときの平均値と、白を表す濃度値255を加えたときの平均値がどちらがより入力画像の濃度値 $f_o(i, j)$ に近いかを比べ、その差が小さくなる2値のうちの一方を採用する。つまり、近傍12画素の総和に濃度値0を加えて平均した値と入力画像の濃度値 $f_o(i, j)$ との差の絶対値を $E_w(i, j)$ 、近傍12画素の総和に濃度値255を加えて平均した値と $f_o(i, j)$ との差の絶対値を $E_b(i, j)$ とすると、式(7)のように表される。

$$F(i, j) = \begin{cases} 1 & (E_w(i, j) \leq E_b(i, j)) \\ 0 & (E_w(i, j) > E_b(i, j)) \end{cases} \quad (7)$$

ただし、

$$E_w(i, j) = \left| f_o(i, j) - \frac{\sum f(p, q) + 255}{13} \right|$$

$$E_b(i, j) = \left| f_o(i, j) - \frac{\sum f(p, q) + 0}{13} \right|$$

近傍演算の際の画素の取り方を図2に示す。人間がある画像を見ると、1画素ずつ見るのではなく、いくつかの画素の集まりを点や線または絵として見ているので、注目画素を2値化するためには、周りの画素の情報も必要である。予備実験として、近傍演算の画素数を変えたり、形状を変えて2値化を行い、12近傍を採用した。

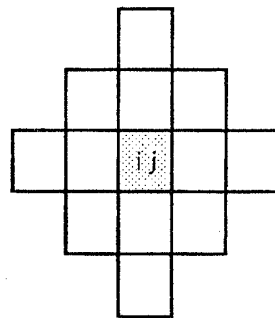


図2 画素の取り方

3 復元法

3.1 多階調画像の復元について

2値画像から多階調画像への変換はいまだ確立された手法がない。そもそも画像というものはそれを見た本人の主観によって評価される。言い方を換えれば、これといった決まった答があるわけではないのである。そこで、著者らは本研究において1つの復元法を試みた。

3.2 平均化による復元法

平均化による復元法とは、図3に示すように、変換したい画素と周りの12画素、合わせて13画素の平均値をとり、それをその画素の階調値とする方法である。この13画素は変換する入力画像の濃度値を利用し、更新値は使用しない。また、画像データの上下左右の境界部における処理は、存在する画素のみの平均をとっている。

この方法の利点は、単純な方法で復元が可能であるということである。つまり、もしこれをチップに組み込む場合に単純な回路で構成できるということである。

近傍処理の形状を図3の形にした理由は、文字というものは比較的縦や横の線が多く、斜めの線などが少ないので縦と横の情報を強調し、それにより文字を比較的綺麗に出すためである。また予備実験により12近傍を採用した。

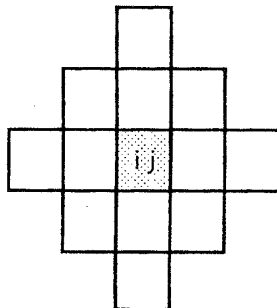


図3 画素の取り方

3.3 平均化による復元法の改善

さらに、出力画像のエッジを強調するため、図4に示す4通りの分類を行い大小2つの三角形に分け、それぞれ大きな三角形と小さな三角形の平均値を出して、2つの三角形の平均値の差を求めた後、4通りのそれぞれの差を比較し、その中で一番差の大きなものを取り出し、その差が基準値より大きな場合には、エッジを検出し、大きな三角形の平均値をその画素の階調値とするように変更した。

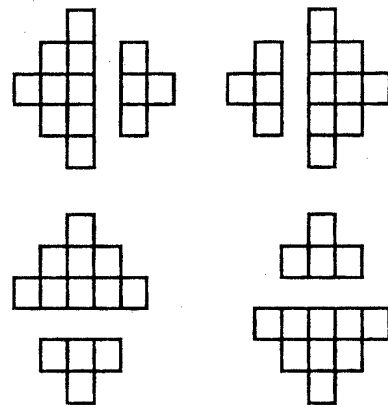


図4 点の取り方

4 2値画像の実行例

4.1 原画像

図5に原画像を示す。この後示す2値画像と復元画像の実行例と比較するものである。



図5 原画像

4.2 単純2値化法

図6に実行例を示す。コントラストははっきりしているが、画像の情報が著しく落ちているのがわかる。



図6 単純二値化法

4.3 組織的ディザ法

図7に実行例を示す。縦じまが出てしまい、あまりよい画像とはいえないようである。これは、ディザマトリクスに依存するものである。



図7 組織的ディザ法

4.4 平均誤差最小法

図8に実行例を示す。細部情報は落ちているが、全体の情報はある程度保たれた画像になっている。



図8 平均誤差最小法

4.5 濃度パターンによる2値化法①

図9に実行例を示す。白黒がはっきりしており、細部情報が多少落ちているようである。



図9 濃度パターン①

4.6 濃度パターンによる2値化②

図10に実行例を示す。濃度パターンによる2値化①よりは、情報の落ち具合は少ないが、そのかわり縦じまが出ている。これは組織的ディザ法と同様の現象である。



図10 濃度パターン②

4.7 Binamics法による2値化

図11に実行例を示す。この画像は計算を6回繰り返したものである。細かいところも比較的よく出ており、情報の落ち具合が少ないことがわかる。さらに、絵も文字も綺麗に出ている画像はこの画像であると思われる。



図11 Binamics法

5 復元画像の実行例

5.1 Binamics法による2値画の復元（13画素）

図12に実行例を示す。



図12 Binamics法による2値画の復元（13画素）

5.2 Binamics法による2値画の復元（9:4画素）

図13に実行例を示す。



図13 Binamics法による2値画の復元（9:4画素）

5.3 Binamics法による2値画の復元(5画素)

図14に実行例を示す。



図14 Binamics法による2値画の復元(5画素)

5.4 Binamics法による2値画の復元(25画素)

図15に実行例を示す。



図15 Binamics法による2値画の復元(25画素)

6 定量的評価

6.1 2値化についてのアンケート調査

図17のグラフにアンケート調査の結果を示す。この結果を見るとまず、単純2値化法では“まあまあ”の人が一番多く、次いで“あまりよくない”というようになっている。本当はかなり情報の

落ちている画像なのだが、ぱっと見た感じが白黒はっきりしているため、よい評価を得ることがあると考えられる。次にディザ法はあまり綺麗な画像を用いたため、“あまりよくない”と“よくない”が圧倒的に多くなっている。平均誤差最小法も“あまりよくない”が多くなっている。これは画像にマトリクスの癖が出力画像に現れたためであると考えられる。濃度パターンでは、①も②も“まあまあ”が一番多くなっている。これはドットの粗さが多少あるものの比較的もとの画像に近い形で濃度変換が行われている為であると考えられる。最後にBinamics法であるが、画像のアンケート調査の結果の中で“とてもよい”が圧倒的に多く、“よい”も合わせると半分以上の人が良さを認めているのがわかる。

これらのアンケート結果からBinamics法を用いた2値化法はかなり実用性が高いものであると思われる。

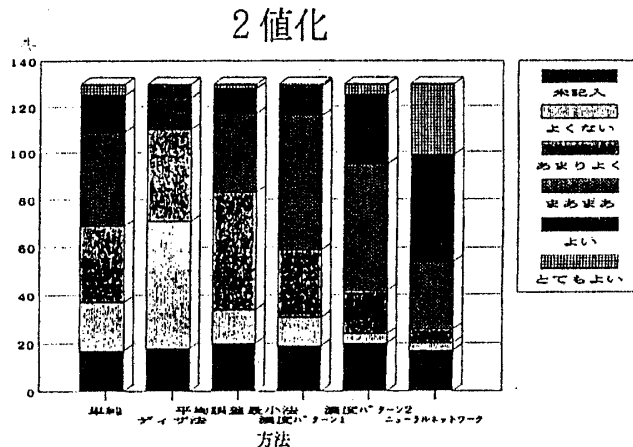


図17 2値化についてのアンケート調査

6.2 復元についてのアンケート調査

図18のグラフにアンケート調査の結果を示す。このグラフから単純2値化法とディザ法、そして濃度パターン②の方法ではあまり良い画像が得られていない。また、2値化ではあまり良い評価を得られなかった平均誤差最小法が、復元ではまあまあの評価を受けている。濃度パターン①の方法

が比較的良好な評価を得ているのは画像全体を評価したためであると考えられる。細部情報はかなり劣化しており、もっと低い評価になったのではないと思う。Binamics法の復元は2値化した画像の質がよい為に復元しやすく評価も良いものになっている。

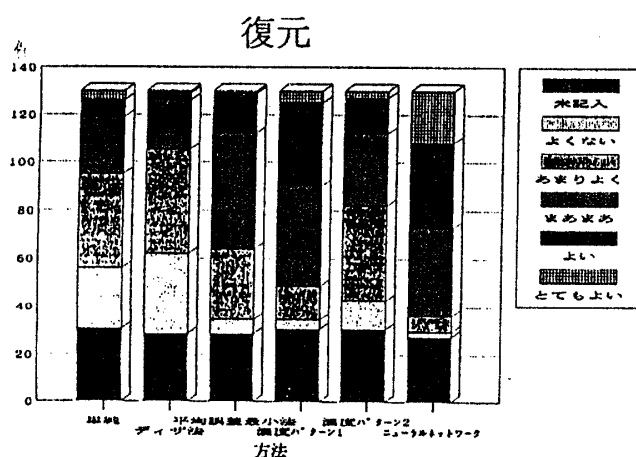


図18 復元についてのアンケート調査

7 結論

今回の研究において、平均化による復元法の評価は、ほぼ2値化した画像の質の評価と同一のものになると言えるであろう。つまり、いかにうまく情報を残しつつ2値化するかが最大の課題であり、この点において、Binamics法による2値化は、有効性が高いものであると思われる。

また、画像の復元法について、本研究では、単純平均を用いた手法によるものとその手法にエッジを強調する処理を追加した手法により実験を行い、文字領域については、エッジ強調がある程度有効であったが、絵柄領域はほとんど変化が見られなかった。

文字と絵柄領域の混在する画像の2値化法として、提案する手法が有効なものであることが確認され、また、2値化法の優劣により単純平均に基づく画像の復元の有効性も左右されることになる。絵柄領域のエッジ強調を実現することが今後の課

題である。

謝辞

今回の研究にあたって、アンケート調査を担当していただいた近松康子女史、およびアンケートにご協力いただいた多数の方々に感謝する。

参考文献

- [1] 高木幹雄, 下田陽久: 画像解析ハンドブック, 東京大学出版会(1991).
- [2] 千種康民, 服部泰造, 池上宗光, 田中衛: "Binamics法による文字混在濃淡画像の2値化" 信学秋期大会, D-203 (1991).