

目次

- Hough 変換
- テンプレートマッチング
- パターン認識とは
- 機械学習とは
- 教師あり学習と教師なし学習

★9 パターン情報処理の応用例 (2) — 画像処理いろいろ (承前)

★9.3 Hough 変換

Hough 変換 (☆1) は、画像の中から直線や円といった図形を検出する方法の一つである。図 1 に、Hough 変換による直線検出の原理と方法を示す。

☆1) Hough 変換 (ハフ変換, Hough transform): P. Hough が考案した。

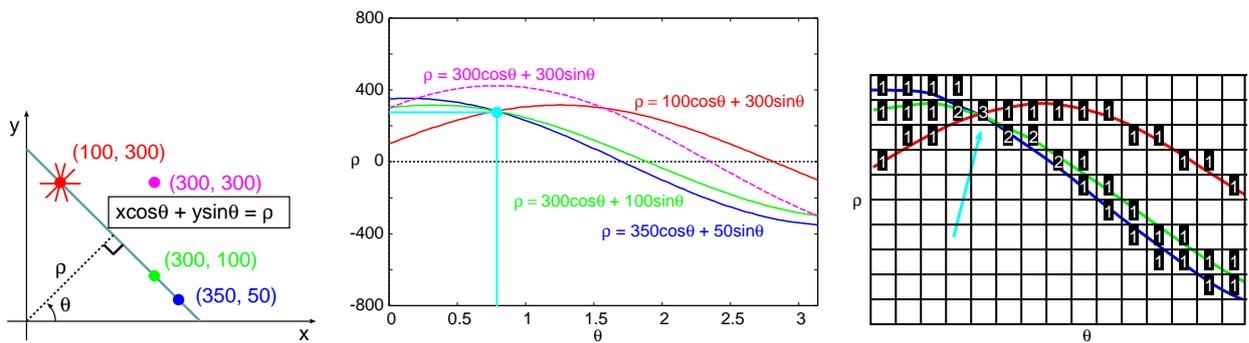


図 1: Hough 変換。

図 1 左: 平面上の点 (x, y) を通る直線は、図に示す ρ, θ という二つの値を用いて、

$$x \cos \theta + y \sin \theta = \rho \tag{1}$$

という式で表せる (☆2)(☆3)。図中の灰色は、点 $(100, 300), (300, 100), (350, 50)$ を通る直線であり、この例では、 $(\theta, \rho) = (\frac{\pi}{4}, 200\sqrt{2})$ となる。

図 1 中: ある点を通る直線は、様々な ρ と θ の組み合わせで無数にあるが、それら全体は (θ, ρ) 平面上で正弦曲線を描く。例えば、点 $(100, 300), (300, 100), (350, 50)$ のそれぞれを通る直線たちは図に示すような 3 つの曲線を描き、その交点の (θ, ρ) の値は、この 3 つの点を通る直線を表す。

図 1 右: (θ, ρ) 平面を格子状に分割したものに対応する二次元配列を用意し、画像中の各点について、その点を通る直線たちを表す曲線上の配列要素に一票を加える、という投票を行うと、画像中に存在する直線に対応した配列要素に票が集まることになる。

図 1 では少数の点しか描いていないが、実際には画像から何らかの方法 (☆4) で抽出した多数の特徴点のそれぞれについて (θ, ρ) 平面上で投票を行い、得票数の多い (θ, ρ) を選び出す処理を行うことになる。この場合、1 位のみでなく、一定以上の票数を集めた (θ, ρ) を全て直線の候補とするといった扱いを考えると、画像中から複数の直線を検出することもできる。図 2 に実際の画像での直線検出例を示す。

☆2) $y = ax + b$ のような表し方も考えられるが、その場合、 a の動く範囲を無限大までとらないと任意の直線を表せないため実用的でない。式 (1) の θ, ρ ならいずれも有限の範囲で済む。

☆3) 画像上の座標は、左下を原点にして右と下に x 軸 y 軸それぞれの正の向きをとって表すことが多いが、ここでは簡単のため上に y 軸正の向きを選んでいる。

☆4) フィルタリングによって輪郭部分を抽出し、二値化する等。

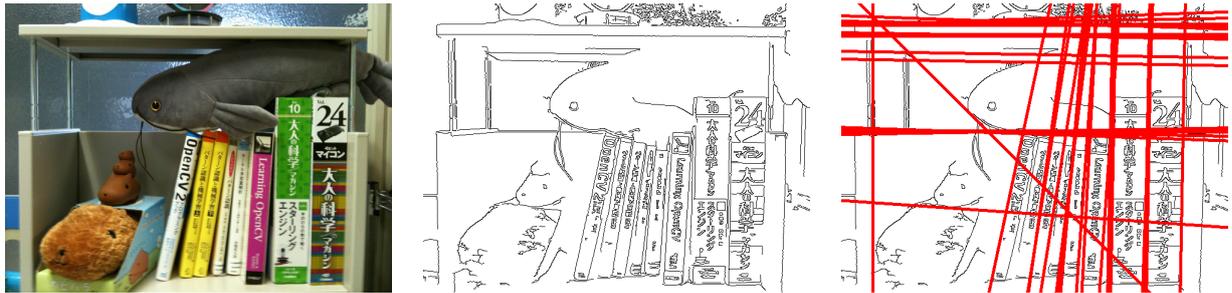


図 2: Hough 変換による直線検出例. 左から, 入力画像, フィルタリングによって抽出した特徴点, 直線検出結果.

ここでは直線検出を例にあげたが, Hough 変換は, いくつかの変数の組み合わせで表せるような図形であれば適用可能である (☆5). ただし, 変数の数が多くなると投票する空間の次元数が大きくなって難しくなる. また, 単純な直線検出の場合でも, 前処理である特徴点抽出の結果によって検出結果が大きく変わるので, 有効な結果を得るためには前処理の方法等も注意深く設定する必要がある.

☆5) 例えば円ならば, 中心の座標と半径の計 3 つの変数の組で表せる.

★9.4 テンプレートマッチング

画像中にあるものが存在するかどうかを調べる方法を考えよう. 最も原始的な方法は, 見本画像 (テンプレートと呼ぶ) を用意し, 入力画像中にそれと似た領域がないかどうか比較する, というものである. テンプレートと入力画像との比較は, 位置を少しずつずらしながら何度も行なう (図 3 左上). このような方法を, **テンプレートマッチング**という.

テンプレートマッチングでは, テンプレートと入力画像の部分領域との**類似度** (似ている割合) を位置毎に測り, この値があらかじめ定めた基準を満たせば, テンプレートに写っているのと同じものがその位置に存在すると判断する. 類似度の計算には, テンプレートと入力画像の一部との**平均二乗誤差** (MSE) がよく用いられる (☆6). これは次式で与えられる.

$$\text{MSE}(x, y) = \frac{1}{wh} \sum_{i=0}^{w-1} \sum_{j=0}^{h-1} (I(x+i, y+j) - T(i, j))^2 \quad (2)$$

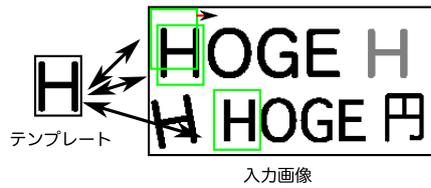
ここで, $\text{MSE}(x, y)$ は, 入力の (x, y) を左上隅とする $w \times h$ 画素の領域とテンプレートとの MSE であり, 値が小さいほど類似していると考ええる. また, $I(a, b), T(a, b)$ は入力とテンプレートの位置 (a, b) の画素値を表し, w, h はテンプレートの幅と高さを表す. 図 3 にテンプレートマッチングの結果の例を示す (☆7).

テンプレートマッチング: template matching.

類似度: similarity.
平均二乗誤差: mean squared error.

☆6) MSE が小さい方が類似度が大きい. この他, 類似度としては正規化相関係数もよく用いられる.

☆7) この実験ではテンプレートの位置をずらしているだけであるが, 実用的には, 物体の大きさや向きの変化にも対応するため, 入力画像を拡大縮小したり回転したものともマッチングするなど, より複雑な処理が必要となる.



左の図は、各位置での MSE の値を明るさで表したものの。黒いほど MSE が小さい (類似度が高い)。

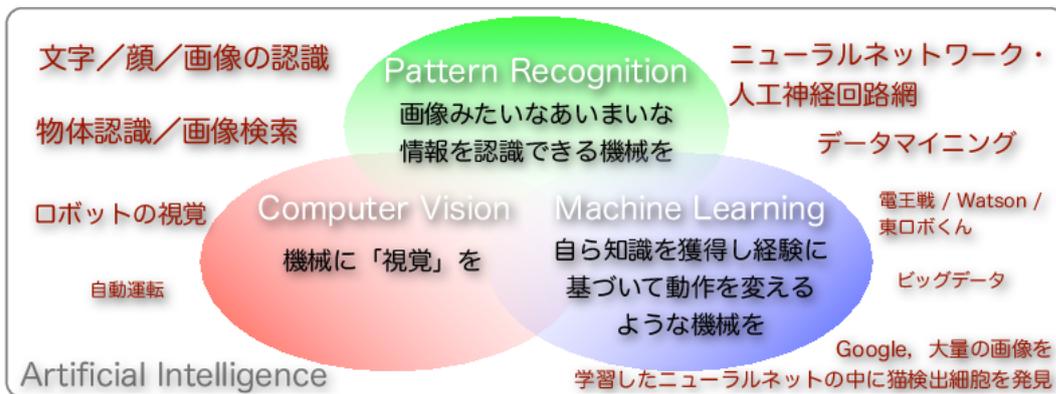


図 3: テンプレートマッチング。下段の 3 つの画像では、MSE の値が基準より小さかった位置にテンプレートと同サイズの赤枠を表示している。左のものほど MSE の基準が小さい (厳しい)。

★ 10 パターン認識と機械学習 (1) — パターン認識, 機械学習とは

今回から、**パターン認識** (☆8) と **機械学習** (☆9) を取り上げる。以下の図は、これら 2 つの分野と関連の深い **コンピュータビジョン** (☆10) を加えて、これらの分野とその周辺のキーワード (のうち特に最近のもののごく一部) を示したものである (手抜きして高橋の卒業研究説明会資料用の図を貼付けただけ (^_^;))。

- ☆8) パターン認識: pattern recognition
- ☆9) 機械学習: machine learning
- ☆10) コンピュータビジョン: computer vision



★ 10.1 パターン認識とは

パターン認識とは、コンピュータを用いて、画像や音声のようなパターン情報の中の規則性を自動的に見つけ出し、それに基づいてパターンをいくつかのカテゴリに分類するような情報処理のことをいう (第 1 回講義資料も参照のこと)。音声認識, 文字認識, 画像認識などが典型例である。例えば...

- 例 PR1: 母音の音声波形から「a,i,u,e,o」を識別する (波形を「a」, 「i」, 「u」, 「e」, 「o」の 5 つのカテゴリに分類する)
- 例 PR2: 人の顔画像をいくつかのカテゴリに分類する (カテゴリの設定の仕方次第で様々な応用が考えられる)
- 例 PR3: 様々な動物の写った画像が与えられたときに、それぞれの画像に写っている動物の種類を識別したい (ネコ, ペンギン, カピバラ,...)

パターンをカテゴリ分類する処理とは少し異なるが、類似の情報処理として、次のような問題もパターン認識で扱う対象となる。

例 PR4: 顔画像から年齢を推定する

例 PR5: 大量の画像の中から、ある画像と似た画像を見つけ出す

近年のパターン認識では、研究レベルでも実用レベルでも、後述する機械学習を利用した方法を採用することがほとんどである。

★ 10.2 機械学習とは

機械学習とは、人間のように自ら知識を獲得し経験に基づいて動作を変えるような機械/コンピュータ/アルゴリズムの実現を目指す研究分野またはその技術を指す。人間が行なっているような知的な情報処理の仕組みを人工的に実現することを目指す**人工知能**(☆11)の一分野であるが、現在ではより幅広く、データの中から何らかの知識や法則性を自動的に見つけ出したり、データに基づいて何らかの判断を自動的に下したりするような情報処理全般を対象とするものとなっている。大量のデータが与えられたときに、そこに潜む規則性を自動的に見つけ出し、そこから何らかの情報を取り出すような情報処理は、機械学習の応用の例である。例えば... (以下の括弧の中は、取り出す情報がどのようなものを表す)

例 ML1 時々刻々変化する気温、湿度、電力使用量の数値から、少し未来の電力使用量の推定値を求めたい (未来の電力使用量 (実数値))

例 ML2 通販サイトの購買履歴データをもとに、顧客を買い物傾向の似たグループに分類したい (例えば「グループ 1」から「グループ K 」までの K 通りのいずれか)

先のパターン認識の例も、機械学習の問題の一つと考えることができる。

例 PR1: 母音の識別 ('a', 'i', 'u', 'e', 'o' の 5 つのカテゴリのいずれか)

例 PR2: 顔画像の認識 (顔/非顔, 男性/女性, A さん/B さん/C さん..., etc.)

例 PR3: 画像中の動物の認識 (ネコ/ペンギン/カピバラ/...)

例 PR4: 顔画像から年齢を推定する (年齢の数値)

これらの処理は共通して、何らかのデータ x が入力されると適当な値 y を出力する関数として次式のように表すことができる。

$$y \leftarrow \boxed{f} \leftarrow x \qquad y = f(x; w) \qquad (3)$$

入力 x は、問題に応じて様々な形式で与えられる。出力 y (☆12) は、例 ML1 や PR4 のように多様な数値をとる場合もあれば、例 ML2 などのように何通りかの値しかとらない場合もある。ここで w は、関数 f に含まれるパラメータ (変数) (☆13) である (☆14)。パラメータを変えると、同じ入力に対しても異なる出力を与えることになる。このように式で表現すると、機械学習とは、「入力に対して「望ましい出力」が得られるように、パラメータ w をデータから自動的に決定する方法」に関する分野/技術と言うことができる。

機械学習において、望ましい出力が得られるようにパラメータを調節する過程を**学習**(☆15)といい、その際に用いるデータを**学習データ**(☆16)という。学習データで望ましい出力が得られることは当然大事であるが、機械学習の目標は、学習後に、学習時に遭遇したことの無い未知のデータが入力されても、望ましい出力が得られる(これを「汎化能力がある」という)ようにすることである。

☆11) 人工知能: artificial intelligence, AI

☆12) この例は全て出力が一つの値をとるものであるが、一般には出力が複数の値の組であってもよいので、ここでは y をベクトルして表している。

☆13) パラメータ: parameter, 助変数, 媒介変数。

☆14) 例えば、 $y = f(x) = ax + b$ では a と b がパラメータ。

☆15) 学習: learning, 訓練 (training) とも。

☆16) 訓練データともいう。

★ 10.3 機械学習のアルゴリズム・手法の分類

機械学習の方法は、学習データの与え方によって次の 2 つに大別される。

教師あり学習 (☆ 17) 個々の学習データが、入力とそれに対する出力の正解のペアとして与えられる。前述の例 PR4 では、顔画像 (入力) とその人の年齢 (出力の正解) のペアを複数集めたものを学習データとすればよいだろう。特殊な場合として、**強化学習** (☆ 18) を含む。

教師なし学習 (☆ 19) 学習データは入力のみで構成され、出力の正解は与えられない。どのような出力が望ましいかが学習アルゴリズムに反映されており、データの分布などに応じて出力が自己組織 (☆ 20) される。**自己組織学習**と呼ぶこともある。

例 ML1 や PR4 は、教師あり学習の典型例である。一方、ML2, PR1,2,3 の場合、学習データとしてどのような分類が正解かという情報を与える (「このお客さんはグループ 1」, 「この顔画像は女性」, etc.) なら教師あり学習の問題であるし、入力データだけを与えて自動的に分類させるならば教師なし学習の問題となる。

☆ 17) 教師あり学習: supervised learning.

☆ 18) 強化学習: reinforcement learning. 出力の正解が直接与えられるのではなく、ある入力に対する出力が「良かった」「悪かった」というような評価の値のみが与えられる形態の学習。ロボットの行動の学習などによく用いられる。

☆ 19) 教師なし学習: unsupervised learning.

☆ 20) 自己組織: self-organization.

★ 10.3.1 教師あり学習

教師あり学習の問題は、式 (3) の出力 y がどのような値であるかによって、**回帰** (☆ 21) と**識別** (または**分類**) (☆ 22) の二つに大別できる。

回帰 出力として量的な値を扱う場合。例 ML1 と PR4 が該当する。

識別 入力をいくつかのグループ (**クラス**または**カテゴリ**という) に分ける問題。上記の 2 つの例以外の上述の例は、教師あり学習の問題として扱うならば全てこちらに入る。例えば例 PR1 なら 'a', 'i', 'u', 'e', 'o' の 5 クラスを識別する問題といえる。

☆ 21) 回帰: regression

☆ 22) 識別/分類: classification

★ 10.3.2 教師なし学習

教師なし学習の適用対象となる問題の代表例を以下にあげる。それぞれについて様々な学習アルゴリズムが存在する。

クラスタリング データをいくつかのかたまり (**クラスタ** (☆ 23)) に分類する。

特徴抽出/次元圧縮 高次元のデータから本質的な情報を抽出する、情報を保持したまま低次元に変換する。この授業でパターン成分分析の手法として学んできたもの (ベクトルの直交展開, フーリエ変換等) と関連が深い。

頻出パターン抽出 データ中に一定の頻度以上で出現するパターンを見つけ出す。

これらは、大量のデータの中から有益な情報を見つけ出す**データマイニング** (☆ 24) の対象とする問題でもある。教師なし学習とデータマイニングという 2 つの分野は大きく重なっている。

☆ 23) クラスタ: cluster, クラスタリング: clustering.

☆ 24) データマイニング: data mining. 英語の動詞 mine には、(鉱石などを) 採掘する/掘り出すという意味がある。