

新たなカラー特徴量を用いた 画像認識技術

様々な対象に適用可能な特徴量で 画像認識がより簡単に

近年、デジタルカメラなどの製品では、顔検出機能、ペット検出機能、個人認識機能、シーン認識機能など、数多くの画像認識技術が搭載されるようになってきました。しかし、画像認識技術の開発は、認識に有効な特徴量を認識対象ごとに考えて作り込む必要があり、非常に手間が掛かる作業でした。

東芝は、この手間の発生を防ぐために、様々な認識対象に共通して使用できる特徴量を開発しました。この特徴量を用いることで、従来よりも簡単に画像認識技術を開発できるようになりました。今後、新たな特徴量を用いた画像認識技術を応用して、様々な製品が登場してくることが期待されます。

画像認識技術

画像に映っている“もの”が何であるかを答えることは、人間にとっては簡単なことですが、コンピュータにとっては非常に困難な問題でした。しかし、近年のコンピュータの計算能力向上と画像認識技術の目覚ましい進歩によって、特定の“もの”に限れば答えることができるようになってきました。例えば、最新のデジタルカメラでは個人認識機能が搭載されており、あらかじめ認識したい人の顔を登録しておくことで、映っている顔が誰であるかを判別することができるようになりました。

今後、顔に限らず様々な“もの”を認識できるようになれば、幅広い応用が期待できます。例えば、セキュリティへの応用では、監視カメラの映像から自

動的に人物を検出し(図1)提示することで、警備員の手間を軽減でき警備コストの削減が期待できます。あるいは、映像中の人物を認識して追跡することで、その人物の移動経路が把握でき、セキュリティだけでなくマーケティングなどにも役立てることができます。また、人物だけでなく、自然や人工物なども認識する(図2)ことで、パソコンに保存した個人撮影の写真を自動的に分類して管理できるようになるなどの応用も考えられます。

東芝は、このような応用の実現を目指して、画像認識技術の開発に取り組んでいます。

画像認識の仕組み

同じ“もの”であっても、撮影の方向や“もの”自身の変形などの要因によ



図1. 監視カメラの映像への応用例 — 監視カメラの映像中の人物を自動的に検出することで、セキュリティやマーケティングへの応用が期待できます。

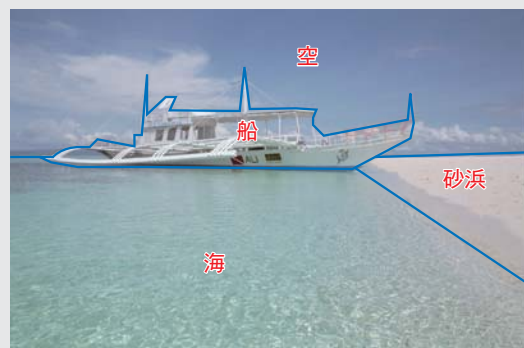


図2. 個人で撮影した写真への応用例 — 写真中の“もの”を認識することで、個人で撮影したたくさんの画像を自動的に分類して管理するなどの応用が期待できます。

て見た目は大きく変動します。したがって、そのまま画像の画素値を観察しても、認識対象の“もの”なのかどうかを判別することは容易ではありません。

そこで、画像認識技術では、“もの”の見た目の変動の影響を軽減するために、“もの”の性質を表す模様(テクスチャ)、形状、色などの特徴の度合いを表す量(特徴量)を画像から抽出して利用します。まず、あらかじめ収集しておいた認識対象の“もの”の画像から特徴量を抽出し、統計的手法を用いて自動的に“もの”を表すモデルを生成して保存します。

画像に“もの”が映っているかどうかを認識するときは、“もの”のモデルを生成したときと同様に画像から特徴量を抽出します。そして、抽出した特徴量と、あらかじめ生成して保存してあるモデル

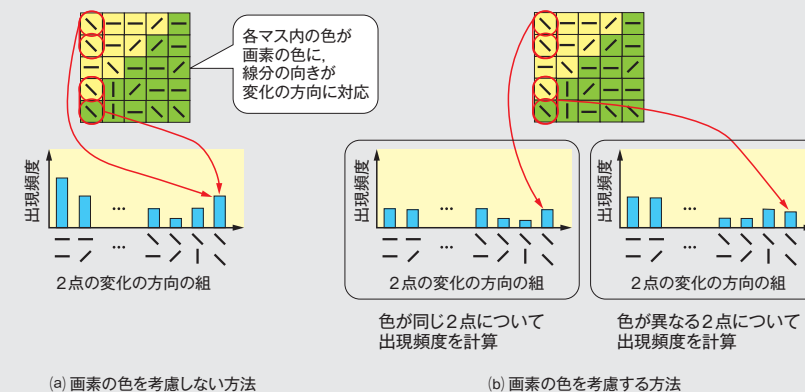


図3. カラー特徴量の一つであるcolor-CoHOGの概念 — 2点の変化の方向を1組とした出現頻度を求める際に2点の色の関係を考慮することで、より詳細に“もの”を表現できるようになりました。

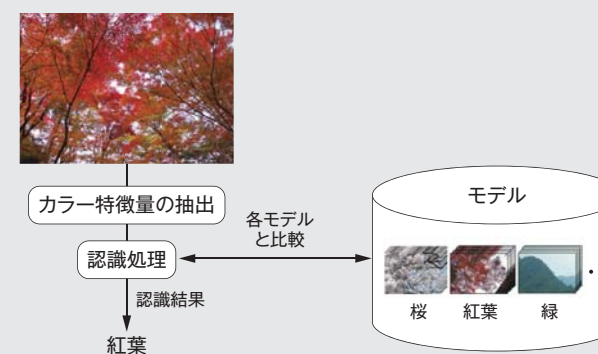


図4. カラー特徴量を用いた画像認識の例 — 提示された写真を、あらかじめ生成してある桜、紅葉、緑などのモデルと比較することで、どんな光景を撮影した写真であるかを認識します。

を比較することで、画像中に“もの”が映っているかどうかを判別します。複数種類のモデルが保存されている場合は、それぞれのモデルと比較して、いちばんもっともらしいモデルが表す“もの”であると判別します。

この画像認識の仕組みでは、どのような特徴量を画像から抽出して利用するかによって、画像認識の性能が大きく左右されることになります。例えば、“人物”に対しては縦長の柱状の形状を特徴量として利用することは有効ですが、“空”に対して同じ特徴量を利用することは有効ではありません。したがって、“空”に対して別途、有効な特徴量を用意する必要がありますが、“もの”の種類は無数にあるため、認識に有効な特徴量を“もの”ごとに考えて作り込むことは、非常に手間が掛かる作業にな

り、現実的ではありません。

カラー画像の特徴量

カラー画像は小さな点(画素)の集合で、各画素は赤(R)、緑(G)、青(B)の各色の強さの度合いを表す数値を持っています。したがって、画素の値が画像中で変化する方向の出現頻度を調べることで、テクスチャや形状の局所的な特徴量を抽出できます。しかし、1点の変化の方向の出現頻度では、異なる“もの”でも同じ出現頻度となる可能性が高くなり、その結果、正しく“もの”を認識できないことが多くなります。そこで、2点の変化の方向を1組として出現頻度を求めることで、表現できるテクスチャや形状の種類を増やします(図3(a))。更に、組を成す2点と同じ色かどうかで場合分けを行い、それぞれの場合につ

いて出現頻度を求めることで、より詳細にテクスチャや形状を表現できるようになります(図3(b))。また、2点の組合せを複数用意することや異なる種類の変化を調べることで、より詳しく“もの”を表現できるようになります。

この考え方に基づいて、当社はカラー画像のための新しい三つの特徴量、color-CoHOG (Co-occurrence Histograms of Oriented Gradients)、CoHED (Co-occurrence Histograms of Pairs of Edge Orientations and Color Differences)、及びCoHD (Co-occurrence Histograms of Color Differences)を開発しました。これらの特徴量は特定の認識対象を想定していないため、あらゆる“もの”に対して適用できます。これによって、“もの”ごとに特徴量を作り込むことなく、様々な認識対象に共通して使用できる特徴量を抽出することができ、従来よりも簡単に画像認識技術を開発できるようになりました(図4)。

今後の展望

画像認識技術は、画像から抽出される特徴量と“もの”という意味的にレベルが大きく異なる二つのものを結び付ける技術です。

今後、認識したい“もの”の種類が増えてくるにしたがって、特徴量と“もの”とを直接的に結び付けることは難しくなり、それらを仲介するような中間表現が重要になってくると考えられます。今後は、この中間表現についても開発を進めていきます。

伊藤 聡

研究開発センター
マルチメディアラボラトリー