

人間は、人物を認識するのに、この輪郭とか目を先ず認知します。そうして、耳とか、鼻とかを認知して、顔を把握します。そういて、頭を把握し、胴体を把握します。全て、輪郭線が中心の解析です。さらに、輪郭線のうち、目の輪郭や内部構造、外部の眉毛、他の顔部品との配置という属性の基に、優しい目であるとか、怒っている・笑っているなどの判断を下します。顔から身体全体の認知に至り、何をしているかというシーンの認知が行われます。

人間の場合、画像・図形認識は線画に先ずイメージ（2値画像）を落とすことから始まります。物事は、オブジェクトと運動と、オブジェクト間の配置関係である格とオブジェクト・運動・格の属性で表現され、それらは入れ子構造をしており、枠とその属する部品群という関係を持って把握され、枠は輪郭線として認識されることを基本とするから、線画というものが認識の第1位にあるものだと言えるのです。

線画は線分が連なったものです。線分の列として線画認識されるものです。そうして、線分間の関係・・・位置関係という情報と曲率とかコーナーの角度とか長さとかの属性が(key,value)という関係で把握（マップ上にこの(key,value)が分布する）されます。線分間の配置関係はポテンシャル法で解析できます。属性は画素解析で得られます。

この(key,value)セットは、多くの画像を入力されると、重なり具合が造られていきます。一つのkeyに振られるvalueが密度分布をなすのです。そこで、keyについて画像種にたいしてクラスタリングを行い、画像を分類できるようになっていきます。ここで、(key,value)のkeyの設定ですが、枠の中の一つの部品を表しているとしているだけで、部品もゴミなのか、大事な本当の部品なのか精確には判断できません。曖昧なのです。だから、部品を特定していくこともクラスタリングになるのです。そのクラスタリングの根幹は、極座標による配置の把握です。枠の中の部品候補の重心を原点として、その部品候補との距離と角度（配置の向き）を(key,value)として部品であることの確信度をクラスタリングで得ていくのです。そうして、明確な部品の名前を決定していくこととなります。

枠の認識は閉じた曲線の検出で行います。枠の根本は、総体的に配置が不変なものの集まりの境界ということです。それは、画像では、輪郭線になるのです。しかし、これも見方によって、曖昧な物です。滲みや断線があります。二重線ということも、藪ですと線分の塊の何とはなしの境界というものもあります。このため、枠の推定は、部品と輪郭線として認識する線分の量ということで、ディープラーニングで決定していくこととなります。