

04 マシンビジョン／検査

マシンビジョンという名前のとおり、この分野ではカメラとコンピューターを機械の視覚として扱い、ヒトがこなしてきた作業の機械での代替を目的としています。マシンビジョンとは産業界でのコンピュータービジョン分野の技術利用を指します。

◎ 進化する産業界の外観検査

産業界においては人間による視覚的なタスクが多く、外観検査のエキスパートが良品と不良品とを目視で見分けています。これらは、形状のいびつさ、バリの有無、凹み、塗装ムラ、異物混入など多様な不良品に対して各々のエキスパートが長年かけて修得した技術であり、同一のタスクでも同じ方法で検知しているとは限りません。

技術を若手に継承するアプローチでは、視線計測やモーション解析などが用いられていますが、それらの情報だけで体得できるとは限りません。エキスパートの方法論に基づいて、画像処理技術者が特徴量を選び、機械学習で識別器を作る検討も行われてきました。また、深層学習を利用した分類器の利用も試みられています。

◎ 外観検査・検品タスクと画像処理技術の関係

工業分野において「形状の正しさ」「滑らかさ」「均一さ」などが視覚的なタスクとして挙げられます。「形状の正しさ」については、手本となるサンプルと同一か確認するものです。よって「同一」は厳格な一致を意味しますので、カメラからの距離が若干変わる、あるいは照明が変わるだけで判定が変わってしまいます。したがって、対象物、照明、カメラの配置や対象物の向きを固定する、背景に余計なものを入れない、といった制限により**マシンビジョン**の難易度を下げる工夫が必要です。

「滑らかさ」はプレス加工における打痕やせん断加工でのバリなど、対象物

の表面あるいは断面が滑らかに保たれているかを確認します。金属のようなよく反射する素材の場合には、不自然な反射の有無を検知することになります。

「均一さ」は明るさ、色などについて均一さを確認します。カメラにおける入射光と濃度値の関係が線形で、撮像素子のダイナミックレンジと分解能が十分であれば、明るさの不均一さを画像上の濃度値の違いとして捉えられます。

ただし、ヒトの感じる明るさや色（測光量）は、画像での赤、緑、青の濃度値からは求めることができません。これは、ヒトとカメラとでは光の波長に対する応答関数（分光感度）が異なり、ヒトの目と同等な分光感度を有する特殊なカメラを使わない限り、絶対量の計測が原理的に不可能なためです。

色については赤、緑、青の濃度値から均等色空間への変換により相対的な色差を利用するか、特定の条件下においてあらかじめ色見本を撮影したうえで測光量への換算やデバイス間での色再現を行う場合があります。なお、厳密な測色は波長ごとの光を計測したうえで測光量を算出する方式やヒトの目と同等な分光感度を持つセンサーで光を計測する方式による、分光測色計や色彩輝度計を用いるのが一般的といえます。

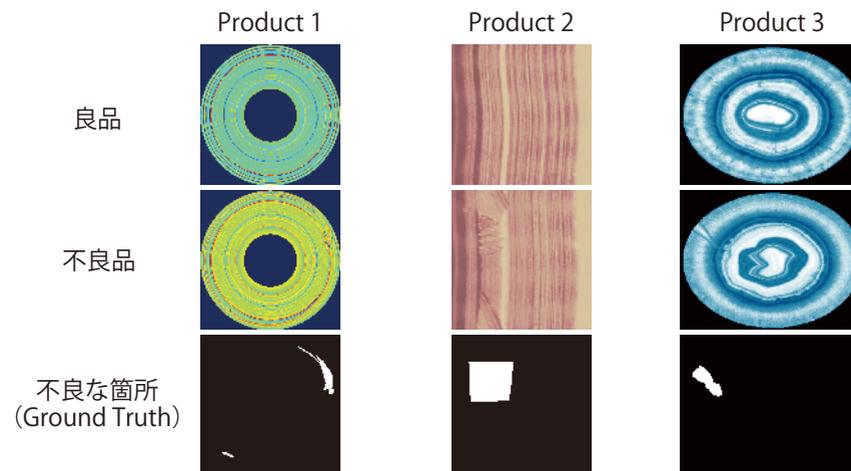


図 04-1 画像検査のベンチマーク [8]

11 イメージセンサー

イメージセンサーは、レンズで集光した光信号を電気信号に変換し、画像情報として記録する大変重要な役割を担っています。ここでは、イメージセンサーの基礎知識と撮像方式について説明します。

半導体イメージセンサーとCCDセンサー

映像を電気信号として扱うデバイスは、半導体プロセスの高密度化が進むにつれて電子ビームの走査を利用するものから、固体撮像素子である**CCDイメージセンサー**を用いたものに置き換わり、さらには**CMOSイメージセンサー**へと置き換わっています。CCDイメージセンサーやCMOSイメージセンサーはどちらもフォトダイオードで光の量を測っていますが、読み取り方式が異なります。

CCDイメージセンサーは構造が複雑なものの、各画素のフォトダイオードの出力をバケツリレー方式で転送するので、暗流が少ない、グローバルシャッターが利用可能、電気信号の変化を強調する増幅器が出口に1つで、画素単位の特性のばらつきが少ないといった利点から長らく主流でしたが、CMOSイメージセンサーの高性能化によりその座を明け渡しています。

- 1：光を電荷に変換し蓄積する
- 2：電荷を転送する
- 3：電荷を電気信号に変換する

外部から取り込んだ光の情報を、
デジタル処理するための形に変換する

図 11-1 イメージセンサーの役割

CMOSイメージセンサー

CMOSイメージセンサーは、増幅器や出力するか否かを制御するゲートを画素ごとに搭載する方式で、登場した当初から消費電力の少なさが利点でしたが、1画素あたりの受光面積が小さくなりがちで暗所での撮影には向かず、かつ1画素ごとに増幅器の特性が異なりノイズが多いといった問題もありました。

また、撮影の際はゲートを順次オンオフするため、画素間の撮影のタイミングが同時ではなく、ローリングシャッターと呼ばれる現象も発生しました。現在では製造プロセスの容易さから高度なノイズ除去処理や、1画素ごとにメモリを搭載したのもも開発されており、これらの欠点は克服されたといえるでしょう。

CCDイメージセンサーでは、すべてのフォトダイオードの電荷が垂直転送CCDに移され、その後水平転送CCDにより1行ずつ出力されます。

CMOSイメージセンサーでは、行・列選択回路で出力するフォトダイオードを1つずつ指定します(図11-2)。

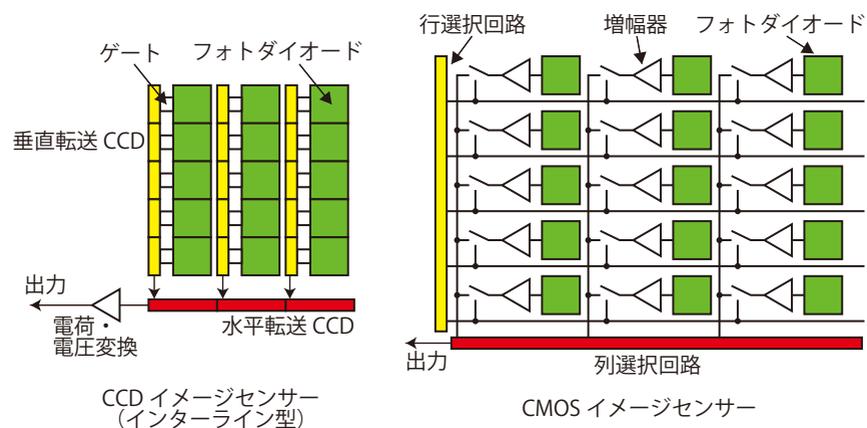


図 11-2 CCDとCMOSの仕組み

16 カラー情報の扱い

ピクセルに色を記録するとカラー画像となります。一筋縄ではいかない色情報の扱い方について少し丁寧に説明します。

物理現象としての色の扱い

色情報の扱いは一筋縄ではいきません。「光源」「物体」「視覚」の3つが揃わないと、色を扱うことはできないのです。また色は、物理現象としての「光の色」の扱いと人の視覚に「映る色」の扱いの2つがあります。さらには、印刷やテレビ、デザインなどそれぞれの用途特有の歴史と流儀があります。

画像データをカラー画像として扱う物理現象の手掛かりの始まりは、図16-1のように、**光の三原色**R（赤）、G（緑）、B（青）にあります。色を光の三原色である赤、緑、青（Red、Green、Blueの頭文字をとって**RGB**と表記されます）のスペクトル強度の組み合わせで色を表現します。

RGB表色系では、原色をR（赤・700nm）、G（緑・546.1nm）、B（青・435.8nm）とする表色系を、国際照明委員会（CIE）のRGB表色系といいます。パソコンにおいては最も多く用いられ、ディスプレイデバイスなどへの出力指示などに用いられます。

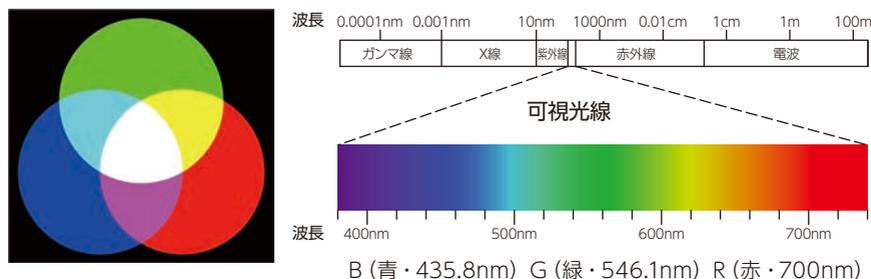


図 16-1 光の三原色 [21] と光のスペクトル

国際照明委員会は、色の表示を標準化するために[R][G][B]という原色を、赤原色[R] = 700nm (nm=ナノメートル)、緑原色[G] = 546.1nm、青原色[B] = 435.8nmの単波長光と定め、これらの原色を用いた「等色実験」を重ね、「等色関数(人間の標準的な色覚を数値化した関数)」を定義しています。「等色実験」は三原色を調整しながら混色し、等色した(等しい色に見えた)その瞬間の三原色の各混色量を測るものです。

それぞれの赤と緑と青を何段階で表現するかによって、色の階調とデータサイズが変わります。グレースケールと同様に、階調はデータのサイズ数がnビットの場合、 2^n (2のn乗)となります。8bitなら各色256階調なのでフルカラー(full color)で16,777,216種類の色表現ができます。

人の視覚に映る色の扱い

画像データの色表現、つまり表色系の技術は多岐にわたりますが、その基本的枠組みについて要点を説明します。

まず、表色系には大きくマンセル表色系(Munsell)とオストワルト表色系(Ostwald)の2つがあります。また、目で見たい色の属性として、色相、明度、彩度という3つの尺度が設けられていて、それらは次のとおりです。図16-2はマンセル色相環、図16-3は色相、明度、彩度の立体図です。

●色相 (Hue)

5色相(R・Y・G・B・P)に分け、それぞれの間をさらに5色相(YR・GY・BG・PB・RP)を加えて補完し、大まかな10色相が設定されています。

●明度 (Value、Brightness)

理想的な黒を0、理想的な白を10とする11段階で設定されていますが、理想的な白や黒は物理的に表現できないため、現実的な数字として黒は1程度、白は9.5程度が用いられます。

●彩度 (Chroma、Saturation)

無彩色を0、有彩色は鮮やかになるほど数値は高くなります。色相や明度に

18 濃淡変換処理

画素ごとに明るさのレベルを濃淡値で表現したものを濃淡（グレースケール）画像と呼びます。ここでは、画像処理の基本となる濃淡画像の扱い、ヒストグラムや濃淡変換処理について説明します。

◎ 画像の濃淡とヒストグラム

デジタル画像を構成する最小要素が画素（ピクセル）であり、そこに格納された値を画素値と呼びます。ここでは、さまざまな目的により、この画素値を変換する処理について説明します。

話を簡単にするため、カラー画像ではなく明るさの度合いが各画素に格納されている**濃淡画像（グレースケール画像）**を考えます。例えば、明るさの度合い（＝濃淡レベル）が8bit（＝256階調）で表現されているとしましょう。その場合、画素値0は真っ黒、画素値255が真っ白を表すことになります。図18-1に、8bit階調で表現された濃淡画像の例を示します。また、この画像の**ヒストグラム**を図18-2に示します。

◎ ヒストグラムから見る画像の統計量

この画像ヒストグラムからは、その画像が持つさまざまな情報を読み取ることができます。画像から最小値、最大値、平均値、中央値、最頻値、分散、標準偏差といった画像の特徴を表す**統計情報**を計算することができますが、ヒストグラム分布を確認するだけでも、こういった統計情報のある程度推測することが可能です。

まず、画像中の画素値中で最小のもの（一番暗い画素）を最小値、最大のもの（一番明るい画素）を最大値といいます。8bitの濃淡画像の場合、0～255の画素値を表現できますが、最小値が0の画素が多数存在するような画像、グラフの山が横軸の左端で切れてしまっているような場合は、画像中の暗い部分が

つぶれてしまっている状態で「黒つぶれ」といいます。逆に最大値が255の画素に集中しているような場合は、明るい部分の情報がつぶれてしまっている状態で「白とび」といいます。

画像撮影時に、ヒストグラムを参照しながら撮影することで、適切な階調範囲に収めることができます。最頻値は、画素値の中で最も頻度が高い値のことをいいます。図18-2においては濃淡値32付近が最頻値となっており、これは画像中の黒板の暗い部分の画素が占める領域が多いためです。その他、平均値は平均的な画像の明るさ、分散、標準偏差からは、画素値のバラツキ具合（集中度合い）がわかります。

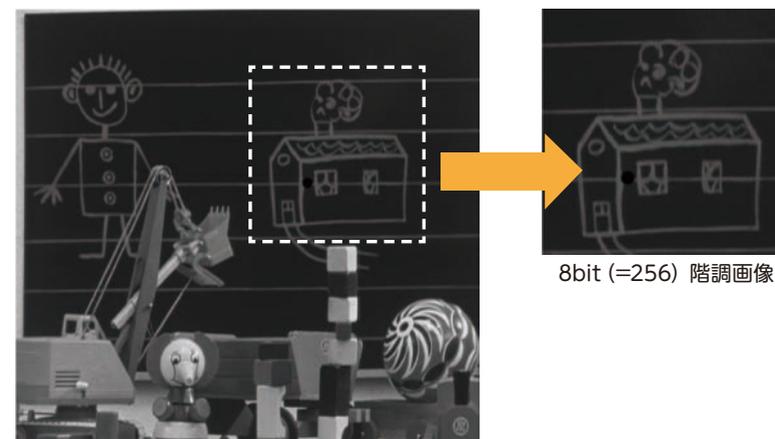


図 18-1 8bit 濃淡画像の例

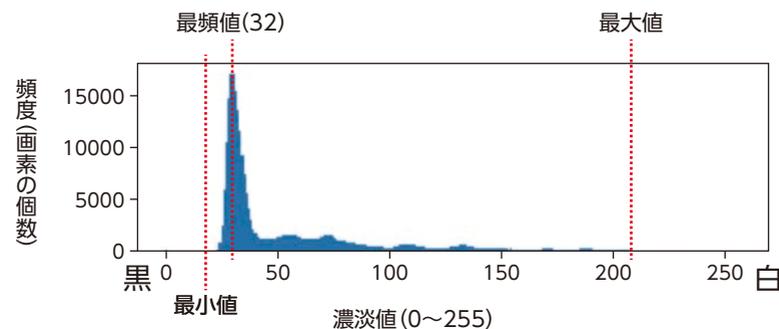


図 18-2 濃淡画像のヒストグラム表示

22

さまざまな画像特徴量

フーリエ変換による周波数特徴量の抽出は、画像解析やフィルタリング、画像復元・圧縮など幅広い分野で利用されています。画像を周波数成分に分解することができるため、周波数空間において問題を解決できます。

○ 画像における空間周波数

2次元の正弦波では、平面を定義する2つの変数があります。そして、関数の値はその平面内の位置に応じて上昇または下降します。図22-1に示す明暗の縞は、ある方向に沿って振幅が正弦波状に変化する2次元の表現です。方向以外に縞模様を制御するパラメータとしては、**振幅**（明暗変化の度合）、**空間周波数**（明暗変化の細かさ）もあります。任意の画像をこのような空間周波数の波の組み合わせで表すことができ、図22-2はさまざまなパラメータを持つ空間周波数の波から画像を構築するイメージを表しています。

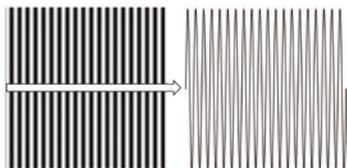


図 22-1 2次元の正弦波上の任意の水平線は1次元のサイン関数になります

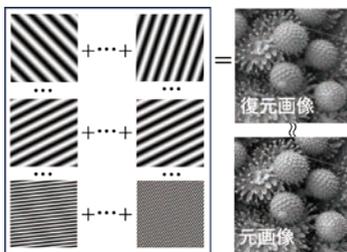


図 22-2 さまざまな空間周波数から画像を復元するイメージ

○ フーリエ変換を利用した周波数解析

フーリエ変換は、ある波形を異なる振幅や周波数、位相を持つ数多くの正弦波に分解できます。画像を2次元の波関数として見なせば、2次元フーリエ変換を用いることで任意の画像空間周波数成分に分解することができます。つまり、図22-2に示される画像復元の逆の工程を辿ることができます。

また、フーリエ変換をコンピューターで実行する際には有限回の計算で終わる必要があるため、画像処理の場合、**2次元離散的フーリエ変換** (Discrete Fourier Transform : **DFT**) を1次元DFTの繰り返しで実現することになります。さらに計算量を減少させるために**高速フーリエ変換** (Fast Fourier Transform : **FFT**) が一般的に画像処理の分野でも用いられています。

FFTを行うことにより、画像から**空間周波数スペクトル**という周波数特徴量を抽出することができます。周波数領域でのフィルタ処理を便利にするため、図22-3で示すスペクトルが一般的に用いられています。

図22-3(右)に示す空間周波数スペクトルの場合、中心から離れるにしたがって高周波数成分になるスペクトルとなっており、画像データの周波数分布を表しています。それぞれの位置に周波数成分が現れます。つまり、周波数の高低と空間周波数の成分配置の関係を示しています。中心付近に白い画素が多いほど画像に低周波成分が多いこととなりますが、図22-3(左)のような細かい模様が多い画像の場合、高い空間周波数成分が大きくなるため、中心付近における白い画素の集中は目立ちません。逆に、濃淡変化が緩やかな画像の場合、高い空間周波数成分が小さくなるため、四隅付近に黒い画素が集中しやすくなります。

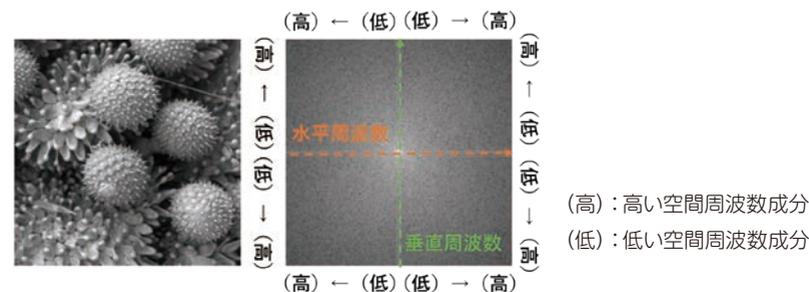


図 22-3 FFT および周波数成分の並び替えによる空間周波数スペクトルの抽出例

26 形状マッチング

ここでは、物体の形状に注目したマッチング方法について説明します。このマッチング方法を活用することで、共通の形状を持つ物体、例えば道路標識や人間の黒目の部分である虹彩などを検出することができます。

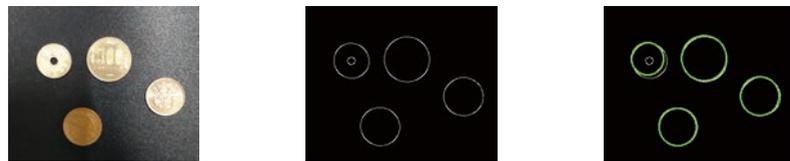
ハーフ変換による円の検出

ハーフ変換は画像中の円や直線を検出できる手法として知られています。ここでは円の検出を例に説明します。

例えば図26-1(a)に示すように硬貨が複数枚映っており、その枚数を数えたいとします。すべての硬貨に共通する特徴としては、形状が円であることです。したがって、画像からすべての円を取得できれば、枚数を数えることができます。

ハーフ変換で円検出をする前には、図26-1(b)に示すようにエッジ画像(物体の輪郭線を抽出した画像)を作るのが一般的です。エッジ以外の画素は不要だからです。このエッジ画像からそれぞれの円の方程式を推定します。

詳細は割愛しますが、半径 r 、中心座標 (a,b) の円を構成しているすべての画素の x, y 座標を (a,b,r) 空間に射影すると、すべて同じ1点にプロットされるという特徴を使用しています。逆に、円を構成していない画素は別な点にプロットされます。同じ1点にプロットされている数が多いほど円である可能性が高いため、その (a,b,r) 座標から円の方程式を導くことで、円を検出できます。図26-1(c)に検出結果例を示します。すべての硬貨を検出できていることがわかります。



(a) 対象画像

(b) Canny 法によるエッジ検出結果

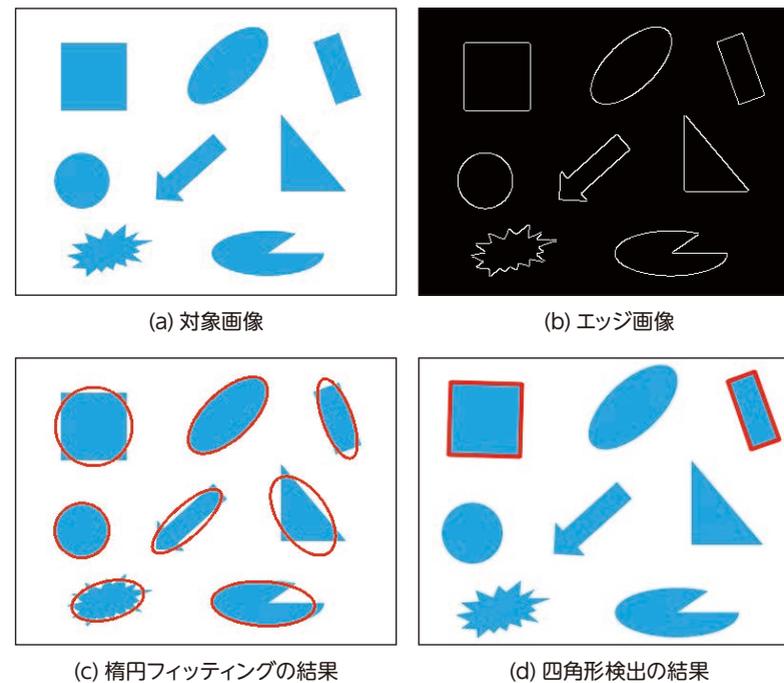
(c) ハーフ変換による円の検出結果

図 26-1 ハーフ変換による円検出の例

楕円と四角形の取得

用途によっては、円以外に楕円や四角形を画像から取得したい場合があります。例えば図26-2(a)の画像から、楕円と四角形を取得することを考えます。まず前処理として、ハーフ変換による円検出と同様に、事前にエッジ検出といった形状情報のみを残す処理を実行します(図26-2(b))。次に、エッジ画像から輪郭検出処理を実施し、楕円の場合は得られた輪郭に対して最小二乗法により楕円を表す式を推定します。このように図形群がクラスター化されている場合は最小二乗法も有効です。その結果を図26-2(c)に示します。

取得された各図形の輪郭に対して最も誤差が最小となるような楕円が取得されていることがわかります。図26-2(d)は四角形の検出結果を示しています。こちらは4つの頂点があるという情報を利用しているため、四角形のみ検出できています。



(a) 対象画像

(b) エッジ画像

(c) 楕円フィッティングの結果

(d) 四角形検出の結果

図 26-2 楕円や四角形検出の例

32

顔画像認識

機械学習の適用によって、飛躍的に認識性能が向上した事例として、顔画像認識があります。ここでは、機械学習による画像認識事例として身近な、顔画像検出と顔画像認証について解説します。

○ 顔検出

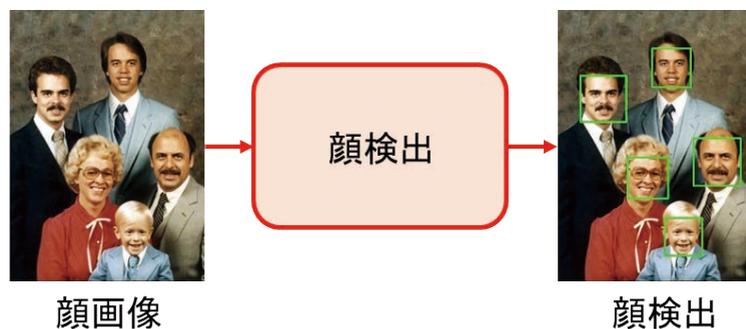


図 32-1 顔検出 (画像は WIDER FACE [35])

図32-1のような画像から1人ずつの顔の領域を検出する顔検出 (Facial Detection) は、デジタルカメラにも使用されており、身近なものになっています。画像中の顔を検出する研究は、コンピュータービジョンにおいて長く研究されてきたテーマの1つです。

顔の形状や目、鼻、口の位置関係は多くの人で共通しているため、これらの情報を活用すれば画像から顔を検出することは簡単に思えます。しかし、図32-2をご覧ください。この図は顔検出のベンチマークとして公開されているデータセットであり、すべて現実世界で撮影された画像です。

この図を見ると、顔の形状や位置関係だけの情報ですべての顔を検出することは困難であることがわかります。もちろん、証明写真のような固定された環



図 32-2 顔検出のベンチマークとして公開されている WIDER FACE dataset [35]

境で撮影された画像であれば、簡単な画像処理の組み合わせで正確に検出や認識が可能ですが、撮影環境が固定されていない画像で顔検出や認識をする必要がある場合、多様な顔画像を含むデータセットを用いた機械学習手法を用いる必要があります。

顔検出の手法としては、Viola と Jones の2人が考案した **Viola-Jones 顔検出アルゴリズム** [36] が有名です。Viola-Jonesアルゴリズムは、Haar-like特徴、積分画像、Ada-boost学習、およびカスケード分類器の主に4つの処理から成り立っています。画像を小領域に分割し、検出はスライディングウィンドウ方式で小領域ごとに実行されます。特徴量の設計は研究者自身が行っていますが、多くの特徴量の候補の中から、顔の検出に有効なものを学習により選択している点が重要です。

○ 顔認識

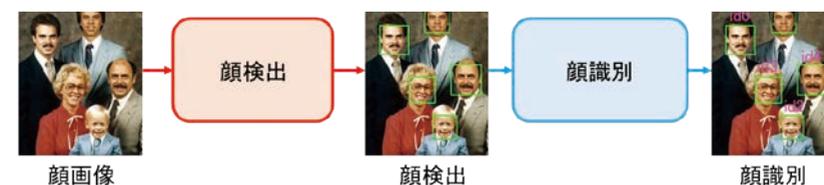


図 32-3 顔認識の流れ (画像は WIDER FACE[35])

図32-3のような複数の人物が含まれる画像に対して顔認識し、1人ずつ区別して認識する (IDを付ける) 場合は、まず図32-3の左のような画像から1人ず

43 3D認識

2012年のAlexNet登場以来、主に2次元(2D)の画像認識への技術進歩が進歩しましたが、2017年のPointNetの提案以降からは徐々に3次元(3D)への物体認識/再構成への注目度が高まっています。

◎ 深層学習による3D認識

3D認識(3D Recognition)とは、3Dデータを入力して対象物体を認識するタスクになります。深層学習を利用した3D認識には図43-1に示すようなボクセル、点群、メッシュなどさまざまな3Dデータ表現が利用されており、近年では陰関数表現まで多岐にわたる表現が用いられています。ボクセルは立方体の集合、点群は3次元座標の集合、メッシュはN角形の集合により、3D物体の形状を表現しています。入力データの表現形式の多様化に伴いニューラルネットワークの構造を変化させる必要があり、それぞれのデータ表現に相応しいニューラルネットワークが提案されています。

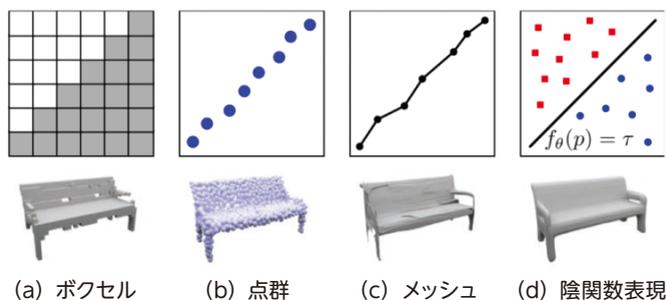


図 43-1 さまざまな 3D データの表現方法 [53]

◎ ボクセルによる3D認識

ボクセルは画像における2次元配列を3次元配列に拡張させた3次元データ表現です。そのため、CNNを3次元に拡張させて使用可能であり、Volume

CNNなどが提案されています。ボクセルでは微細な幾何形状を表現するため高解像度にする必要があります。しかし、高解像度では計算コストの観点から限界があるため、ボクセルには解像度と計算コストの面での課題が大きく、最近では3次元点群のような3Dデータ表現に移行しつつあります。

◎ 3次元点群による3D認識

3次元点群は各座標値の相対的な位置関係により物体の幾何学的構造を表現する不規則なデータ構造です。そのため、画像認識技術をそのまま適用することは困難でしたが、2017年に点群を直接入力可能とした**PointNet**が提案されました。

PointNetでは、点群において重要とされる順序不変性^{*}、剛体不変性^{*}の2つの性質を満たしたうえで、高精度に3D物体を認識することを実現したことが大きな貢献といえます。特に、順序不変性に関してはMax-Poolingを導入することで実現しています。具体的には、図43-2に示すように各点群から特徴抽出後、各点群における特徴ベクトルの最大値のみを利用します。これにより、入力点群の順序が変化しても出力は不変となります。それ以降、点群に対して畳み込み処理を用いるネットワークやグラフ処理を用いるネットワークが数多く提案されています。

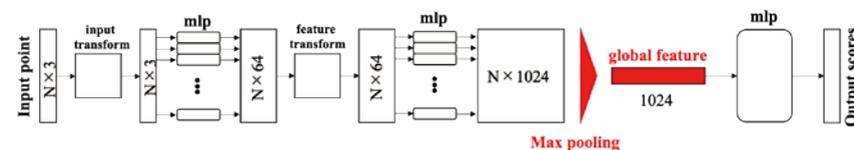


図 43-2 PointNET

まとめ

- ▶ 入力データの表現形式によって最適なネットワーク構造が異なる。
- ▶ 3次元点群の順序不変性、剛体不変性などの性質を満たすPointNetは、高精度な3D物体認識性能を実現している。

^{*}順序不変性：入力点群の順序が変化しても出力は不変である性質。

^{*}剛体不変性：平行移動などのデータ拡張を与え入力座標が変化しても、3D物体の形状に対する出力は不変である性質。

44

異常検知

近年、生産工業における製造工程では、製品にキズなどの欠陥が含まれないか作業員が目視で検査する外観検査の工程を、深層学習を利用した異常検知タスクとして自動化する取り組みが目まぐるしく行われています。

異常検知の概説

異常検知 (Anomaly Detection) とは、図44-1に示すように観測されたデータの大多数が含まれる領域と、ある一定の法則で形成される分布に合致しない領域に存在する異常なパターン (**外れ値**) を特定するタスクです。

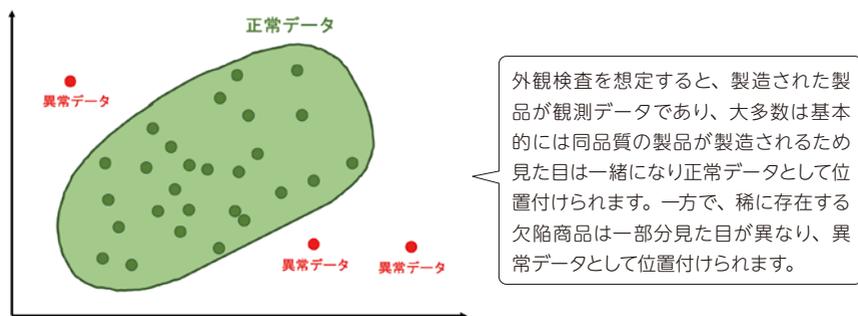


図 44-1 正常データと異常データの分布例

深層学習を用いることで高精度に画像認識することが可能ですが、それには膨大な学習データが必要となります。しかし、欠陥商品すなわち異常データを取得することは稀であり、正常製品と比較して非常に少数のデータになります。そのため、最近では異常検知手法の大多数は教師なし学習問題として扱われています。

教師なし異常検知では、学習時に正常データを学習して、推論時に正常 or 異常を判定する問題設定として定式化されています。さまざまな教師なし異常検知手法が提案されていますが、ここでは基本的な **AutoEncoder** ベースの異常検知手法を紹介します。

AutoEncoder を用いた教師なし異常検知手法

画像生成などのタスクで多様に用いられている AutoEncoder ベースの異常検知手法では、正常データで再構成学習したモデルは異常データに対しては再構成誤差が高くなるという前提に基づいています。

基本的には、図44-2に示すように正常画像を入力し、Encoder-Decoder を介して再構成画像を生成します。正常画像のみを使用して Encoder-Decoder を学習させることにより、学習モデルは正常画像のみに過剰適合することになります。そのため、もし異常画像が入力された場合には、欠陥部分を再構成することは難しくなります。つまり、異常画像は正常画像の場合よりも再構成誤差が大きくなります。再構成誤差が大きい画像に対して異常と判定することで異常画像を検出しています。

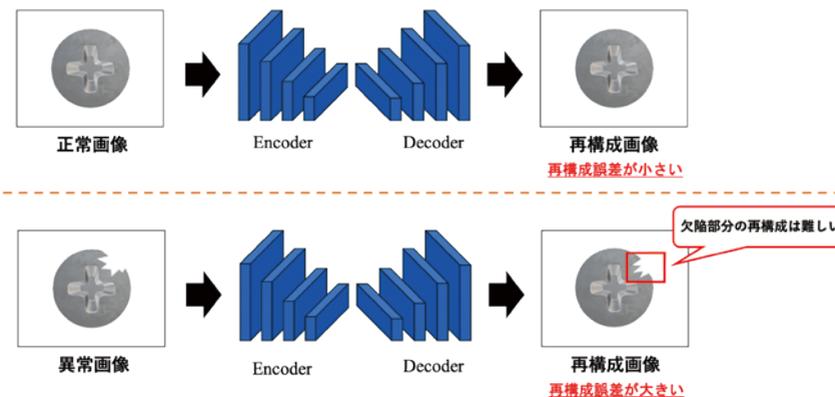


図 44-2 AutoEncoder を用いた教師なし異常検知

まとめ

- ▶ 異常検知は、大多数の観測データが含まれる領域と、その分布に合致しない領域に存在する異常パターンを特定するタスクである。
- ▶ AutoEncoder は、画像再構成誤差の大小により画像上の異常部位を検知することができる。

52

ハードとソフト④
新しいカメラ

画像センシング技術の進歩と合わせて、新しい機能や特徴を持ったカメラも次々と登場しています。ここでは、カラー画像と共に奥行き情報（デプス情報）も取得可能なデプスカメラについて解説します。

◎ デプスカメラ／Kinect

家庭用ゲーム機用の体感型ゲームのためのジェスチャー・音声認識デバイスとして **Kinect** が登場してから約10年経過しています。Kinectの原理は、近赤外線のスポン光を大量に照射して、それらの関係から奥行き（**デプス**）を推定するLight Coding方式です。適用場面は原理的には屋内環境に限られますが、家庭用ゲーム機用のデバイスでありながら高解像度なRGB-D画像が得られるとともに、多人数での姿勢推定を可能としました（図52-1）。



図 52-1 Kinect



図 52-2 OAK-D

Kinect以前から、デプスの取得にはステレオ視やアクティブステレオなどさまざまな方法が考案されていました。一般に光を照射してから反射して戻るまでに時間が反射物体までの距離に依存するのを利用して画素ごとの距離を測る **ToF** (Time of Flight) を利用するデバイス単体でデプスが測れるカメラは非常に高価でした。

わずか数万円のデバイスでVGAサイズのデプス画像が高FPSで取得可能なKinectの登場は、画像処理やロボット関係の研究者にとって衝撃的でした。Kinectの深度センサーはPrimesense社が提供したものでしたが、同社は後にApple社に買収されており、Kinect v2ではToF方式を採用し、より高解像度かつ検出範囲の拡大と性能はアップしたにもかかわらず、初代Kinectのような支持を得られず販売が終息していました。現在はAzure Kinect DKが発売されており、デプスカメラと姿勢推定のほか、音声認識、画像認識のサービスも受けられます。なお、産業用途としてのデプスカメラはこれらの低価格化の余波を受けて、高価であったデプスカメラがBasler Blazeのように数十万円の価格で入手可能となっています。

◎ デプスカメラ／RealSense

Intelは、オールインワンなプロトタイプデバイスであるEuclid Development Kitや、RealSenseシリーズとしてステレオカメラやLiDAR Cameraやトラッキングカメラ、顔認識カメラなどさまざまな方式によるデプスカメラをリリースしていました。しかし残念ながらEuclidは早々に販売終了となり、RealSenseシリーズについても2021年にビジネス的理由でステレオカメラ以外の開発・提供終了を発表しています。

Intel Movidius Myriad X VPUを搭載したSoM (System-on-Module) を利用したLUXonis社のOAK-Dは、新しいデプスカメラです。OAK-Dは、120FPSのステレオ視用カメラ2個と4Kカメラ1個を搭載し、RGB-D画像はもちろん、OpenVINO経由でVPUを使用した推論もカメラ本体のリソースのみで可能です（図52-2）。

まとめ

- ▶ カラー画像と奥行き情報を同時に取得可能なデプスカメラは、性能向上と低価格化が進み、さまざまな用途で活用されている。