

■杉本晃宏 コンテンツ科学研究系 教授

【タイトル】
コンピューター・ビジョンを離散幾何学の立場で再構築する

【本文】

コンピューター・ビジョンの手法に、数理工学の立場から厳密な検討を加え、精度などの議論がきちんとできるようにしたいと思っています。

コンピューター・ビジョンとは、デジタルカメラで撮影して得た 2 次元画像から、そこに写っている対象が何かを判断する技術とっていいでしょう。この技術は大きく進歩して一般のデジタルカメラにも応用されるほど身近になりました。例えば、ファインダー内にある対象を背景と区別してそこにピントを合わせる、人の顔を認識してその部分がきれいに写るよう露出を調整するといった技術がコンピューター・ビジョンの成果です。私は数理の立場から、対象が何であるかを認識するための手法を研究してきました。

モデルを用いた認識と画像の再構築

3 次元の形状情報が 2 次元画像として撮像される過程を数理的にとらえ、カメラで撮った画像から対象を認識するためのモデルを作っておく方法があります。いくつかの形状ごとにモデルを作っておいて 2 次元画像がどれに合うかを判定すれば、照明や視点が変わっても、何が見えているかが分かるようになります。

対象が点集合で表現されているとします。そして、空間に対象を表すいくつかの点がある視点から見えていて、別の視点からやはり同じようにいくつかの点が見えるとします。どの点とどの点に対応するかが分かれば、カメラで撮った 2 枚の画像を基に、まったく別の新しい視点から見たときどのように見えるかを正確に予測することができます。

こうしたコンピューター・ビジョンの理論を深めながら現実の問題に適用するうちに、1 つの問題に気が付きました。2 次元画像から 3 次元の情報を復元しようとした場合に、復元した長さや角度に対して精度保証をするためのきちんとした理論がないのです。

離散幾何学を使って精度保証をする

デジタルカメラで撮るということは、連続している 3 次元の情報を離散化（デジタル化）した 2 次元情報に変換することです。例えば、現実の平面上に直線を描いたものをカメラで撮影すると、コンピューター上では基盤の目を塗ったように変換されます。どれほど解像度を上げようと、デジタル化された情報である以上は何らかの誤差（離散化誤差）が伴います。このデジタル化された 2 次元画像の情報から対象の 3 次元空間での位置や向きを推定するのですが、推定された値にどれくらいの誤差が含まれるのかということが分からなかった。そういう突っ込んだ議論を置き去りにして実験的に、つまりカメラの前に実際に何かを置いてみて、精度を調べるということが多かったのです。

ここで、離散幾何という数学の助けを借りると、かっちりとした精度を求めることができます。今考えているのは、離散幾何の考えに基づいてコンピューター・ビジョンで使われてきた幾何学（多視点幾何）を新しい「離散多視点幾何」として厳密に再構築することです。

離散多視点幾何ができれば、対象の大きさとカメラの配置、そしてどのくらいの誤差を許容するかを決めれば、どのくらいの性能のカメラが必要となるかが理論的に分かります。解像度の違うカメラを複数組み合わせる場合でも、同じ対象を見ていることが分かればその位置を厳密に求めることもできます。別な言い方をすると、対象物とカメラの性能から得られる精度の限界が決まってしまうのに、それに気づかず計算方法をいじってみたりしている場合に、はっきり「それは無理です」と言ってあげられるようになります。

（取材・構成 齋藤 淳）