

Introduction

白い紙に黒く印刷されたものをビデオカメラで撮影しても、本当の意味での白黒ではない。カメラがどこを向いていてもグレースケール（またはカラー）なのだ。机の上の照明をうまくコントロールしないと、机の上に置かれた紙をビデオで撮影しても、実際の紙のイメージとはかけ離れたものになってしまう。スキャナーやコピー機の内部とは異なり、机の上では均一な照明を確保することが非常に難しい。机の上には、デスクランプや間接照明、窓、動く影などがあり、それらの影響でページ全体の明るさが変わってしまうことがあります。人間の視覚はこれを補正しますが、机の上の紙のイメージがこれらの変化を考慮せずに機械で操作されると、非常に悪い結果になります。例えば、図5-1では、右側の「in case of emergency」のボックスの背景が周囲の背景と一致せず、はっきりとした縁が残っています。右側のボックスには周囲の背景が表示されています。左から右に向かって濃くなっているのは、紙の左側に光が当たっていたために左から右に向かって暗くなっているのは、この画像を取り込んだ紙の左側に光が当たっていたからです。

この問題は、コントラストの高い線画やテキストを扱うときに顕著に現れます。なぜなら、元の原稿は純粋な白黒であるにもかかわらず、カメラはさまざまなレベルのグレーの画像を生成するからです。この問題は、コントラストの高い線画や文字を扱うときに顕著に現れます。DigitalDeskで試作された多くのアプリケーションは、線画を投影するために、画像のどの部分が黒か白かを明確に判断しなければなりません。線画を机の上に投影したり、テキスト画像を光学式文字認識（OCD）に渡したりするためにこれは、線画を机の上に投影したり、テキスト画像を光学式文字認識（OCR）サーバーに渡したりするためです。このシステムではグレーの画像（通常は1ピクセルあたり8ビット）はシステムで使えないため、白黒の画像（1ピクセルあたり1ビット）に変換する必要があります。これにはさまざまな方法があります。場合によっては、人が見ることを想定して、できるだけグレースケールの画像のように見えるように、ディザやハーフトーン（[Floy76]のような手法）を用いて、以下のような結果を得ることができます。図5-1。この画像はグレースケール画像のように見えますが、黒と白の小さな斑点が様々な形で存在しています。この画像はグレースケール画像のように見えますが、様々なパターン、密度、サイズの小さな黒と白の斑点で構成されています。人間の視覚は、これらの画像をグレー画像や元の白黒画像とほぼ同じように理解します。しかし、文字認識や選択・貼り付け、複数の画像の合成などの機械処理では、ディザやハーフトーンの画像を使うことはできません。これらの画像は、1画素に1ビットしかないにもかかわらず、本来のグレースケール画像よりも機械処理に適していません。そのため、単純なパターンの線や文字、比較的大きな白と黒のパッチが必要となります。グレー画像からこれらのタイプの白黒画像を生成するプロセスは、一般的に「閾値処理」と呼ばれており、デジタルデスクシステムの重要な要件となっています。

画像の閾値処理には様々な方法がありますが、基本的なプロセスは、各グレーピクセルを検査し、それを黒にするか白にするかを決定することです。本レポートでは、DigitalDeskでの閾値処理のために

開発・テストされた様々な技術について説明し、最後に適切と思われるアルゴリズムについて説明しています。このアルゴリズム（ここでは "Quick Adaptive Thresholding "と呼びます）は、最良のものではないかもしれませんが、十分に機能しますので、ここで紹介した段階では、DigitalDeskの他の問題の方が重要になりました。

Global Thresholding

ある意味では、閾値処理はコントラスト強調の極端な形、つまり明るいピクセルをより明るく、暗いピクセルをより暗くすることと捉えることができます。画像を閾値処理する最も簡単な（そして最も一般的な）方法は、あるグレーレベル以下のすべてのピクセルを黒に設定し、それ以外のピクセルを白にすることです。問題は、このグレーレベルをどのように選択するかです。例えば、ピクセルが8ビット（0～255）であれば、128が選択されることになります。この方法は、画像のすべての「暗い」ピクセルが128以下の値を持ち、明るいピクセルが128以上の値を持つ場合には有効ですが、画像の露出が過剰または不足している場合には、結果としてすべてが白または黒になってしまいます。閾値を決める際には、可能性のある値ではなく、実際の値の範囲を見たほうがよいでしょう。画像の各ピクセルの最大値と最小値を求め、その中間値をしきい値とすることができます。さらに良い方法は、実際の値の範囲だけでなく、その分布を見てしきい値を決めることです。例えば、白地に黒の線画や文字が描かれている画像の場合、ほとんどのピクセルは背景の濃さになり、小さいながらも濃いインクの割合が多くなると考えられます。画素の強度のヒストグラムは、図5-2のようになります。

背景の大きなピークと、ダークインクの小さなピークが見えるはずですが、周囲の明るさによっては、曲線全体が左右にずれることもあります。いずれにしても、2つのピーク間の局所的な最小値を閾値として選ぶのが最適です。理論的には問題ありませんが、実際にはどのように機能するのでしょうか。図5-3は、この手法が有効なフレームグラフ画像とそのヒストグラムです。平滑化されたヒストグラムには2つの顕著なピークがあり、このヒストグラムに曲線を当てはめたり、2つのピークの間点を取ったりして、理想に近いしきい値を計算することは難しくありません。しかし、この画像は白と黒の両方が多く含まれており、典型的な画像ではありません。DigitalDeskでは、図5-4のような画像の閾値も設定できなければなりません。この画像のヒストグラムでは、小さい「暗い」ピークがノイズに紛れてしまっているため、ピーク間の局所的な最小値を確実に見つけることができません。そのため、画像を閾値処理する際には、次のような戦略が有効です。

1. ヒストグラムを算出する（図5-4の点で示す）。
2. ヒストグラム曲線の移動平均の最大値を求めることで、大きなピークを見つけることができます。移動平均は、図5-3と図5-4の曲線のように、最大値に対するノイズの影響を滑らかにします。
3. このピークと最小値の間の距離に一定の割合でしきい値を選択します。このピークと最小値（ゼロカウントを除く）の間の距離の一定割合で閾値を選択します。

実験によると、この距離の1/2であれば、非常に明るい場所からほとんど真っ暗な場所まで、幅広い照明条件で非常に良い結果が得られるようです。例えば、図5-4では、ピークが215で、最も低い記録値が75なので、しきい値は145になります。次の図（図5-5）は、4つの異なる照明条件で撮影した画像に、このヒストグラム・グローバル・スレッシュホールド・アルゴリズムを適用した結果を示しています。ヒストグラムからわかるように、さまざまな照明条件にもかかわらず、このアルゴリズムはそれぞれのケースで適切なしきい値を選択しており、それぞれのしきい値の結果はほぼ同じになっています。

このヒストグラムベースのグローバル閾値処理技術は、光が均一に当たっている画像や、上の例のように、光があまり変化しない紙のごく小さな部分を表す画像では、非常にうまく機能します。しかし、より広い領域や、机の上で明るさの範囲が大きく異なる通常のオフィスの照明下では、良い結果を得ることができません。また、画像全体に単一のグローバルしきい値を使用するため、画像の一部が白すぎたり、暗すぎたりします。その結果、図5-6のように、テキストの多くが読めなくなります。

不均一に光っている紙から良好なしきい値画像を生成するには、適応型しきい値アルゴリズムが必要です。この技術は、各ピクセルの背景の照度に応じて、画像全体のしきい値を変化させるものです。以下では、図5-6の画像に適用したアルゴリズムの結果を説明します。この画像は横から照明されており、白地に黒の文字（「PaperWorks」）、黒地に白の文字（「XEROX」）、白地に灰色の文字（「The best way...」）、さまざまな影、「PaperWorks」の文字の下に細い水平の黒線があるため、難易度の高いテストとなっています。PaperWorks "の文字の下に細い水平の黒い線があります。