

2015 年度春学期 画像情報処理 第 15 (補) 回 パターン認識についての話題

局所特徴量

第 15 回の講義では、パターン認識の基本的な考え方について説明し、パターン認識とはパターンから得られた特徴ベクトルを分類することである、と述べました。

では、どのような特徴量を使って特徴ベクトルを作ればよいのでしょうか。これにはさまざまな考え方があります。特定の対象に関する問題であれば、それに応じて特徴量を考えることができます。例えば、画像で顔を認識する問題なら、顔の目・鼻・口といった各部分の位置関係を測定して、特徴量を構成しています。

一方、より一般的な問題、さまざまな問題に応用できる特徴量を構成するにはどうすればよいでしょうか。理想的な特徴量は、次のような性質を持つべきだと考えられます。

- **識別力 (discriminability)** が高いこと。パターン認識とは分類ですから、ある特徴量を用いて分類したとき、別のカテゴリーに分類されてもらいたいパターンが同じカテゴリーに分類されてしまうようでは、特徴量の能力は低いことになります。
- **不変性 (invariance)** があること。顔を認識するのに、少し顔が傾いたら別人と認識されてしまうようでは、認識装置としては使えません。他にも、大きさの違いなど、影響を受けないほうがよい変形はいろいろあります。
- **頑健性 (robustness)** があること。画像に雑音が入っていたり、照明のむらがあるだけで、正しい認識ができなくなってしまうようでは困ります。

これらの性質は、必ずしも同時に満たされるものではなく、むしろ相反する要求であるといえます。これらの望ましい性質をなるべく持つような、一般的な問題に応用できる特徴量が、さまざまなやり方で研究されています。

それらのうち、近年広く使われるようになってきているのが**局所特徴量**です。「顔」などといった特定の物体の認識ではなく一般的な問題に対応できるということは、画像中の物体を広い範囲で見た時にはその形はさまざま、特定の特徴を取り出すことはできません。しかし、画像を狭い範囲で局所的に見れば、どの画像にも共通の特徴を見出すことができます。そこで、画像中の各部分についてその近傍のみの情報で特徴量を構成するという方法が考えられてきました。ここでは、そのうち 2 つの手法を紹介します。

高次局所自己相関

ひとつの画像中に含まれる各画素値をすべて合計すると、それを画素数で割ることで画素値の平均、すなわち画像全体の平均的な明るさを求めることができます。これを拡張して、各画素値についての特定の方向に隣接した画素との関係の、画像全体での平均を求めるのが、高次局所自己相関特徴 (Higher-order Local AutoCorrelation, HLAC) です。

HLAC をもとめるには、各画素値について特定の相対位置 (たとえば「右上」) にある画素値との積を求め、これを画像全体で合計します。相対位置を一通り指定すると、ひとつの特徴量が得られます。相対位置が 1 カ所だと通常の自己相関ですが、2 カ所以上になると「高次相関」が得られます。画像においては、各画素値の周囲 3×3 画素を対象とし、そこから 1 つまたは 2 つの相対位置を指定します。

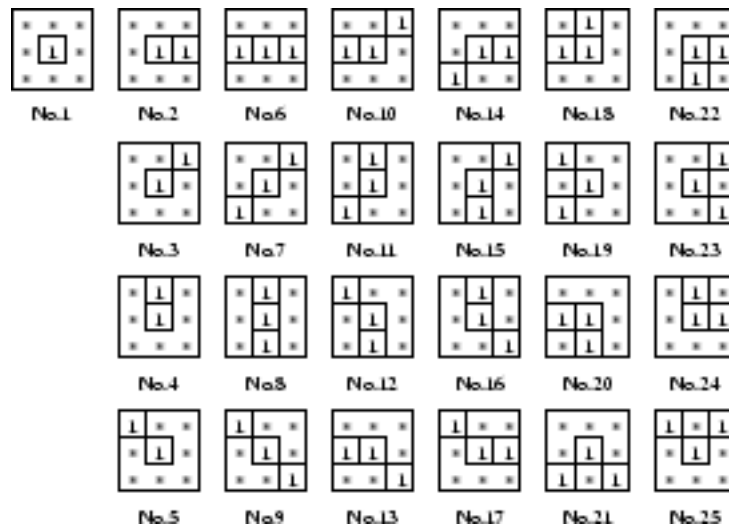


図 1: HLAC のための相対位置の指定。これらで 25 個の特徴量が得られる。(http://home.hiroshima-u.ac.jp/tkurita/lecture/statimage/node6.html より)

この方法では，相対位置に対応して特徴量を求めているので，移動に対して不変です。計算としては大変簡単で，高速に演算できる特徴量です。しかし，「一定方向の隣接画素との関係」というのは，画像の見た目の特徴をよく表しているらしく，さまざまな画像認識の問題でよい結果を出しています。このアイデア自体は 30 年以上前に提案されたものですが，近年広く使われるようになり，また，画素値の空間的關係だけでなく，色の関係や動画像に対しての時間的關係にも応用されています。

文献

- N.Otsu, and T.Kurita, “A new scheme for practical, flexible and intelligent vision systems,” Proc. IAPR Workshop on Computer Vision, 431-435, 1988.
- 小林匠，大津展之，画像特徴量 [VI・完]：高次局所自己相関に着目した画像特徴量と画像認識への応用，電子情報通信学会誌，94, 4, 335-340, 2011.

SIFT

SIFT(Scale-Invariant Feature Transform) 特徴量は，画像中から認識に必要と思われる特徴点を抽出し，その特徴点の周辺での画素値の勾配を特徴量とするものです。この方法は，次の 2 段階からなります。

- **特徴点の抽出** これには，ガウス関数による平滑化を，平滑化の度合いを変えながら行い，平滑化の度合いに対する差分をとる DoG(Difference-of-Gaussian) 演算を用います。ある度合いで差分が大きくなったとすると，それはその度合いで画像中のエッジ（物体の輪郭）などの特徴ある部分が平滑化されたことを意味します。そこで，その差分を選び，さらにその画像の中で平滑化により変化している点を選びます¹。
- **特徴量の算出** 各特徴点について，そのまわりの画素値の勾配を，特徴点から各方向に向けて調べます。そして，勾配が極大となる方向を基準方向とし，特徴点の周辺領域での画素値の勾配を，基準方向にもとづく相対座標で表して特徴量とします。

¹実際には，この方法で選ばれる特徴点は多すぎるので，さらに種々の条件で絞り込みます。

この方法は、あらかじめ特徴点を絞り込んでいることで識別力が高く、また特徴量を画素値勾配の大きな方向を基準とした相対量で表しているため、回転や大きさの変化に対する不変性があります。また、特徴量は各特徴点で正規化されているため、照明変化に対する頑健性もあります。これらの長所があるため、種々の画像認識や、変化する画像の対応付け（マッチング）に用いられています。

文献

- D. G. Lowe, “Object recognition from local scale- invariant features”, Proc. of IEEE International Conference on Computer Vision (ICCV), pp. 1150-1157, 1999.
- 藤吉弘亘, “Gradient ベースの特徴抽出 - SIFT と HOG - ”, 情報処理学会研究報告 CVIM, 160, 211-224, 2007.
<http://www.vision.cs.chubu.ac.jp/SIFT/>

なお、この手法は Lowe と Univ. British Columbia が米国特許を保持しています。(Method and apparatus for identifying scale invariant features in an image and use of same for locating an object in an image, US Patent 6,711,293 B1)

ブースティング

第14回で説明した「誤差逆伝播法」や「カーネル法」は、線形分離ではない複雑な分類問題を解くための工夫として開発されたものです。これらの方法は、学習・分類を行うシステム（学習器）はひとつで、境界を複雑に「曲げる」ことで問題に対応しようとしたものです。

これに対して、もっと単純な分類手法による学習器（弱学習器）を集めて、それらの結果を「投票による合議」で合わせることで、高い分類能力を得ようという考えが**ブースティング**です。この考え方は、学習器のもつアルゴリズムを組み合わせてより上手に運用しようという手法で、アルゴリズムのアルゴリズム、すなわち「メタアルゴリズム」とよばれるものです。

ブースティングでは、それぞれ弱学習器に分類を行わせ、それらの弱学習器のもつ合議の際の「発言力」を調整します。これは、ある事例に対して間違った分類をした弱学習器には、その事例に対する発言力を低下させ、別の弱学習器にその事例を重点的に学習させる、といった方法で行われます。

ブースティングを含めて、「ランダムに答えを出すよりはまし」という弱い学習器を組み合わせて、高い識別能力を得ようとする考え方を**アンサンブル学習**といいます。

文献

- Y. Freund and R. E. Schapire, “Experiments with a New Boosting Algorithm,” Proc. 13th Int'l Conf. on Machine Learning, 148-156, 1996.
- Y. Freund, R. Schapire, 安倍直樹, “ブースティング入門（特集・計算学習理論の進展と応用可能性）”, 人工知能学会誌, 14, 5, 771-780, 1999.