

の長いレンズで撮影すると遠くの背景もかなり大きく写るが、焦点距離の短いレンズで撮影すると遠くのものほど急速に小さく縮んで写るので、このことを利用すればよい。実はこれが今日の先端技術であるロボット工学と重要な関係がある。これを説明するために、まずカメラの仕組みから始めよう。

よく知られたように、レンズによる結像の原理は図 2 のようになっている。実際のカメラではもっと複雑な仕組みになっていることが多いが、カメラの内部は問題にしないでそれを単に使うって撮影する立場からは、すべてのカメラがこのようになっていると考えてよい。レンズの回転対称軸を光軸と呼ぶ。今、ある物体にピントがあっていて、レンズの中心 O からの距離が a のフィルム面に像を結んでいるとする。 O から物体面までの距離を b とすると、次のレンズ方程式が成り立つ。

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f} \quad (1)$$

ここで f がレンズの焦点距離である。後で述べるが、ロボットの視覚認識として重要なのはレンズの中心からフィルム面までの距離 a である。しかし、もし物体までの距離 b がカメラの大きさに比べて十分大きければ近似的に $1/a \approx 1/f$ であり、したがって $a \approx f$ とみ

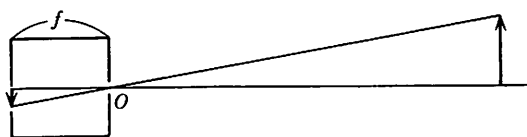
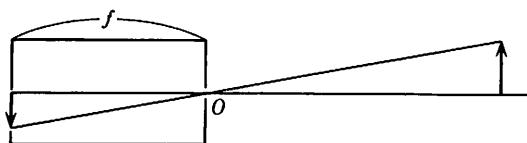


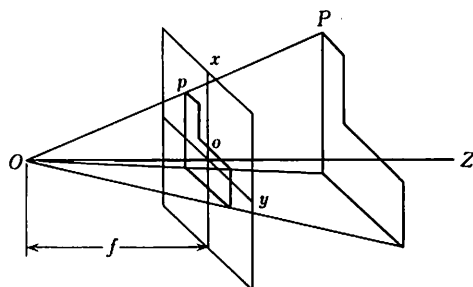
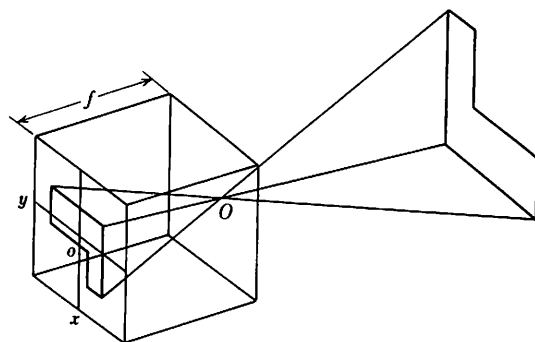
图 3

なしてよい。通常のロボットの作業環境ではこの条件が成立しているので、以下ではレンズの中心からフィルム面までの距離を単に「焦点距離」と呼んで、記号 f で表す。この用語は多少まぎらわしいが、新語を導入すると煩雑なので最近ではよく用いられている。

以上のことは、ピンホールカメラを考えるときわかりやすい。ピンホールカメラにはレンズがないが、ピンホールからフィルム面までの距離を“焦点距離”と定義する。また、ピンホールを通り、フィルム面に垂直に交わる直線を“光軸”と定義する。図3(a)のような焦点距離の長いものがいわば“望遠ピンホールカメラ”であり、図3(b)のような焦点距離の短いものが“広角ピンホールカメラ”である。

3 透視変換によって3次元がわかる

しかし、図2のモデルにしても図3のピンホールカメラにしても、像は原点 O の後方に上下左右が反転して写るので考えにくい。これを考えやすくするには、像を原点 O の周りに 180° だけ回転して、像があたかもレンズまたはピンホールの前方にあると思えばよい。



☒ 4

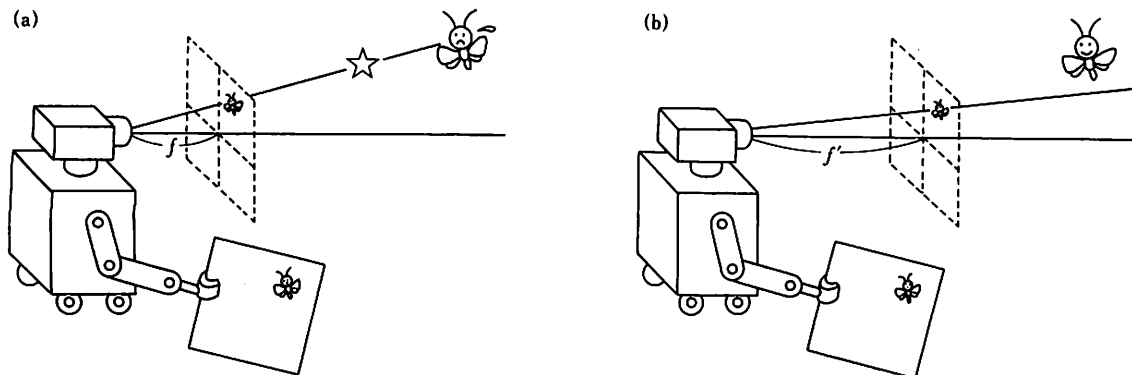


図5 (a)正しい3次元解釈。(b)誤った3次元解釈。

(図4(a))。このことは、原点の前方に焦点距離 f と等しい距離をもつ仮想的な面を考えて、空間中の点がその点と原点 O を通る直線との交点に像を結ぶと考えることに相当する(図4(b))。そのような仮想的な面を以後、画像面と呼ぶ。このように考えても等価であることは図4から明らかであろう。

図4(b)のように空間中の物体点を画像面上に写像することを透視変換と呼ぶ。これはちょうど人の目の前の一定の距離に透明なガラス板を置いて、それを透して見える景色をそのままガラス上に描くことに相当している。このようにすれば真に迫った絵が描けることはルネサンス時代の画家達に知られていたことである。このことから、原点 O と物体点を結ぶ直線を視線と呼び、原点 O を視点と呼ぶ。画像面の視点との距離 f はやはり焦点距離と呼ぶ。

ロボットが画像から3次元空間を理解できるのはこの原理に基づいている。ロボットに搭載されたカメラから画像が得られたとすると、そのカメラのレンズの中心 O を視点、その光軸を Z 軸とする透視変換を想定すればよい。すなわち、その画像を光軸に垂直に、レンズの中心 O から焦点距離 f だけ離れた位置に仮想的に置いたと考えればよい。すると、画像上の点 p に対応する空間中の物体点 P はその視点 O から点 p を通って伸ばした視線にあることになる。したがって、ロボットがその点に到達したければその視線方向に進行すればよいし、ロボットアームでその物体をつかみたいなら、その方向にアームを移動すればよい。

このことから、ロボットの正確な動作のためには焦点距離 f が正確にわかっている必要がないことが理解できる(図5(a))。もし誤った焦点距離 f' を用

いると、それから判断した3次元空間は誤ったものになり、ロボットは正しい動作ができない(図5(b))。

4 デジタル画像のキャリブレーションとは？

しかし、実際問題としては焦点距離を正確に知ることにはなかなか難しい。最大の理由はカメラとして、フィルムを用いる普通のカメラではなく電気信号として画像を取り出すビデオカメラを用いることである。その信号はコンピュータに入力され、ディスプレイに表示されて画像として初めて観測される。ディスプレイに表示される画像は普通は格子上に配置された多数の点からなっている(例えば 512×512 個)。そのような各点を画素と呼ぶ。コンピュータの内部では各画素それぞれに対するメモリがあって、モノクロ画像ではその点での濃淡値、カラー画像では赤(R)、緑(G)、青(B)三原色の各成分の強度値が数値データとして記憶されている。そのようなデータの構造をもつ電子的な画像のことをデジタル画像と呼ぶ。

したがって、たとえレンズの光学的な焦点距離がわからなくても、それとデジタル画像の大きさとの比がわからないと無意味である。このことはフィルムを用いるカメラでも同じである。もし焦点距離 50 mm のレンズで撮影したフィルムをそのままのサイズで焼き付けた写真があれば、その写真を視点から 50 mm 離れた位置に置いて視線を計算すれば3次的に正しい方向が得られる。しかし、もしその写真が何倍かに引き伸ばされたものなら、焦点距離も同じ倍率で拡大しないと正しい3次元方向が計算できない。引き伸ばしの倍率がわからなければレンズの光学的な焦点距離がわかっ

[illegible]

でも意味がない。すなわち、焦点距離は物理的な長さとしてではなく、画像の大きさをもとにして測らなくてはならない。

すでに述べたように、ディジタル画像は画素の格子状の配列であるから、ディジタル画像上での距離を測定するのに、隣り合う画素の間隔を1とする単位を用いるのが合理的である。この単位もまた画素と呼んでも混乱は生じないであろう(長さ20画素の線分、等)。したがって、問題は与えられたディジタル画像から焦点距離を画素の単位で推定することである(図6)。一般に、カメラや撮影系の特性をあらかじめ正しく測定しておくことをキャリブレーションという。問題は焦点距離のキャリブレーションである。

画像がデジタル画像であることに加えて、キャリブレーションが必要となるもう一つの理由に、ズームレンズの使用がある。家庭用ビデオカメラでも最近ほとんどズームレンズが用いられている。ズームレンズとは焦点距離を自由に変えることのできるレンズである。ロボットがズームレンズで撮影した画像から3次元方向を計算しようとする、その画像を撮影した状態での焦点距離をあらかじめ調べておかなければならない。このように、焦点距離のキャリブレーションは画像によるロボットの制御において欠かすことができない。

ところで、気がついた人もいるかもしれないが、画像を透視変換のモデルに当てはめるためには焦点距離だけでなく、画像の中心、すなわちカメラの光軸と交わる点の位置も知る必要がある。通常のカメラではフレーム中心(長方形画像の対角線の交点)が光軸に対応

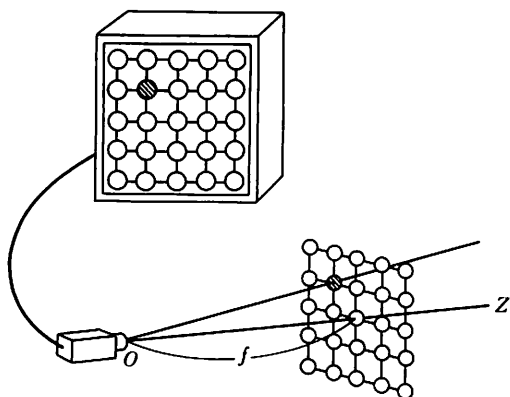


図6 デジタル画像の焦点距離のキャリブレーション
(5×5画素の例).

しているが、場合によっては少しずれていることもある。したがって、画像の中心を定めることもカメラキャリブレーションの重要なテーマである。しかし、これを定めるには焦点距離を定める方法とはまったく違う原理を用いるので、ここでは触れない。以下では、フレームの中心が光軸に対応していると考える。

5 空間中の計測で焦点距離がわかるが…

3 節に示した透視変換の幾何学的関係から、物体の大きさとその位置がわかれば焦点距離が定まる。例えば大きさ h cm のテストパタンを画像面に平行に、視点 O から距離 l cm のところに置いたとき、その投影像の大きさが r (画素) であったとする。焦点距離を f (画素) とすると図 7 (a) の関係から $r/f = h/l$ であり、したがって f が

$$f = \frac{l}{h} r \quad (2)$$

と求まる。距離 l と大きさ n は比さえ計算できればよいので、同じ単位を使う限りどんな単位で測ってもよい。

しかし、この方法は実際的ではない。困るのは視点 O からテストパタンまでの距離 l を測ることである。すでに述べたように、視点 O はカメラのレンズの中心に当たっているが、完成品のカメラでは外部から見てそれがどこにあるかわからない。また正確にカメラの画像面に平行になるようにテストパタンを置くことも難しい。

距離 l の測定を避ける方法は、テストパタンを平行に保ったまま距離 d だけ後退させることである。その

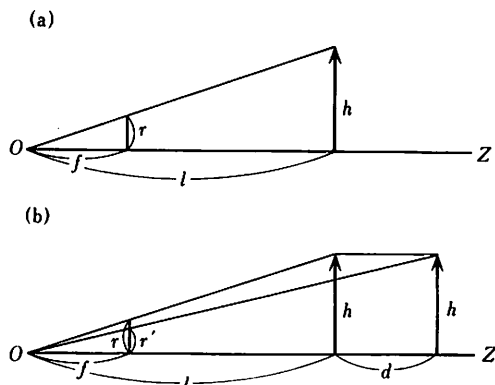


図7 空間中の計測による焦点距離の計算

投影像の大きさが r' (画素) になったとすれば、式(2)と同じ関係から

$$f = \frac{l+d}{h} r' \quad (3)$$

となる。これと式(2)から l を消去すれば、焦点距離 f が次のように求まる。

$$f = \frac{d}{h} / \left(\frac{1}{r'} - \frac{1}{r} \right) \quad (4)$$

したがって、視点から始めのテストパタンまでの距離 l を知る必要がない。

しかし、それでも正確にカメラの画像面に平行となるようにテストパタンを置かなければならない。また、それを一定距離だけ正しく平行移動するのも難しい。このようなことは理屈ではなく、実際にやってみて初めて痛感することである。整備された実験室で精密な計測装置を用いれば簡単にできるかもしれないが、いつでもどこでも手軽にというわけにはいかない。できればロボットの作業する現場で、空間中の計測を含まず、単に画像の解析だけから焦点距離が計算できることが望ましい。その手掛かりを与えるのが消失点である。

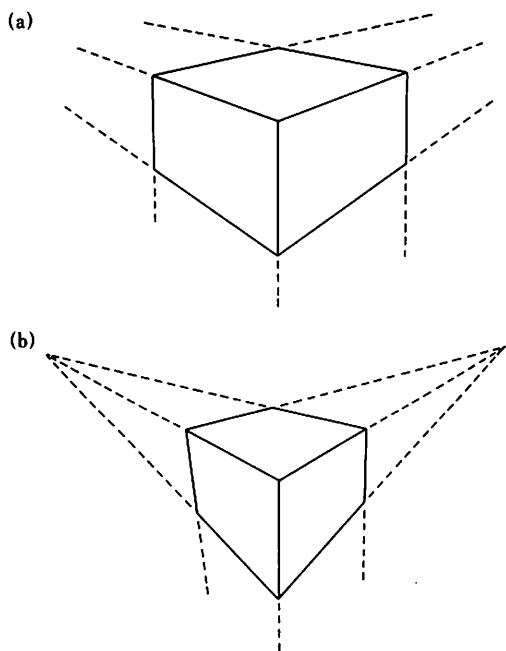


図8 消失点。 (a) 焦点距離が長い場合。 (b) 焦点距離が短い場合。

6 消失点から3次元を推定する

いま空間中で奥行きのある平行線が画像上に写っているとする。その間隔は空間中ではどこでも一定だが、視点から遠ざかるにつれて画像上では間隔が縮んでいく。したがって、その平行線を限りなく延長すればついにはある一点で交わる。この点を消失点と呼ぶ(図8)。

望遠レンズでは奥行きが増しても像の大きさはあまり縮まないのに対して、広角レンズでは奥行きが増すと像の大きさが急速に縮む。ということは、平行線の像は空間中で同じ方向に伸びていても、焦点距離が長いほど画像の中心から離れたところにあり(図8(a)), 焦点距離が短いほど画像の中心に近いところにあることになる(図8(b))。これは昔から透視変換による物体像を製図する技法としてよく知られている。工事中のビルディングの完成予想図などではよく立体感を誇張するために、消失点を画像の中心に近づけた焦点距離の短いものに相当する図が描かれる。

それではこの性質を利用して実際に焦点距離を推定するにはどうしたらよいだろうか。その原理は次の事実にある。それは、消失点を通る視線が空間中の平行線と平行になっているということである。これは図9から明らかであろう。空間中の直線上を一定速度で動く点を考えてみればよい。その点とその直線に沿って視点から限りなく遠方に動くと、その像はその直線とちょうど平行に伸びる視線と画像面との交点に収束することがわかる。

ということは、焦点距離がわかっていると、消失点を見つければ3次元方向が計算できる。視点から焦点距離だけ離れた画像面を考え、視点とその消失点を通る視線を作れば、それがそのまま空間中の平行線の3次元方向になっているからである。このことは画像を解析して3次元環境を理解するコンピュータビジョンという研究における3次元認識の基本原理の一つでも

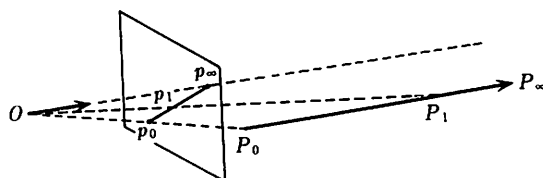


図9 消失点は空間中の3次元方向を示す。

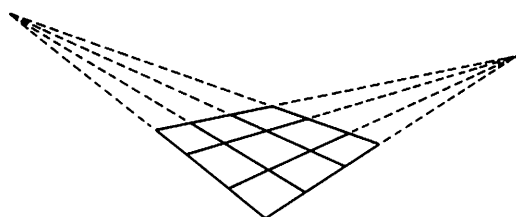
ある。

この事実から焦点距離のキャリブレーションとして次の方法を思いつく。まず、平行線からなるテストパターンを作り、カメラに対してある角度を保つように空間中に配置する。そのカメラ画像を撮り、パタンの直線を延長して消失点の位置を計算する。そして焦点距離を適当に仮定し、パタンの平行線の3次元方向を計算してみる。計算した3次元方向が配置した3次元方向とが一致していれば仮定は正しかったことになるし、一致していなければ一致するように焦点距離の値を修正すればよい。

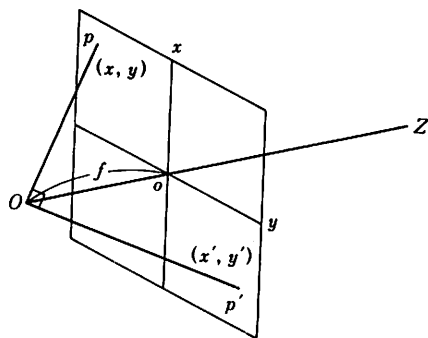
しかし、この方法でもやはりパタンのカメラに対する傾きをあらかじめ正しく計測しておかなければならない。したがって、そのための測定装置が必要となる。整備された実験室でなら可能かもしれないが、いつでもどこでもというわけにはいかない。“工学”の研究ではこのような実的な考慮を欠かすことができない。

7 格子パターンによるキャリブレーション

画像のみから焦点距離を知る便利な方法は互いに直交する平行線からなる格子パターンを用いることである。図10(a)のような画像が得られたら、それぞれの方向の直線を延長して二つの消失点の位置を定める。



(a) 格子パターンによる焦点距離のキャリブレーション



(b) 消失点と焦点距離の関係

図 10

焦点距離を適当に仮定して、それぞれの3次元方向を別々に計算する。計算した3次元方向が互いに直交していれば、仮定した焦点距離が正しかったことになるし、直交していなければ、直交するように焦点距離の値を修正すればよい。

実はこのことは、人間が図8を見て焦点距離を判断するときに暗黙に使っていることである。図8のような多面体の図を見ると、普通の人は頂点で辺が直交していると感じる。3次元形状が想像できるのはこのような角度の情報を暗黙のうちに用いているためと思われる。このような人間の視覚認識のメカニズムについては、コンピュータビジョンの研究者だけでなく、心理学者も深い関心をもっている。

具体的に焦点距離を計算する公式を導いてみよう。図10(b)のように、画像面上に xy 座標系をとり、二つの消失点 p, p' の座標をそれぞれ $(x, y), (x', y')$ とする。視点を O とし、焦点距離を f とすると、ベクトル $\vec{Op}$ は (x, y, f) となり、ベクトル $\vec{Op'}$ は (x', y', f) となる。これらが互いに直交していなければならないから、二つのベクトルの内積が0である。したがって

$$xx' + yy' + f^2 = 0 \quad (5)$$

の関係がある。これから焦点距離を計算する公式が次のように得られる。

$$f = \sqrt{-xx' - yy'} \quad (6)$$

実際にはこの公式はコンピュータで計算する際にいくつかの点で不便である。例えば、消失点が画像の中心からはるか遠方にあるときはその座標値が大きくなり過ぎてオーバーフローを起こす可能性がある。また以下で述べるが、計算の信頼性を理論的に解析するにも都合が悪い。したがって、これを視線方向を変数とする理論的な取り扱いに適した形に書き直す必要がある。しかし、そのためには数学的なことがいろいろからんでくるので、これ以上立ち入らない。

さて、このようにして焦点距離を計算するとして、その精度を左右する要因を考えてみる。まず大切なことは、テストパターンをカメラの画像面に対して大きく傾けなければならないことである。すでに指摘したように、焦点距離が求まる原理は奥行き方向に対してその投影像が縮む比率にあるから、テストパターンを画像面に平行に置いたのでは何の情報も得られない。また、傾け方が足りないと、画像上に現われる透視変換の効果が少なく、画像の解析の誤差の効果が相対的に大き

ORに親しみORを役立てる

オペレーションズ・リサーチ

1993年9月号

特集 事例研究・研究レポート

ネットワーク構造に基づく道路の重要度評価
区間AHPを用いるDEAの改良型領域限定法
CVS向け弁当の需要予測
アラル海の水位——微分方程式による考察
地域における利便性と地価との関係

1993年8月号

特集 交通のDSS

通勤鉄道の列車ダイヤの評価
鉄道波動輸送計画
バスの路線計画
鉄道における車両運用計画
航空機の乗務員スケジューリングシステム

■日本OR学会入会のご案内■

正会員(個人)	年会費	12,000円
	入会金	1,200円
学生会員(個人)	年会費	5,000円
	入会金	600円
賛助会員(法人)	年会費	A種 95,000円
		B種 48,000円
	入会金	不要

オペレーションズ・リサーチ誌および論文誌が配布されます。

■入会手続き■

入会申込みのご希望があれば、会費振込用紙、原簿等の必要書類をお送りいたします。

各号1部 950円(本体922円) 年間購読料 10,800円(税・郵送料含)
お申し込みは下記に

(社)日本オペレーションズ・リサーチ学会

113 東京都文京区弥生2-4-16 学会センタービル

電話 03-3815-3351(代)

~~~~~

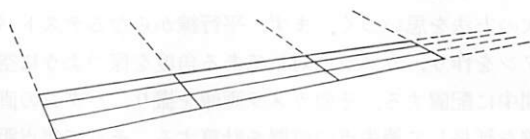


図11 傾け過ぎたパタンの像

くなって、結果の信頼性が低下する。

一方、傾け過ぎると画像処理の問題が生じる。そのためには画像処理の手順について知らなければならない。コンピュータによる処理では入力したデジタル画像に2値比、エッジ検出、細線化といった処理を行なう。そして、格子を作る二組の直線を、それを構成する画素の列としてメモリに記憶する。次に、各直線を構成する画素の列に最もよく当てはまる直線の方程式を計算し、さらに、各方向ごとにその方向の直線の共通の交点として最も適切な点を推定して消失点とする。したがって、格子パターンを傾け過ぎると、傾けた方向の格子線が短く写るので、それに当てはめた直線の信頼性が低下する。同時に、それと直交する格子線の間隔が狭く写るので、共通の交点の推定の信頼性が低下する(図11)。

これらの考察から、計算される焦点距離の信頼性を左右する両方の要因がちょうどバランスする格子パタンの傾きが存在するはずである。その傾きを理論的に求めるには、上に述べた各要因の計算の信頼性に与える影響を定量的に評価しなければならない。このことについては次回に述べる。

### 参考文献

- 1) 出口光一郎:「画像と空間」, 昭晃堂, 1991.
- 2) 井口征士, 佐藤宏介:「三次元画像計測」, 昭晃堂, 1990.
- 3) 金谷健一:「画像理解/3次元認識の数理」, 森北出版, 1990.
- 4) K. Kanatani: Geometric Computation for Machine Vision, Oxford University Press, Oxford, 1993.

(かなたに・けんいち, 群馬大学・工学部)