

# 先見的情報を用いた画像陰影からの多面体の形状復元手法

## コンピュータビジョンと画像理解の第一歩

文教大学大学院情報学研究科 教授 下平 丕作士†

Hisashi Shimodaira†

あらまし コンピュータビジョンと画像理解とは何か、その難しさについて説明した後、対象を多面体に限定し、画像陰影から多面体の形状を復元するための著者の研究<sup>1)</sup>の概要を紹介する。

キーワード：コンピュータビジョン，画像処理，陰影，形状復元，多面体

### 1. はじめに

ビジョンとは視覚のことである。人間は眼からの情報から3次元の世界(シーン)でなにが起こっているかを知ることができる。カメラを通じて得られた観測データに基づいてコンピュータにこれをやらせたいというのがコンピュータビジョン(CV)であり、観測データが画像である場合が画像理解である。

CVの研究の歴史は古く、1965年頃MITのMinsky教授が大学院生に夏休みの宿題として課題を出したそうである。その頃から、数多くの研究がなされてきたが、40年経ってもいまだに、全然と言ってもいいほど未解決である。人間は、1枚の写真を与えられたとき、その中に何が写っているか、写っているものの形状やおおよその大きさなどを、特別な予備知識なしで理解することができる。コンピュータでは、一般的な状況では、それがほとんどできない。しかし、シーンや対象について限定された場合には、それが可能であり、実際、郵便物の宛先の認識、製品の組み立てロボットなどで実用化され、その有用性が認められている。

本稿では、CVにおいて解決すべき課題について説明した後、対象を多面体に限定し、一般的な状況における画像陰影から多面体の形状を復元するための著者の研究<sup>1)</sup>の概要を紹介する。

### 2. コンピュータビジョンはなぜ難しいか<sup>2)</sup>

画像は3次元の世界を2次元の画像面に射影したもので、奥行き情報つまり3次元情報を完全に失っている。ビジョ

ンは、2次元の画像から3次元の世界の様子を復元しようとするのである。対象を多面体に限っても、ある画像を生じさせる3次元物体は、無数に存在する。CVでは、画像として与えられた情報を用いて、その中の1つを探索し推論しなければならない。これは与えられた条件に最もよく適合する最適解を探索する問題に帰着するが、画像として与えられた情報のみでは、探索範囲が広すぎて解を見つけることが困難である。

人間は、1枚の写真を与えられたとき、その写真がどのようなシーンを撮ったものか(コンテクスト)を経験的知識に基づいて理解することができる。また、3次元物体についての先見の知識を持っており、さらにそれらを様々な方向や環境で見たときにどのように見えるかを経験的知識として持っている。これらの知識を無意識のうちに用いて、1枚の写真を与えられたとき、その中に何が写っているか、写っているものの形状やおおよその大きさなどを認識することができるのである。CVにおいても、コンピュータにこのような知識を与えることが必要になるが、一般的な状況では、このような知識はほとんど無限とっていいほど存在するため、知識をどのように表現し利用するかが難しいのである。

### 3. 画像陰影からの多面体の形状復元

ここでは、画像情報のうちの陰影、すなわち画素値の変化を用いて多面体の大きさを含む形状を復元する手法について説明する。これは、広く画像陰影からの形状復元(Shape-From-Shading:SFS)と言われる分野の一部である。多面体の画像を解析しエッジを検出することにより、多面体の稜線を表す線画が得られる。例えば、図1の画像について図2のような線画が得られる。それぞれの直線は多面体を構成する平面の交線である。SFSは、画像中の各面の明るさと交線のデータにマッチするような3次元空間における多

2007年1月10日受付

〒253-8550 神奈川県茅ヶ崎市行谷1100

shimodaira-h@hi-ho.ne.jp

†Graduate School of Information and Communication,  
Bunkyo University

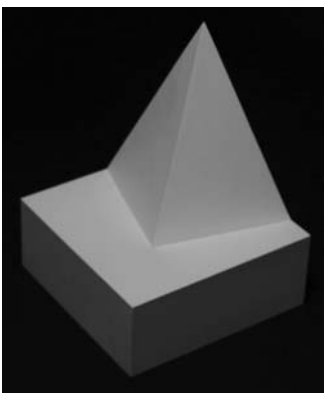


図 1 多面体の画像

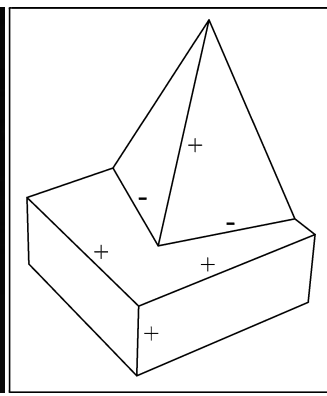


図 2 画像から抽出した線画

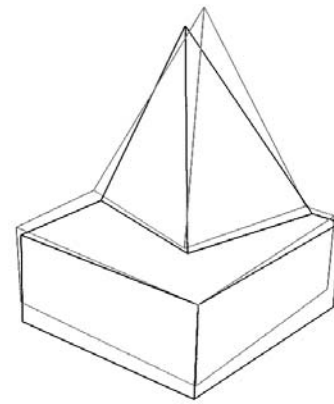


図 3 計算結果

面体の形状を逆算する問題である．実世界では多面体で近似できる物体が多く，画像理解における重要な課題である．

稜線は凸 (+)，凹 (-)，または面と背景との境界 ( , ) の 4 種類に分類され，ラベル付けされる．多面体の線画は，これらの 4 種類のラベルで全体として矛盾のないようにラベル付けができることが明らかにされた．このような矛盾のないラベル付けは通常複数存在し，ラベル付けだけでは，ごく限られた場合にしか元の形状を特定できない．杉原厚吉 (1984) は上記のラベル付けに基づいて線画がみたくべき代数式を導き，これを用いて多面体の形状を計算により求める手法を提案した．この手法は，数学的には厳密であるが，画像中のノイズから生ずる誤差により解が大きく影響される．また，Shimshoni ら (1997) は，上記のラベル付けに基づいて，線画がみたくべき代数式にノイズから生ずる誤差を未知数として含める手法を提案した．この手法では，探索する未知数が著しく多くなるという難点がある．

著者の手法は，画像情報と先見の情報からなる目的関数を大域的最適化手法により探索し，目的関数が最小になる解を求める．多面体を構成する平面の誤差を考慮した未知数の設定方法を工夫し，先見の情報として稜線のラベル付けと平面の水平性・垂直性を利用して探索範囲を狭める手法に特徴がある．1 枚の画像データのみでは 3 次元物体の大きさを逆算できないので，カメラから物体までのおおよその距離と撮影時のカメラの焦点距離をデータとして与える必要がある．以下に，その特徴について説明する．

画像から抽出された線画の頂点の座標値は誤差を必ず含んでいる．3 次元空間における平面は 3 点によって決定される．したがって，多面体を構成する平面の頂点が 4 個以上あり，座標値に誤差が含まれている場合，全ての頂点が同一平面に存在することはできない．提案手法では，復元する面をそれぞればらばらな面とし，各面が少なくとも 1 自由度 (未知数) を有するように順次組み立て，面間のギャップを非線形最適化により最小化する．

線画のラベル付けについては，矛盾のないラベル付けを計算するアルゴリズムが開発されており，実世界に合うような解を選ぶことができる．また，実世界における多面体

は，例えば，家具，建物など多くのものは水平面上に存在し，多くの場合垂直面を持っている．したがって，シーンのコンテキストに合うようにこれらの情報を利用できる．

このような複雑な非線形最適化問題は，局所的な最小値が無数に存在し，真の最小値を求めることは難しい．ここでは，DIRECT (Dividing rectangle) という，大域的探索能力の優れたアルゴリズムを用いている．

例として，提案手法により，図 1 の画像から抽出した図 2 の線画に基づいて，下部の直方体の各面の水平性と垂直性を用いて得られた計算結果を，図 3 に示す．薄い線が実物であり，濃い線が復元した形状である．

#### 4. む す び

提案手法は，SFS の第 1 歩であり，実用的には未完成である．次のような今後の課題がある．実世界の画像中には，必ず他の物体の影が生ずるので，影を考慮した線画の抽出を行う．また，物体の面の明るさは，光源からの光の反射光だけでなく，各面の反射光が相互に影響しあった結果であるため，相互反射の影響を考慮する．

#### 〔文 献〕

- 1) Hisashi Shimodaira: A Shape-from-Shading Method of Polyhedral Objects Using Prior Information, IEEE Transactions on Pattern Analysis and Machine Intelligence, 28-4(2006), 612-624.
- 2) 金出武雄: ロボット視覚, 情報処理, 44-11(2003), 1130-1137.



しもだいら ひ さ し  
下平 丕作士 1941 年生. 1971 年 3 月 東京都立大学大学院修士課程終了. NTT などを経て, 1996 年 4 月 文教大学情報学部教授に就任. 工学博士. コンピュータビジョン, 画像処理, 人工知能, 計算知能などの研究を行っている. 文教大学大学院情報学研究科では「アルゴリズム特論」, 「画像情報処理演習」を, 情報学部では「画像処理」, 「マルチメディア技術」, 「メディア処理プログラミング」, 「コンピュータ推論」を担当.