

「C言語によるパターン認識プログラミング」 第2週

○パターン認識とは概略（復習）

パターン認識とは、画像に映っているパターンが、予め準備したパターンのどれに当たるかを識別することである。なお、準備パターンはあくまでもサンプルであり、たとえば▲パターンでも、実際に識別する画像には、雑音や誤差等が生じていると考える。なお、回転や大きさの変化を許すかどうかなどは目的によって予め決めておく。たとえば、■と◆は回転させれば相似形である。これを違うパターンとするか、同じ「四角」パターンとするかは、決めておく必要がある。

画像処理によるパターン認識には、大きく分けて二つあり、一つはテンプレートマッチング、もう一つは構造や特徴などの構造マッチングである。どちらの方法も一長一短がある。

テンプレートマッチングは、識別すべき物体のサンプル画像を切り抜いてテンプレートとし、認識をしたい画像に当てはめていき、ぴったりと重なったら識別できたとする方法である。

構造マッチングは、テンプレートマッチングと違って、画像そのもので照合（マッチング）を行わず、一度数値などの特徴を抽出して、識別サンプル（つまりデータベース）と識別対象（つまりテストデータ）を比較していく方法である。特徴としては様々なものが考えられ、たとえば物体の面積といった幾何学的特徴などが利用される。

ここでは簡単のため、2値画像（白か黒のモノクロ画像）を対象とし、白を0、黒を1として表現するものとし、また画像の大きさも小さいものを扱う。

○エッジ抽出概略（復習）

エッジとは、画像における変化する部分である。つまり白から黒、もしくは黒から白に変化するような部分をエッジと呼ぶ。エッジを抽出する方法の基本は微分である。画像処理における微分は通常引き算が利用される。たとえば白(0)から黒(1)に変化する場合、 $1-0=1$ を微分値とする。但し画像データは2次元であるので、縦横方向の変化を同時に考える必要がある。

○ヒストグラム（度数分布）概略（復習）

画像データの幾何学的特徴を算出するのに利用される手法の一つである。

画像データは二次元であるので、縦方向・横方向の二種類のヒストグラムが考えられる。また原画像だけでなく、エッジ画像のヒストグラムも抽出可能である。

例えば▲画像の縦方向のヒストグラムは▲になり、横方向のヒストグラムは∟のようになる。また△画像（▲のエッジ画像）のヒストグラムは、縦方向では横に低く平らで”_”のようになり、横方向だと”J”のようになる。

○面積と上下面積比、周囲長概略（復習）

面積は画像パターンの面積そのものである。一方周囲長は、画像パターンの周囲の長さである。たとえば同じ大きさの正方形内にぴったり収まるパターンを考えた場合、■と●では面積も周囲長も違ってくる。

面積を算出するには、単純にパターンの画素数を数え上げればよい。

上下面積比は、上部分の面積と下部分の面積の比率のことである。上下に対称な図形の場合は、この比は1.0に近くなる。なお、面積と同じように、上や下部分の面積をそのまま特徴として扱うことも可能である。

周囲長は、エッジ等の周囲を求めてからこの長さを求めるが、理想的に周囲が抽出されていて隣の画素との長さを上下左右だけでなく、斜めでも1とするならば、エッジ画像の面積、つまりエッジ画像のエッジ画素数を数え上げればよい。

○パターン認識の方法

パターン認識をするには識別したいパターンの特徴を把握しておく必要がある。たとえば、パターン A は面積が大きく & 周囲長が短い、またパターン B は面積が小さく & 周囲長が長いという特徴があれば、識別したいテスト画像の面積と周囲長を観測して、どちらのパターンかを決定させる。なお、この場合、どちらのパターンでもない、例えば面積が小さく & 周囲長も短ければ、どちらのパターンでもないとするのが自然である。但し、テストパターンとして A か B のどちらかしか出現しないことが分かっている場合は、どちらに近いかで判定する。

・二つのパターンに識別する場合を考えてみよう

より具体的な方法を考えてみよう。

まずパターン A とパターン B の二つを対象に識別したい場合を考える。ここではパターン A は▲、パターン B は●としよう。

この二つのパターンは、理想的な形状の場合、面積で識別できそうなことは前回に説明した。今、認識対象とする画像は、18x18 にほぼ一杯に収まるものとする。なおプログラムでは都合上 20x20 画素を用意して、一番外側の周りの画素にはパターンは現れないものとする。

そこで、もしパターン A▲とパターン B●の二つを識別したいとした場合、各々パターンがどの程度の面積になるのか、ノイズや誤差を考慮して、具体的な数値を考えて想定しなければいけない。

また、パターン C を◆とした場合、これとパターン B●の二つを識別できる特徴も、上記と全く同じように、面積で識別できるはずである。

では、パターン A▲とパターン C◆の 2 パターンを識別することを考えてみよう。前回説明したように、理想的なパターンだとすると、この二つは同じ面積になるので、残念ながら面積ではこれらを区別することはできない。よって、別の特徴を考える必要がある。

特徴としてヒストグラムを考えてみよう。ヒストグラムは、{原画像、エッジ画像} × {縦方向、横方向} の 4 種類が考えられるが、A, C 二つのパターンに対するこれらヒストグラムの形は、下図のようになる

パターン	原画像	原画縦 Hist	原画横 Hist	エッジ画像	エッジ縦 Hist	エッジ横 Hist
A	▲	▲ 小大小	▲ 小 中 大	△	— 小小小	└ 小 小 大
C	◆	▲ 小大小	◀ 小 大 小	◇	— 小小小	├ 小 小 小

原画像縦方向のヒストグラムは A, C は同じで区別できない。またエッジ画像縦方向のヒストグラムも A, C 共にほぼ同じで区別できない。一方原画像横方向、エッジ画像横方向のヒストグラムは A, C 別の形であるので区別可能そうである。

原画像横方向ヒストグラムについて考えてみると、A は上の方は小さく（低く）段々増え下

の方ではかなり大きく（高く）なる。一方 C は中ほどが大きい上下部分では小さい。つまり、横方向ヒストグラムを三つの部分 {上、中、下} に分けて考えた場合、A=小・中・大であり、C=小・大・小となる。

同じようにエッジ画像横方向ヒストグラムに注目してみると、C は全体的に平坦であるが、A は下部分が大きくなっている。つまり、横方向ヒストグラムを三つの部分 {上、中、下} に分けて考えた場合、A=小・小・大であり、B=小・小・小となる。

より具体的に考えると、たとえばヒストグラムの上部分とは、 $i=1\sim5$ の行とし、その 5 行の度数合計が 10 以下ならば小、10～60 ならば中、60～100 ならば大などと判定することが考えられる。さて、このような数値はどのように決めればよいのであろうか。本来これらの判定基準値は、数学的かつ統計的に求めるべきであるが、ここではアルゴリズム作成者（つまり自分自身）が十分に考えた上で妥当かつ適切に決めればよい。いずれにしても「上部分」や「下部分」や、「大きい」とか「小さい」といった定性的な表現を、具体的にどのようにするかを決めたり、数値に置き換えたりする必要がある。

なお、周囲長についても同様に各々のパターンで違いが現われるのか考えてみるべきである。

ちなみに周囲長は、今回のプログラムでは結構誤差が大きいことにも注意が必要である。微妙な判定はできないと思った方がよい。

※ヒストグラムによる判定プログラムについて

ヒストグラムを 3 か所に分類してそれぞれに対して大中小を判定している。判定の基準は適宜設定する必要がある。注意が必要なのは一組の基準しか設定できないこと。つまり原画像のヒストグラムとエッジのヒストグラムの両方を利用するにはそれなりにプログラムを改変する必要がある。エッジのヒストグラムの方が明らかに基準値が小さいが、両方を設定できるようにはなっていないから。

・三つのパターンに識別する場合を考えてみよう

さて、上記のように、パターン A▲と B●、B●と C◆、また A▲と C◆がどのような特徴で区別できるのかを検討した。

それではパターン A,B,C の三つのパターンを識別させることを考えてみよう。まずヒストグラムを考えてみよう。先ほどと同じようにパターン A,B,C のヒストグラムがどのようになりそうかを下図に示す。

パターン	原画像	原画縦 Hist	原画横 Hist	エッジ画像	エッジ縦 Hist	エッジ横 Hist
A	▲	▲	▲	△	—	┘
B	●	●	●	○	—	
C	◆	▲	◀	◇	—	

このように A,B,C のヒストグラムを考えると、残念ながらどのヒストグラムを見ても B,C を明確に区別できない。よって、ヒストグラムは、B,C を区別するにはよい特徴ではないと言えるので、別の特徴を利用する必要がある。

もし B と C を区別する特徴があれば、たとえば、ヒストグラムで「A▲」と「B●もしくは C◆」をまず区別し、後者の場合は、更に B と C を区別する特徴で認識させればよいことになる。

○より工夫を！より柔軟で正確な識別を！！

このようにどのような特徴を使うかは非常に重要である。

上記では、エッジ横ヒストにより、A と {B,C} を分け、次に B と C を分けるのに面積を利用することを考えた。しかし、面積は B と C を区別するには実際は微妙かもしれない。理想的には確かに B●の面積はそこそこ大きいはずが、ノイズや誤差があると微妙となってくる。よって、B と C をより明確に区別する特徴がないかを考えるとよい。

例えば、ヒストグラムは縦方向と横方向しか考えていなかったが、斜め方向を考えてみるのも手である。エッジ画像の斜め方向のヒストグラムは、B と C を区別するのにかなり有効な特徴であることがわかる。

また、3つのパターンをよくみると、B,C は点対称な図形であるのに対し、A だけは点対称ではない。これを利用するにはどうしたらよいであろうか？図形の偏り具合を数学的に表現することも十分に可能であるが、ここでは面積を少し工夫して考えてみよう。たとえば、画像を上半分と下半分に分けて考える。パターン A▲は、上半分の面積は小さく下半分は大きい、一方パターン B●も C◆も、上半分と下半分の面積はほぼ同じである。つまり、上面積／下面積を計算すると、パターン A は小さな数字 (0.5 程度?) となり、パターン B,C はほぼ 1.0 になるはずである。

勿論この上下面積比も万能ではない。今、パターンは回転を許さず、■はパターン C◆とは違うパターンであるとした場合は、この上下面積比は有効であるが、回転を許し、■もパターン C◆の変形であるとした場合には、この特徴は利用できなくなる。

いずれにしても、様々な変形・回転や雑音などを許して、柔軟にパターン識別させることは、非常に難しく、人間のパターン識別能力の素晴らしさには脱帽せざるを得ない。近年パターン識別、特に一般物体識別に対する新たな効果的手法が開発されて盛んに研究&実用化されてきた。たとえば、「一般的な物体」として「人の顔」を検出するデジカメなどはかなり普及してきている。それでも識別・抽出の正確さ、識別処理時間や、学習数、学習時間などなど様々な面で問題点は残されており、コンピュータの目が、人間に近づくにはまだまだ遠い未来と言えよう。

一方で、難しそうなことも、ちょっとした工夫によって簡単な方法で解決できることがあることも知っておこう。