

問題提起

3D ユークリッドか 2D 非ユークリッドか

—画像理解の方法論—

正員 金谷 健一[†]

3D Euclidean or 2D Non-Euclidean ?

—Methodology of Image Understanding—

Ken-ichi KANATANI[†], Member

[†]群馬大学工学部情報工学科, 桐生市

Department of Computer Science, Gunma University, Kiryu-shi, 376
Japan

あらまし コンピュータビジョンにおける3D復元問題の数学理論のありかたについて問題提起を行う。

1. まえがき

コンピュータビジョンにおいて、物体の形状や運動を投影された画像から復元する研究が数多くあるが、理論の構成法を大別すると、3次元空間での幾何学的関係からベクトル解析により求める“3D ユークリッド幾何学の立場”と、投影された画像面をテンソル解析、微分幾何学などによって解析する“2D 非ユークリッド幾何学の立場”がある。いろいろの例をあげてこれを論じる。

2. 3D 復元問題の構成

空間中に物体があり、それを写したカメラから入力した画像がえられたとする。そして、カメラの位置、向き、特性は既知であるとする。通常は単純な透視変換が仮定される。この時、投影された画像から物体の特性(形状、位置、向き、運動など)を推定する問題が3D 復元問題である。

問題を定量的に扱うために、物体に特定のモデルを仮定するのが普通である。モデルとしては、直線、平面、2次曲面、球、円柱、およびそれらを結合したもの(例えば多面体など)が通常はよく用いられる。このようなモデルは少数のパラメータで指定することができる。また、位置は位置ベクトルで、向きは方向ベクトルで指定され、運動は並進ベクトルと回転軸ベクトルとその回りの角速度で指定できる。従って、物体の3次元特性は幾つかのパラメータを指定することで決定される。そのようなパラメータを“物体パラメータ”と呼ぼう。

一方、画像に対しても特徴点抽出、領域分割、境界線検出などのいろいろな画像処理を行い、いろいろなデータとして特徴づけることができる。このようにして、画像上で計測されるデータを“画像の特徴量”と呼

ぼう。

結局、3D 復元問題を解くには、物体パラメータと画像の特徴量との間の関係式を導き、これを解けばよい。そのような式を“復元方程式”と呼ぼう。これが複雑な非線形連立方程式になれば、以後の処理が困難になる。式の形は簡単なほどよいし、解析的な解が得られるのが最も望ましい。どのようにして簡単な式を導くかが問題である。

3. 3D ユークリッド幾何学の立場

一つの考え方はすべての関係を3次元空間の中で記述することである。視点 O を3次元空間の一点とみなし、透視する画像面が3次元空間中に置かれていると考える。画像上の各点は3次元空間中の点と同一視する。そして、与えられた画像を3次元空間に“逆投影”する。例えば画像上の一点 P に対応する空間中の点の位置は、視点 O を始点とし、画像上の点 P へ向かうベクトルを \overrightarrow{OP} とすると $r\overrightarrow{OP}$ と表せる。ここに r は未知変数である。同様に、画像上の直線 l に対応する空間中の直線は、視点 O と直線 l の定義する平面上にある。このように、画像上のデータ(画像の特徴量)をもとにして、未知変数(物体パラメータ)を導入することによって、画像上に描かれた図形に対応する3次元物体を3次元空間中に考える。そして、その物体がもっていないなければならない“拘束条件”を満足するように物体パラメータの値を決定する。すなわち、拘束条件を物体パラメータによって書き表したものが復元方程式となる。

拘束条件は考えているモデルによって与えられる。ある直線が別の直線と交わる、ある線分がある特定の長さを持っている、ある辺が別の辺とある角度をなす、あるいは直交するなどの条件である。また、物体の表面のある部分が平面であるというモデルであれば、平面性の条件が課されるし、2次曲面のモデルであれば2次方程式が課される。物体が剛体であって運動している場合には、剛体運動の条件が課される。

これまでの3D 復元問題に関するコンピュータビジョンの論文では、ほとんどといっていいほどこの立場から書かれている。まず得られた画像に奥行きを考えて、つじつまの合うような解釈をさがすというのは最も率直で素朴な発想であろう。用いる数字も3次元ベクトル解析のみで十分であり、直交性、平行性などの3次元ユークリッド幾何学の概念だけですむ。これは、ビジョンの研究が人間の視覚認識を研究する知覚心理学から始まったという歴史的事情に由来するのかもしれ

ない。しかし、それをそのままコンピュータビジョンにもってくと、得られる式は形が大変複雑になり、解は解析的な形ではなく、単に解法のアロリズムとしてしか与えられないことが多い。

復元方程式を複雑にする最大の原因は物体パラメータのとりかたである。画像面から出発するため、どうしても画像面に相対的な量(例えば画像面からの距離)をとることになり、その物体固有の不変な意味をもつものをとりにくいからである。そして方程式は普通3次元ベクトル関係式で表される。どの成分も同等の意味があるときにはベクトル関係式は非常に簡単な形になることが多いが、3D復元問題では視点を通り画像面に垂直な透視の軸方向が特別な役割を果たしている。これを z 軸にとるなら、 x , y , z 成分は同等な意味をもっていない。

4. 2D非ユークリッド幾何学の立場

対立する考え方はまず最初に物体パラメータをもつ物体のモデルを考えておいて、それがどのように画像面に投影されるかを考えることである。透視変換は簡単な式で書き表せるから、これは容易である。そして、物体パラメータを変化させたとき、投影像がどう変化するかを調べる。ここで肝心なことは、あくまで2次元画像としての変化を考えることである。すなわち、元の3次元空間は忘れてしまうのである。

このようにすると、単にパラメータによって変化する2次元画像を考えていることになり、画像面上に x, y 座標系をとれば、すべての量が x, y を用いて記述できる。濃淡画像であれば2変数関数 $f(x, y)$ として扱える。次に、そのような2次元図形を特徴づける画像の特徴量を理論的に計算する。これは物体パラメータの関数である。一方、実際に得られた画像に対してその特徴量を計測する。両者を等しいと置いたものが復元方程式になる。

3次元空間中の直線はやはり直線として画像上に投影されるが、平行な、あるいは直交した直線は画像上ではもはや(2次元ユークリッド空間としては)平行ではなく、あるいは直交しない。また、投影された線分の(2次元ユークリッド的)長さはもとの3次元空間での長さを保存していない。しかし、画像面上に適当な計量を導入することによって、すなわち、画像面を“非ユークリッド空間”とみなすことによって、元の意味での長さ、平行性、直交性を画像上で考えることができる。

このように2次元の世界だけで考えると復元方程式

の形が簡単になることが期待される。すべての量が x, y 座標によって表され、もし透視の軸と交わる位置を座標の原点に選べば、 x 軸、 y 軸は完全に同等であるから、 x 座標、 y 座標について対称な関係式が得られるはずである。更に、座標系を原点の回りに回転しても全く同等な座標系が得られるから、すべての関係式は座標回転に不変な意味をもつような形に表せるはずである。このことから、座標回転に不変な意味をもつような特徴量を計測した場合のみ3D復元が可能であると結論することができる。

5. 3Dユークリッド幾何学の立場との比較

問題を2次元に還元すると、微分幾何学、テンソル解析などのさまざまな数学的考察を行うことができる。特に物体が平面であれば、物体面から画像面への変換は2次元射影変換であり、画像面に対して射影幾何学を適用することができる。また、座標系の回転を考えることは、画像面上の2次元回転群を考えることにはかならないが、変換群としては座標回転によるもの以外も考えられる。一例はカメラをレンズ中心の回りに回転して得られる画像面の変換である^{(4),(6)}。また、画像から出発するのと違って、物体の3次元モデルから出発するため、物体パラメータとしてその物体固有の不変な意味をもつものがとれる。その幾何学的性質に着目して、復元方程式の解析解が得られることもしばしばある。

3Dユークリッド幾何学の立場(以下、略して“3Dの立場”)と異なる最大の点は、3Dの立場では、与えられた画像データから出発するのに対して、2D非ユークリッド幾何学の立場(以下、略して“2Dの立場”)では、物体の3次元モデルから出発するため、画像上で何を計測したらよいか“結論”として得られることである。

また、その2次元図形が備えていなければならない“適合条件”も得られる。もし、与えられた画像がこれを満していなければ、それは“にせの画像”、すなわち3次元物体を投影することによっては得られないもの、と判定できる。

3次元モデルから導いた理想的な画像は当然適合条件を満たしているはずであるが、実際の画像ではノイズのために満たされていないことがありうる。そのような画像から直接に出発する3Dの立場では、未知数を用いても可能な解が構成できなかつたり、できても真の解を含んでいなかつたりする可能性がある。一方、2Dの立場は本質的には画像のマッチングであるので、このようなことは処理しやすい、とくに用いる特徴量

として観測値の総和や平均の意味をもつようなものを選べば、ノイズの影響を抑えることができる。これは工学的には大変重要なことである。

6. 3D 復元問題の例

画像上に複数の線分が描かれているとする。それらの3次元空間での対応する長さやなす角度が既知であるとすると、どのような3D復元が可能であろうか。これに対する典型的な3Dの立場からの考察を尺長・金子⁽¹⁴⁾が行っている。これに対してKanatani⁽¹⁰⁾はカメラ回転による画像の変換を2次元空間の変換として導入する2Dの立場による定式化を与え、いかに考え方が単純になるかを示している。それによると、角度情報に関しては中心投影画像と平行投影画像とは簡単な変換で等価に移り合うので、平行投影のみ考えればよいことがわかる。

Witkin⁽²⁰⁾はテクスチャを構成する線分要素の方向分布が等方向であれば、画像上の線分要素の方向分布のヒストグラムを作ることによって面の傾きとその方向が計算できることを示した。その考え方は、単純には分類できないが、基本的には3Dの立場である。同じ問題をKanatani⁽⁹⁾は完全に2Dの立場から定式化した。そこでは面の傾きという3Dの意味を忘れて、画像のアフィン変換を考え、テンソル解析によるフーリエ解析を行って、画像の特徴量を導いている。更に、ステレオロジー(積分幾何学)を応用した計測法も提案している。

Aloimonos and Swain⁽¹⁾はテクスチャを構成する点要素や線分要素の分布が一様であれば、画像上の場所場所での点要素の数や線分要素の長さを計算して、面の傾きとその方向が求まることを示した。その導き方は3Dと2Dが折衷したものであるが、同じ問題に対してKanatani and Chou⁽¹¹⁾は徹底した2Dの立場から最も一般的な定式化を与えた。まず対象の面の“第一基本形式”を画像面上の xy 座標で表して、それを用いて一般的な画像の特徴量を定義している。

Sugihara and Sugie⁽¹⁸⁾は物体の剛体運動を、平行投影によって得られる画面上のオプティカルフローから復元するアルゴリズムを示した。これは典型的な3Dの立場に基づくものである。これに対してKanatani⁽⁶⁾は2Dの立場より、物体の平面部分のオプティカルフローを画像面上で特徴づけるパラメータ(画像の特徴量)から座標回転不変量(複素数)を構成して、3D復元の解析的な式を与えた。中心投影の場合の解析的な解は、やはり平面部分に着目してLonguet-Higgins⁽¹³⁾や

Subbarao and Waxman⁽¹⁶⁾が導いている。これらは基本的には2Dの立場であるが、Kanatani⁽⁶⁾は徹底した2Dの立場より、やはりフローパラメータから構成した複素不変量を用いて3D復元の解析解を与えた。

物体の有限運動に関しては、すべての情報を画像上の物体像の移動のみから引き出すもの(Tsai and Huang⁽¹⁹⁾)、標準的な位置からの変位を求めるもの(Silberberg, et al.⁽¹⁵⁾)などいろいろ提案されているが、ほとんどが3Dの立場である。これに対してKanatani and Chou⁽¹²⁾は2Dの立場から画像上の特徴量に基づき、画像間の対応点を用いずに3D運動を復元する方法を提案している。これは金谷⁽⁷⁾の無限小運動の場合の有限運動への拡張である。

7. むすび

3Dの立場と2Dの立場の違いは復元方程式を3次元空間での関係式とみなすか2次元画面上の関係式とみなすかということであるから、必ずしも厳密なものではない。また、3Dの立場が便利な問題もある。問題ごとに最も適した定式化を用いるべきであろう。

Horn⁽²⁾は彼の著書のまえがきで次のように述べている。コンピュータビジョンの初期の研究では素朴な方法のみによって多くのめざましい成果が得られたので、単純な数学だけでコンピュータビジョンの研究が進展するだろうという楽観的期待が抱かれた。しかし、やがてその限界が分かってくると、やはり深い数学が必要であるという認識が次第に広まってきた。しかし、その変化には時間がかかる。Hornは同じ著書の中でSugiharaの多面体の線画の解釈の理論(Sugihara⁽¹⁷⁾参照)を例にとり、従来の定式化の枠組みに入らないことと、なじみの薄い数学を用いていることにより、当然に値する注目を受けていないと述べている。

いうまでもなく、従来の定式化と異なるアプローチを試みることや、新しい数学を導入することは科学技術の発展のためには欠くことができない。コンピュータビジョンの今後の発展のためには、解さえ得られればよいという態度ではなく、これまで単純な方法でも解決した問題を含めて、もう一度、数学的枠組を再検討しなければならない時期に達したのではないと思われる。その意味で、ここでの議論が今後のコンピュータビジョンの発展の基礎となることを期待したい。

謝辞 本文の内容の一部は筆者が米国Maryland大学に滞在して行った研究をもとにしている。3D復元問題についての討論に加わって頂いたAzriel Rosenfeld, Larry S. Davis, John Aloimonos, Murali

Subbarao, Tsai-Chia Chou の諸氏に感謝する。また、東京大学工学部計数工学科の甘利俊一教授、杉原厚吉助教からも有益なコメントを頂いた。

文 献

- (1) Aloimonos and M. J. Swain : "Shape from texture", Proc. 9th Int. Joint. Conf. Artif. Intell., Los Angeles, pp. 926-931 (1985-08).
- (2) B. K. P. Horn : "Robot Vision", MIT Press, Cambridge, MA (1986).
- (3) K. Kanatani : "Detection of surface orientation and motion from texture by a stereological technique", Artif. Intell., 23, pp. 213-237 (1984).
- (4) 金谷健一 : "オプティカルフローの変換則と不変分解", 信学技報, PRL 85-9 (1985-06).
- (5) K. Kanatani : "Structure and motion from optical flow under orthographic projection", Comput. Vision Graphics Image Process., 35, pp. 181-199 (1986).
- (6) K. Kanatani : "Structure and motion from optical flow under perspective projection", Comput. Vision Graphics Image Process., in print.
- (7) 金谷健一 : "対応点を用いない物体の運動認識の理論", 情処学論, 27, 3, pp. 373-382 (昭 62-03).
- (8) K. Kanatani : "Camera rotation invariance of image characteristics", Comput. Vision Graphics Image Process., in print.
- (9) 金谷健一 : "3次元復元のための座標回転不変量の構成", 信学論(D), J70-D, 5, pp. 937-945 (昭 62-05).
- (10) K. Kanatani : "Constraints on length and angle", Comput. Vision Graphics Image Process., in print.
- (11) K. Kanatani and T.-C. Chou : "Shape from texture : General principle", 第17回画像工学コンファレンス論文集, pp. 129-132 (昭61-12).
- (12) K. Kanatani and T.-C. Chou : "Tracing finite motions without correspondence", Proc. Int. Workshop Industrial Appl. Machine Vision Machine Intell., Tokyo, pp. 118-123 (1987-02).
- (13) H. C. Longuet-Higgins : "The visual ambiguity of a moving plane, Proc. Roy. Soc. Lond. B-223, pp. 165-175 (1984).
- (14) 尺長, 金子 : "透視角度変換による3次元情報抽出", 信学技報, PRU86-83 (1987-01).
- (15) T. M. Silberberg, D. Harwood and L. S. Davis : "Three dimensional object recognition using oriented model points", Techniques for 3-D Machine Perception (ed. A. Rosenfeld), Elsevier, Amsterdam, pp. 271-319 (1986).
- (16) M. Subbarao and A. M. Waxman : "Closed form solution to image flow equations for planar surface in motion", Comput. Vision Graphics Image Process., 36, pp. 208-228 (1986).
- (17) K. Sugihara : "Machine Interpretation of Line Drawings", MIT Press, Cambridge, MA (1986).
- (18) K. Sugihara and N. Sugie : "Recovery of rigid structure from orthographically projected optical flow", Comput. Vision Graphics Image Process., 27, pp. 309-320 (1984).
- (19) R. Y. Tsai and T. S. Huang : "Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces", IEEE Trans. Pattern Anal. Machine Intell., PAMI-6, 1, pp. 13-27 (1984).
- (20) A. P. Witkin : "Recovering surface shape and orientation from texture", Artif. Intell., 17, pp. 17-45 (1981).
(昭和 61 年 11 月 19 日受付)