



「人間の目」から「機械の目」へ

長い歴史の中で、農業は人間が主体となって行われてきました。しかし近年では農業の担い手が不足しており、また一方で農業自体も大規模化する傾向にあって、効率化や省力化が求められるようになってきています。

このため車両ロボットなどの機械が農作業に使用されるようになりましたが、そうすると、ロボットが作物と雑草を見分けたり、生育状況を判断できなければなりません。判断できなければ、作物を雑草と認識して刈り取ってしまうでしょう。つまり「人間の目」で判断していたことを「機械の目（マシンビジョン）」で見て、正しく判断しなければならないのです。



色と画像判断の関係

RGB 色空間

人間もコンピュータも、基本的には「画像（映像）」でものを検知してそれが何かを判断します。まず、「デジタル画像」について見てみましょう。

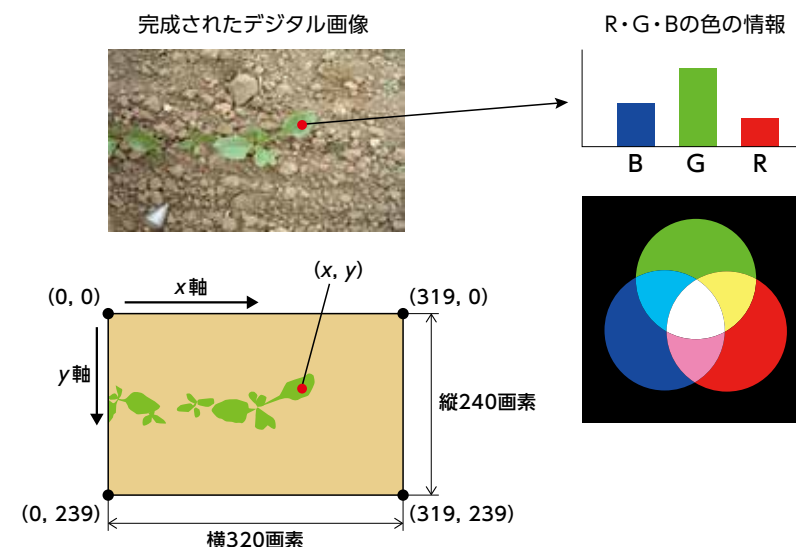
デジタルカメラやスマートフォンのカメラのように私たちがふだん目にしているカメラは、「RGBカメラ」といいます。RGBカメラによって撮影されたデジタル画像は、小さな点がたくさん集まって構成されており、この小さな点のことを「画素（ピクセル）」といいます。そして、この1つの画素の中には、データとして「R（レッド：赤）」「G（グリーン：緑）」「B（ブルー：青）」の3つの色のデータが数字として入っています（図1）。

RGBカメラは学校でも習ったことがある「光の三原色」によって、さまざまな色を生みだしているわけです。このような、赤・緑・青にもとづいてあらゆる色をあらわすものを「RGB色空間」といいます。

それなら、緑色は植物、茶色は地面と、機械にも簡単に判別ができるように思われます。たしかに私たち人間は、地面に生えている植物の写真を見たとき、瞬時に脳が判断して直感的に理解することができますが、「機械の目」は数字データで判断するので簡単ではありません。

さまざまな色は、光の三原色の組み合わせでつくられていますが、「緑

図1 RGB色空間



撮影されたデジタル画像は、たくさんの画素によって構成されている。図は、縦240画素・横320画素を例としている。画面は横がx軸、縦はy軸であらわされ、左上が原点(0,0)で、画面内のある任意の位置の画素は(x, y)であらわされる。1つの画素にはR（赤）、G（緑）、B（青）の3つの色成分（情報）があり、この光の三原色の組み合わせ、つまりそれぞれの色成分のレベルを組み合わせることによって、さまざまな色を表現している。3色が合わさる（R・G・Bが最大）と白で、いずれもない（R・G・Bが0）ところは黒となる。

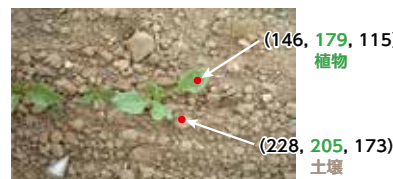
っぱいもの」にも赤や青の成分の情報が含まれており、その相対的なバランスによって色が生み出されています。単純に「緑成分が強い部分を植物と判断する」としてしまうと、もし、土壌の部分の画素の緑成分が他のところより強ければ、コンピュータは土壌の部分を植物と判断する可能性が高くなります(図2)。基本的に、私たち人間の目と同じしくみを使っているのですが、私たちが直感的に判断するのと違い、コンピュータは詳細で膨大な数字情報をもとに画像を判断するわけです。

HSV 色空間

RGB色空間以外にも、色の違いのあらわし方がいくつかあります。その1つが「HSV色空間」で、これは「色相(Hue)」「彩度(Saturation)」「明度(Value)」の値の組み合わせでさまざまな色を表現します(図3)。このうち「色相」とは、色の違いをあらわしています。また「彩度」は鮮やかさをあらわし、低くなるとくすんでいって、最後はどの色も灰色になります。そして「明度」はその名のとおりに明るさをあらわしていて、暗くなっていけば最後には真っ暗、つまり黒になります。

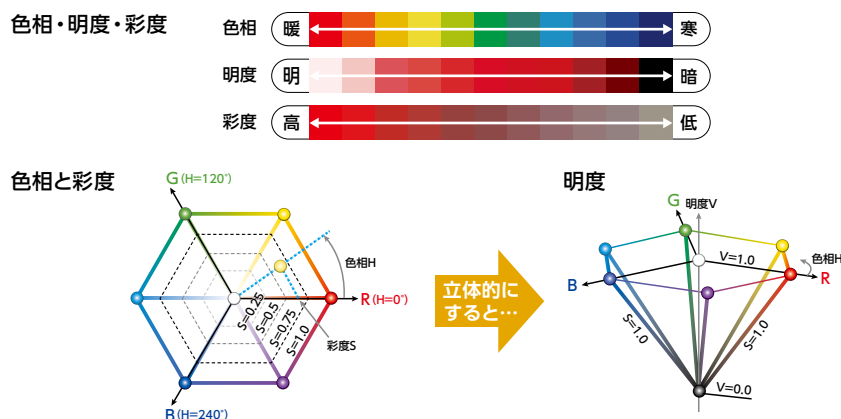
HSV色空間では、たとえば「土と葉っぱを見分けたい」という場合は、色相に注目して「色相の数値が120前後なら緑だ」ということから土と判別できます(図3)。RGB色空間が、色を三原色に分解する人間の目の機能に近いものであるのに対して、HSV色空間は人間が脳の中で色を認識する考え方に近いといえます。

図2 コンピュータによる判断



植物の部分は(146, 179, 115)で、G(緑)が「179」と強い。茶色の土壌部分は(228, 205, 173)でGがほかよりも弱い。しかし、もし「G成分が高い部分を植物とする」としてしまうと、コンピュータは「205」のほう、つまり土壌を植物と判断することになる。

図3 HSV色空間



「色相・彩度・明度」の図はR(赤)を例としているが、他のさまざまな色にも当てはまる。下の2つの図は、色相・彩度・明度をわかりやすく模式図化したもの。色相Rは六角形の0°の位置、G(緑)は120°、B(青)は240°の位置にあり、黄色は60°、水色は180°、紫は300°と移動すると色相が徐々に変わっていく。彩度は、中心(灰色)と外の辺の間の距離となる。これに対し、明度は右の図のように高さのようなもので、下の中心部に行くにつれて暗くなり、いちばん低い位置では黒になる。

図4 2値化



画像処理の基本 —— 2値化とラベリング

「画像に写っている作物の葉を分析したい」とき、背景の地面の画像データは必要ありません。その場合、興味のあるところだけを残して、それ以外を取り除く方法を「2値化^{にちか}」といいます。

図4は2値化された画像で、不要な地面は黒く、分析したい葉の部分は白くなっています。色付きで見たい場合は、白部分を元の画像データに戻せます。2値化は、処理をするときに楽になる他、余分な画像データを取り除くのでデータ量を減らせるというメリットがあります。

2値化によって対象物をしぼっただけでは、まだ終わりではありません。図5を見てください。左側の四角のなかに4つのかたまりがあることが人間にはわかりますが、コンピュータにとっては「それぞれのピクセル(画素)が白か黒か」という情報でしかありません。そこで右側のように、それぞれのかたまりごとに番号をつけることで、「4つのかたまりがある」とコンピュータは認識できるようになります。これを「ラベリング」といい、こうして果実の熟し具合や、虫食いの状況などを個別に判断することができるようになります。



画像処理の基本 ― 形状特徴量

コンピュータなどに機械の目で判断させるには、収穫物などの「形状」も数字であらわさなければいけません。図6のように、丸っこいものとギザギザのものを人間は一目で違うと認識しますが、コンピュータで処理する場合は、「丸っこい」「ギザギザ」のようなあいまいな表現が通用しません。その場合は対象物の「形状特徴量」を見ます。

たとえば、面積はほぼ同じでも、その周囲の長さはギザギザのほうが長いので、「周囲長÷面積」で数値の違いを出して判断させます。このほか、長いほうの径と短いほうの径の比である「長径÷短径」や、対象物の重心が真ん中か、外側に寄っているかを見る「2次モーメント」などがあります。

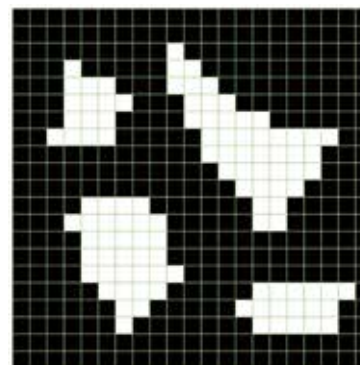


画像処理の基本 ― テクスチャー

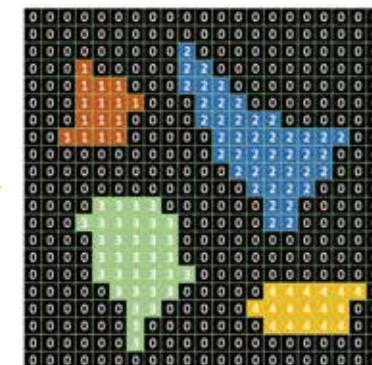
1枚の葉の画像をアップにすると、緑色がのっぺりとフラットに映ります。しかし、牧草地の一部をアップにすると、たくさんの細い植物が折り重なり、すき間は下の地面の暗い部分がところどころ写っていてゴチャゴチャした画像になります。このような、画像の模様を認識する方法が「テクスチャー」です。

テクスチャーを数字で表す方法はいくつかありますが、先の牧草地の例の場合は「明るさのばらつき」を利用します。1枚の葉の明るさはのっぺりとしていて一定ですが、牧草地の画像は一定ではないので、その違いを数

図5 ラベリング

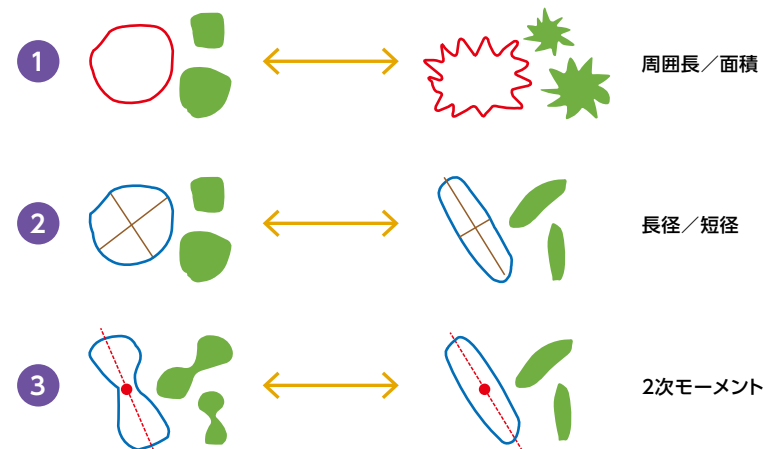


① 2値化により、対象物のかたまりがしぼられる



② ラベリングをすることで、それぞれのかたまりが意味のあるものとなる

図6 形状特徴量



- ① 丸っこい形状とギザギザの形状の面積はほぼ同じだが、周囲の長さはギザギザのほうが長いので、周囲長を面積で割れば違いが数値として出てくる。
- ② 丸っこいほうは長径も短径も同じなので、細長い方との違いがわかる。
- ③ 赤い点はこの形状の重心。ひょうたん型は細長いものにくらべて重心が外側寄りとなっている。

字データとして出します。

今まで説明してきた画像の判断はコンピュータが最終的に行いますが、その際にはAI（人工知能）を用いることができます。AIについては第8節でご説明します。

その他の画像——3次元画像とマルチスペクトル画像

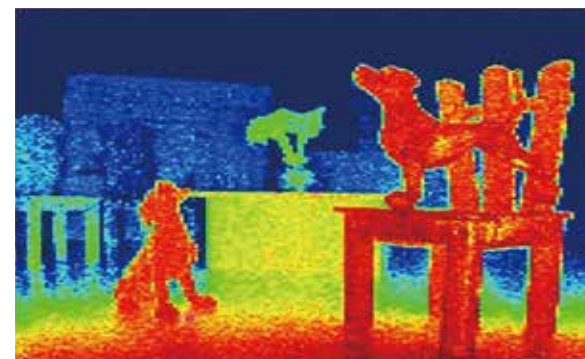
一口に画像といっても、私たちが普段使うデジタルカメラの画像以外にも、さまざまな種類の画像があります。その1つが「3次元画像」です。3次元画像は奥行きを持っています。写真1は奥行きを色で強調した画像の例で、手前にあるものを暖色（赤系）、奥にあるものを寒色（青系）で表現しています。

3次元のカラー画像は、色に関する情報（R・G・B）に距離Dを加えた4つの画素データから成り立っています。距離を測る方法にはいくつかありますが、よく使われるのは、第4節でもご説明した「ステレオビジョン」です。目の前に指を置いて左の目で見、次に右の目で見ると、指の位置がズレて見えます。指が目から近いほどズレが大きく、遠ければ遠いほどズレは小さくなります。ステレオビジョンカメラは、こうした人間の目による距離測定を2つのカメラを使って行います。

またスマート農業では、R・G・Bだけでなく近赤外線などのデータを加えた「マルチスペクトル画像」も利用されています。植物は近赤外線などによく反応するので、このマルチスペクトル画像によってクロロフィル（葉緑素）の量がわかり、それを作物の生育状況を算出する「正規化植生指数（NDVI）」で計算すれば、生育状況をあらわす植生マップ、つまり新しい情報に加工することが可能となります。

3次元画像を撮影できる装置をはじめとする映像機器はかつては数百万円もしたものもあり、たいへん高価でしたが、現在では数万円で購入できる3次元カメラも出てきており、農業分野で活用しやすくなりました。そして、こうした高度な情報を持った画像を処理する技術は第8節で取り上げるAI（人工知能）の発達によって大幅に向上することが期待されています（図7）。

写真1 3次元画像



距離画像 + カラー画像 = 3次元カラー画像

上の写真は3次元カラー画像の一例で、奥行きを寒色系で強調してあらわしたものです。3次元画像を得るには距離を測らなければならないので、撮影には「ステレオビジョン」や「TOF (Time-of-flight)」 「パターン照射方式」などのカメラが必要になる。

図7 画像処理技術とAIが農業の姿を変えていく



人間が目で見えて判断していたことを、これからは機械が見て判断し、より高度な情報を提供してくれるようになるだろう。