

論文

3次元運動推定時におけるモデルの位置と姿勢の自動修正

正 員 山本 正信[†]

Direct Estimation of Pose and 3D Motion from an Image Sequence

Masanobu YAMAMOTO[†], *Member*

あらまし 本論文では、物体の3次元形状が3次元モデルとして与えられているとき、物体のおかれている位置と姿勢および運動パラメータを動画像から推定する問題を扱っている。3次元運動と位置および姿勢を表すパラメータを未知数とし、それらを動画像から同時に直接推定できる非線形連立方程式を導いた。この方程式系を制約なしの関数2乗和の極小化問題として解いた。すなわち、物体の3次元モデルをできるだけ対象物体に近づけておき、この状態を初期値として、逐次近似法により極小化問題の解を得た。この手法を使って、3次元モデルに基づく物体の追跡が行えることを実験で確認した。Loweの手法はモデルと物体の像との照合を基本としているのに対し、我々の手法は、モデルと物体像の照合も連続した画像間の照合も必要としない。

キーワード 動画像解析, 直接法, 3次元運動, 3次元モデル, 位置と姿勢, 追跡

1. ま え が き

連続する画像間の対応付けから、物体の3次元運動や3次元構造を知ることができる。時空間こう配法を使えば、画像間の対応付けを行うことなく、直接3次元運動や3次元構造を推定することが可能である。しかしながら、未知数に物体面までの距離が含まれているとき、常に未知数の総数が方程式の総数を上回ることで推定は困難になる。そこでさまざまな制約条件を導入してこの推定問題を解くことが試みられている。対象物体を平面^{(8),(7)}や2次曲面⁽⁶⁾とみなす場合、物体面の3次元情報が既知の場合^{(9),(2)}、物体の動きが並進運動のみ、あるいは回転運動のみ、または回転運動が既知、などの場合⁽²⁾に3次元構造や3次元運動を推定する手法が提案されている。

なかでも、物体の3次元構造が既知の場合、3次元運動を安定に精度良く推定することができる。このことは、物体の3次元幾何モデル(以下、簡単にモデル)が与えられているとき、物体の3次元追跡が可能であることを意味している。モデルに基づく物体の追跡は運動パラメータの推定と、推定値に基づくモデルの移

動を逐次行うことにより達成可能である。しかし、初期時刻でのモデルの位置・姿勢のずれ、あるいは追跡時に何らかのノイズによりモデルのずれが起こった場合、モデルと物体とのずれは累積し、追跡は事実上不可能になる。

本論文では、モデルと物体のずれを修正する手法を提案する。物体面の3次元形状は既知でありモデルにより与えられるものとする。しかし、このモデルは正しい位置と姿勢からずれているため、正確な3次元情報を与えることはできない。このとき、3次元情報の推定誤差を打ち消すように、動画像からモデルの位置と姿勢を修正し同時に運動パラメータを求める手法を提案する。

モデルに基づく物体の3次元追跡の手法としては、画像とモデルを直接照合する手法^{(4),(5)}が知られている。本手法は画像とモデルの照合も連続した画像間の照合も必要としないのが特徴である。

本章では、正確な3次元情報が与えられているときの、3次元運動パラメータの推定法を概説する。3.では、モデルのずれを修正し同時に運動パラメータを推定する手法を示す。4.では推定式がユニークな解をもつための条件を考察する。5.では、物体の3次元追跡実験を行う。

[†] 新潟大学工学部情報工学科, 新潟市
Faculty of Engineering, Niigata University, Niigata-shi, 950-21
Japan

2. 3次元モデルに基づく運動パラメータの直接推定

対象物体が剛体であるとする。その3次元形状と位置・姿勢が3次元モデルにより与えられるものとする。この剛体の動きが小さいとき、その3次元運動パラメータは、直接法を使って動画像から次のように容易に得られる^{(9),(2)}。

シーンの3次元座標系を (x, y, z) とする。物体はその中心点 $\mathbf{p}_c = (x_c, y_c, z_c)^T$ を通る軸回りに、角速度ベクトル $\mathbf{R} = (R_x, R_y, R_z)^T$ 、並進速度ベクトル $\mathbf{T} = (T_x, T_y, T_z)^T$ で剛体運動しているとする。そのとき、物体上の点 $\mathbf{p} = (x, y, z)^T$ の移動ベクトル $\delta = (\delta x, \delta y, \delta z)^T$ は次式で与えられる。

$$\delta = \mathbf{R} \times (\mathbf{p} - \mathbf{p}_c) + \mathbf{T} \quad (1)$$

物体は、画像面上に中心投影されその像により動きが観測される。投影中心を原点、簡単のため投影面を $z=1$ とし物体上の点 \mathbf{p} は投影面座標 (X, Y) に次式で対応する。

$$X = \frac{x}{z}, \quad Y = \frac{y}{z} \quad (2)$$

式(2)の全微分より、画面上の移動ベクトル $\Delta = (\Delta X, \Delta Y)^T$ と物体上の点の移動ベクトルとの関係式が得られる。

$$\Delta = \frac{1}{z} A \delta \quad (3)$$

但し、

$$A = \begin{bmatrix} 1 & 0 & -X \\ 0 & 1 & -Y \end{bmatrix}$$

式(1)を式(3)に代入すれば画面上の移動ベクトルが運動パラメータで表される。

$$\Delta = \frac{1}{z} A \mathbf{T} + \left(B - \frac{1}{z} B_c \right) \mathbf{R} \quad (4)$$

但し、

$$B = \begin{bmatrix} -XY & 1+X^2 & -Y \\ -(1+Y^2) & XY & X \end{bmatrix}$$

$$B_c = \begin{bmatrix} -y_c X & z_c + x_c X & -y_c \\ -(z_c + y_c Y) & x_c Y & x_c \end{bmatrix}$$

一方、画面上の投影像の明るさを $E(X, Y, t)$ とおく。物体が移動しても、移動先での像の明るさが不変だとすれば、次式が成り立つ。

$$E(X + \Delta X, Y + \Delta Y, t + \Delta t) - E(X, Y, t) = 0 \quad (5)$$

さて、式(4)の $(\Delta X, \Delta Y)$ を式(5)に代入すれば、運動パラメータ \mathbf{T}, \mathbf{R} に関する非線形方程式が得られる。

同じ物体に対応する画面上のいくつかの点から導かれるこの非線形方程式を連立させることにより運動パラメータを得ることができる。しかし、この方程式には一般にノイズが含まれているので共通解は存在しない。そこで、次の評価関数、

$$\sum_{x,y} (E(X + \Delta X, Y + \Delta Y, t + \Delta t) - E(X, Y, t))^2 \quad (6)$$

を最小にするパラメータ値を最適解とする。この問題は制約なしの関数2乗和の極小化問題である。適当な初期値から出発すれば、ニュートン法⁽¹⁾により最適解に達することができる。

初期値を $\mathbf{T}^{(0)}, \mathbf{R}^{(0)}$ とする。ここでは \mathbf{T}, \mathbf{R} とも小さいとしているので、 $\mathbf{T}^{(0)} = \mathbf{0}, \mathbf{R}^{(0)} = \mathbf{0}$ とする。最適解は次の反復公式により得られる。ここで、括弧内の変数は反復回数である。

$$\mathbf{T}^{(n+1)} = \mathbf{T}^{(n)} + \Delta \mathbf{T} \quad (7)$$

$$\mathbf{R}^{(n+1)} = \mathbf{R}^{(n)} + \Delta \mathbf{R} \quad (8)$$

増分

$$\Delta \mathbf{T} = (\Delta T_x, \Delta T_y, \Delta T_z)^T$$

$$\Delta \mathbf{R} = (\Delta R_x, \Delta R_y, \Delta R_z)^T$$

は、次の線形連立1次方程式を解くことにより得られる。

$$\mathbf{J}(\mathbf{T}^{(n)}, \mathbf{R}^{(n)}) \begin{bmatrix} \Delta \mathbf{T} \\ \Delta \mathbf{R} \end{bmatrix} = \mathbf{e}^{(n)} \quad (9)$$

ここで、 $\mathbf{J}(\mathbf{T}, \mathbf{R})$ はヤコビ行列である。

$$\mathbf{J}(\mathbf{T}, \mathbf{R}) = \begin{bmatrix} \vdots \\ \mathbf{J}_{x,y}(\mathbf{T}, \mathbf{R}) \\ \vdots \end{bmatrix}$$

$$\mathbf{J}_{x,y}(\mathbf{T}, \mathbf{R}) = \frac{\partial E(X + \Delta X, Y + \Delta Y, t + \Delta t)}{\partial (\mathbf{T}, \mathbf{R})}$$

$$\mathbf{e}^{(n)} = \begin{bmatrix} \vdots \\ e_{x,y}^{(n)} \\ \vdots \end{bmatrix}$$

$$e_{x,y}^{(n)} = E(X, Y, t) - E(X + \Delta X^{(n)}, Y + \Delta Y^{(n)}, t + \Delta t)$$

連立方程式(9)を構成する線形推定式は次のようになる。

$$\frac{1}{z} \mathbf{a} \cdot \Delta \mathbf{T} + \left(\mathbf{b} - \frac{1}{z} \mathbf{b}_c \right) \cdot \Delta \mathbf{R} = e_{x,y}^{(n)} \quad (10)$$

但し、

$$\mathbf{a} = \begin{bmatrix} E_x \\ E_y \\ -XE_x - YE_y \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} -Y(YE_X + YE_Y) - E_Y \\ X(XE_X + YE_Y) + E_X \\ XE_Y - YE_X \end{bmatrix}$$

$$\mathbf{b}_c = \begin{bmatrix} -y_c(XE_X + YE_Y) - z_c E_Y \\ x_c(XE_X + YE_Y) + z_c E_X \\ x_c E_Y - y_c E_X \end{bmatrix}$$

線形連立方程式(9)はその係数行列 $\mathbf{J}(\mathbf{T}, \mathbf{R})$ が正則である限り一意な解が得られる。

3. モデルの位置・姿勢ずれの自動修正

3次元モデルが物体の形状は正しく表しているが、その位置・姿勢が正しくないとき、3次元情報 z は真の値とは異なる。そのとき、推定された運動パラメータは正しくない。本章では、モデルの位置・姿勢ずれを修正し、同時に正しい運動パラメータを求める手法を簡単な例で示す。そして、一般化する。

3.1 モデルのずれの修正原理

与えられている情報は、動画像および物体の3次元モデルのみである。これだけの情報から、果たしてずれを修正し、正しい3次元運動パラメータを知ることができるのであろうか。

2.で導かれた運動パラメータの推定法において、モデルが正しい位置にあり、ノイズがなく、かつ1回の繰返しで収束するという、理想的な場合を考えてみよう。このとき推定式は、式(10)より、

$$\frac{1}{z} \mathbf{a} \cdot \mathbf{T} + \left(\mathbf{b} - \frac{1}{z} \mathbf{b}_c \right) \cdot \mathbf{R} = e_{x,y}^{(1)} \quad (11)$$

となる。ノイズがないと仮定しているから、運動パラメータの数(ここでは6個)よりも多くの推定式を連立させたとしても、連立方程式の最小2乗解はこれらの推定式の共通解となる。すなわち、得られた最小2乗解を推定式(11)に代入したとしても等号は成立する。

一方、運動パラメータが与えられたならば、式(11)を使って、逆に距離 z の推定値 \hat{z} を知ることができる。

$$\hat{z} = \frac{\mathbf{a} \cdot \mathbf{T} - \mathbf{b}_c \cdot \mathbf{R}}{e_{x,y}^{(1)} - \mathbf{b} \cdot \mathbf{R}} \quad (12)$$

式(11)から推定された運動パラメータの値を上式に代入したとき、推定値 \hat{z} は当然真の値 z に等しくなる。

$$\hat{z} - z = 0$$

さて、モデルが正しい位置・姿勢からずれておかれているとしよう。モデルから与えられる距離 z' は真の値 z とは異なる ($z' \neq z$)。この誤った距離 z' を使って、運動パラメータを推定する。このとき、最小2乗解は、連立方程式の共通解であるとは限らない。従って、得

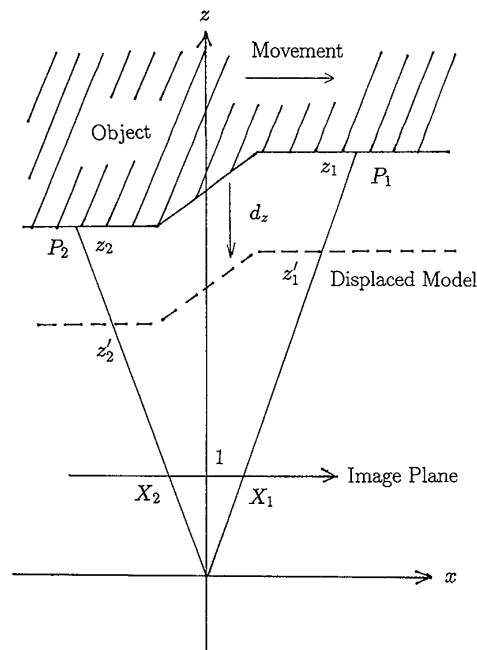


図1 運動している2次元物体と撮像系のモデル
Fig. 1 Dynamic 2D model and measurement geometry.

られた運動パラメータを使って、距離の推定式(12)から得られた距離 \hat{z} は、誤った位置・姿勢におかれたモデルからの距離 z' と一致しない。すなわち、

$$\hat{z} - z' \neq 0$$

である。モデルのずれからくるこの誤差は、我々が動画像とずれたモデルを使って計算できる量であり、ずれがなければ0となる。従って、この誤差を解消するようにモデルの修正を行えばよい。

3.2 簡単な例

最も簡単な場合を考えよう。2次元物体が x - z 平面内を x 軸方向に並進運動しているとする。図1に物体(ハッチングされている)とその撮像系を示す。物体のモデルが図1の点線の位置まで原点方向に d_z ずれているとする。このとき、ずれの方向はわかっているが、ずれの量が未知であるとする。このずれを修正する問題を考える。

前節と同様、まず、モデルが正しい位置にあり、ノイズがなく、1回の繰返しで収束するとする。物体の運動は、 x 軸方向の移動のみであるから、この移動量を推定する方程式は、式(11)を簡単化した。

$$\frac{E_x}{z} T_x + E_t = 0 \quad (13)$$

で与えられる。ここで、 E_t は $-e_{x,y}^{(1)}$ とする。この方程式から移動量 T_x がただちに得られる。得られた移動量は同じ物体上であれば、等しくなければならない。例

えば、図1に示された物体上の2点 P_1 , P_2 に対しそれぞれ次の推定式が与えられる。

$$\frac{E_{X_1}}{z_1} T_x + E_{t_1} = 0 \quad (14)$$

$$\frac{E_{X_2}}{z_2} T_x + E_{t_2} = 0 \quad (15)$$

そのとき、点 P_1 で得られた移動量 $T_x = -z_1 E_{t_1} / E_{X_1}$ と点 P_2 で得られた移動量 $T_x = -z_2 E_{t_2} / E_{X_2}$ は等しくなければならない。そのためには、

$$\frac{z_1}{z_2} = \frac{E_{X_1} E_{t_2}}{E_{X_2} E_{t_1}} \quad (16)$$

が必要である。

逆に、条件式(16)を仮定すれば、推定式(14), (15)を連立させたとき、最小2乗解、

$$T_x = -\frac{\frac{E_{X_1} E_{t_1}}{z_1} + \frac{E_{X_2} E_{t_2}}{z_2}}{\frac{E_{X_1}^2}{z_1^2} + \frac{E_{X_2}^2}{z_2^2}} \quad (17)$$

は、 $-z_1 E_{t_1} / E_{X_1}$ とも $-z_2 E_{t_2} / E_{X_2}$ とも等しくなる。最小2乗解 T_x は二つの方程式(14), (15)の共通解であることがわかる。

一方、移動量 T_x が与えられると、式(13)を z について解くことにより、距離を推定することができる。距離の推定値 \hat{z} は、

$$\hat{z} = -\frac{E_X}{E_t} T_x \quad (18)$$

で与えられる。点 P_1 , P_2 の距離の推定値 \hat{z}_1 , \hat{z}_2 は式(18)に式(17)で与えられる T_x を代入することにより得られる。更に、条件式(16)を使えば、

$$\hat{z}_1 - z_1 = 0 \quad (19)$$

$$\hat{z}_2 - z_2 = 0 \quad (20)$$

であることが証明される。

さて、物体のモデルが図1の点線で示されるように、原点方向に d_z ずれているとする。そのとき、モデルから与えられる物体までの距離は、

$$\begin{aligned} z_1' &= z_1 - d_z \\ z_2' &= z_2 - d_z \end{aligned} \quad (21)$$

となる。このずれた距離情報を使って移動量を推定する。推定式は、

$$\begin{aligned} \frac{E_{X_1}}{z_1'} T_x + E_{t_1} &= 0 \\ \frac{E_{X_2}}{z_2'} T_x + E_{t_2} &= 0 \end{aligned} \quad (22)$$

である。この連立方程式の最小2乗解は

$$T_x = -\frac{\frac{E_{X_1} E_{t_1}}{z_1'} + \frac{E_{X_2} E_{t_2}}{z_2'}}{\frac{E_{X_1}^2}{z_1'^2} + \frac{E_{X_2}^2}{z_2'^2}} \quad (23)$$

となる。今度は、この推定値を式(22)に代入しても等号は成立しない。従って、距離の推定値 \hat{z} もずれたモデルからの距離 z' に等しいとは限らない。

$$\hat{z} - z' \neq 0 \quad (24)$$

モデルのずれの影響が式(24)に現れる。このずれの影響を打ち消すようなモデルの修正量を次のように求める。

モデルをずれた位置から z 軸方向に S_z 移動させるとする。そのとき、点 P_1 , P_2 の距離 z_1'' , z_2'' は、

$$z_1'' = z_1' + S_z \quad (25)$$

$$z_2'' = z_2' + S_z \quad (26)$$

で表される。距離の誤差の式(24)において z' を z'' で置き換え、この誤差が0となるように修正量 S_z を求める。点 P_1 について

$$\hat{z}_1 - z_1'' = \frac{E_{X_1}}{E_{t_1}} \frac{\frac{E_{X_1} E_{t_1}}{z_1''} + \frac{E_{X_2} E_{t_2}}{z_2''}}{\frac{E_{X_1}^2}{z_1''^2} + \frac{E_{X_2}^2}{z_2''^2}} - z_1'' = 0 \quad (27)$$

となり、整理すると、

$$\frac{z_1' + S_z}{z_2' + S_z} = \frac{E_{X_1} E_{t_2}}{E_{X_2} E_{t_1}} \quad (28)$$

を得る。点 P_2 に関する方程式、

$$\hat{z}_2 - z_2'' = 0$$

からも式(28)と全く同じ方程式が導かれる。式(28)を S_z について解く。

$$S_z = \frac{E_{X_1} E_{t_2} z_2' - E_{X_2} E_{t_1} z_1'}{E_{X_1} E_{t_2} - E_{X_2} E_{t_1}} \quad (29)$$

ずれ S_z が動画像とずれたモデルから得られたことになる。このずれを条件式(16)を使って変形し、更に式(21)を代入すれば、

$$S_z = d_z \quad (30)$$

となり、実際にずれが正しく推定されていることがわかる。

以上の解法は、推定式(22)における距離 z' にモデルの修正量 S_z を加え、運動パラメータとモデルの修正量を未知数とする連立方程式を解くことに等しい。すなわち、連立方程式、

$$\begin{aligned} \frac{E_{X_1}}{z_1' + S_z} T_x + E_{t_1} &= 0 \\ \frac{E_{X_2}}{z_2' + S_z} T_x + E_{t_2} &= 0 \end{aligned} \quad (31)$$

を T_x , S_z を未知数として解くのである。この連立方程式の解は

$$T_x = \frac{E_{t_1} E_{t_2} (z_2' - z_1')}{E_{x_1} E_{t_2} - E_{x_2} E_{t_1}}$$

$$S_z = \frac{E_{x_1} E_{t_2} z_2' - E_{x_2} E_{t_1} z_1'}{E_{x_1} E_{t_2} - E_{x_2} E_{t_1}}$$

となる。修正量 S_z は先に求めた値(29)に等しい。これはずれの正しい推定値であった。また、移動量 T_x も条件式(16)を使って変形し、更に式(21)を代入すれば、正しい値であることが証明される。次節では、この手法を3次元運動の推定と位置・姿勢の3次元修正に一般化させる。

3.3 3次元モデルの位置・姿勢ずれの一般的修正

3次元モデルの姿勢のずれは、物体の中心 (x_c, y_c, z_c) を通る軸周りに回転させることにより、また、位置のずれは並進移動させることにより修正することができる。ずれが小さいとすれば、このずれの修正量は回転を表すベクトル $\mathbf{Q} = (Q_x, Q_y, Q_z)^T$ と、並進を表すベクトル $\mathbf{S} = (S_x, S_y, S_z)^T$ で表現できる。このとき、ずれの修正量と運動パラメータの推定問題は次のように定式化される。

動画像 $E(X, Y, t)$ とずれた3次元モデルから得られる物体の3次元情報が与えられたとき、正しい運動パラメータ \mathbf{T} , \mathbf{R} とずれの修正量 \mathbf{S} , \mathbf{Q} を求めよ。

前節の簡単な例題で述べた手法を要約すると次のようになる。(1)距離 z をずれの修正量を未知数とする関数で表す。(2)このとき、2.で導いた運動パラメータの推定式は、運動パラメータとずれの修正量を未知数とする方程式となる。従って、この方程式を解けば問題は解けたことになる。

まず、距離 z をずれの修正量 \mathbf{S} , \mathbf{Q} の関数で表す。記述の複雑さを避けるために、簡単な例での記述とは異なって、ずれたモデルから与えられる3次元情報を z とする。ずれたモデルから与えられる物体の3次元情報が、関数

$$z = f(x, y) \quad (32)$$

で表されているとする。ずれの修正量が与えられたとすれば、3次元情報は次のように修正される。ずれの修正により、モデル上の点 $\mathbf{p} = (x, y, z)^T$ は次式により点 $\mathbf{p}' = (x', y', z')^T$ に移る。

$$\mathbf{p}' = \mathbf{p} + \mathbf{Q} \times (\mathbf{p} - \mathbf{p}_c) + \mathbf{S} \quad (33)$$

式(33)を \mathbf{p} について解く。このとき、ずれの修正量 \mathbf{Q} , \mathbf{S} の要素同士の積は高次の微小量として無視すれば、 \mathbf{p} は次式で近似される。

$$\mathbf{p} = \mathbf{p}' - \mathbf{Q} \times (\mathbf{p}' - \mathbf{p}_c) - \mathbf{S} \quad (34)$$

式(34)を式(32)に代入し、 x' , y' , z' を x , y , z に書き換えれば、

$$z + Q_y(x - x_c) - Q_x(y - y_c) - S_z$$

$$= f(x + Q_z(y - y_c) - Q_y(z - z_c) - S_x,$$

$$y - Q_z(x - x_c) + Q_x(z - z_c) - S_y) \quad (35)$$

を得る。上式を z について解けば、修正された3次元情報

$$z = g(x, y, \mathbf{S}, \mathbf{Q}) \quad (36)$$

が得られる。ずれの修正がない場合には、

$$g(x, y, \mathbf{0}, \mathbf{0}) = f(x, y)$$

である。

式(36)はシーンの座標 (x, y) 上での3次元情報の表現である。しかし、直接観測できるのは投影面座標 (X, Y) であるので、投影面座標上での表現の方が便利なが場合が多い。シーンの座標による表現を投影面座標による表現に変換するには次のようにする。式(2)から $x = Xz$, $y = Yz$ を式(32)に代入し、 z について解くことにより

$$z = F(X, Y) \quad (37)$$

が得られる。同様に、式(35)に $x = Xz$, $y = Yz$ を代入すると次式が得られる。

$$z + Q_y(Xz - x_c) - Q_x(Yz - y_c) - S_z$$

$$= f(Xz + Q_z(Yz - y_c) - Q_y(z - z_c) - S_x,$$

$$Yz - Q_z(Xz - x_c) + Q_x(z - z_c) - S_y) \quad (38)$$

z について解けば、

$$z = G(X, Y, \mathbf{S}, \mathbf{Q}) \quad (39)$$

が得られる。ずれの修正がない場合には、

$$G(X, Y, \mathbf{0}, \mathbf{0}) = F(X, Y)$$

である。

さて、2.で束縛式(5)の $(\Delta X, \Delta Y)$ は式(4)で表される。更に、 z は修正量の関数(39)で表される。従って、同じ物体上の点から導かれる束縛式(5)の集まりは、運動パラメータ \mathbf{T} , \mathbf{S} , モデルの位置・姿勢の修正量 \mathbf{S} , \mathbf{Q} を未知数とする非線形連立方程式となる。

$$E(X + \Delta X, Y + \Delta Y, t + \Delta t) - E(X, Y, t) = 0 \quad (40)$$

但し、

$$\Delta = \frac{1}{z} \mathbf{A} \mathbf{T} + \left(\mathbf{B} - \frac{1}{z} \mathbf{B}_c \right) \mathbf{R}$$

$$z = G(X, Y, \mathbf{S}, \mathbf{Q})$$

2.と同様に、評価関数(6)を最小にするパラメータ値は、適当な初期値から出発して、ニュートン法⁽¹⁾により逐次最適な解に収束させてゆく。初期値は次のように与えておく。運動パラメータについては、ずれたモ

デルにより与えられる3次元情報 $z=F(X, Y)$ を使って推定されるパラメータ値を初期値とする。ずれの修正量については、ずれが小さいとみなして $S^{(0)}=0, Q^{(0)}=0$ とする。反復公式は、以下に与えられる。

$$T^{(n+1)}=T^{(n)}+\Delta T \quad (41)$$

$$R^{(n+1)}=R^{(n)}+\Delta R \quad (42)$$

$$S^{(n+1)}=S^{(n)}+\Delta S \quad (43)$$

$$Q^{(n+1)}=Q^{(n)}+\Delta Q \quad (44)$$

増分 $\Delta T, \Delta R$ および

$$\Delta S=(\Delta S_x, \Delta S_y, \Delta S_z)^T$$

$$\Delta Q=(\Delta Q_x, \Delta Q_y, \Delta Q_z)^T$$

は次の線形連立1次方程式を解くことにより得られる。

$$J(T^{(n)}, R^{(n)}, S^{(n)}, Q^{(n)}) \begin{bmatrix} \Delta T \\ \Delta R \\ \Delta S \\ \Delta Q \end{bmatrix} = e^{(n)} \quad (45)$$

ここで、 $J(T, R, S, Q)$ はヤコビ行列であり、

$$J(T, R, S, Q) = \begin{bmatrix} \vdots \\ J_{X,Y}(T, R, S, Q) \\ \vdots \end{bmatrix}$$

$$J_{X,Y}(T, R, S, Q) = \frac{\partial E(X+\Delta X, Y+\Delta Y, t+\Delta t)}{\partial (T, R, S, Q)}$$

但し、

$$\Delta^{(n)}=(\Delta X^{(n)}, \Delta Y^{(n)})^T$$

$$= \frac{1}{z^{(n)}} A T^{(n)} + \left(B - \frac{1}{z^{(n)}} B_c \right) R^{(n)}$$

$$z^{(n)}=G(X, Y, S^{(n)}, Q^{(n)})$$

ヤコビ行列の要素には、修正量 S, Q についての $z=G(X, Y, S, Q)$ の偏微分値が含まれている。この偏微分値は式(38)を偏微分することによって得られる。従って連立方程式(45)を構成する線形推定式は次のようになる。

$$\begin{aligned} & \frac{1}{z^{(n)}} a \cdot \Delta T + \left(b - \frac{1}{z^{(n)}} b_c \right) \cdot \Delta R \\ & + \frac{a \cdot T^{(n)} - b_c \cdot R^{(n)}}{(z^{(n)})^2} (c \cdot \Delta S - (d - d_c) \cdot \Delta Q) = e_{X,Y}^{(n)} \end{aligned} \quad (46)$$

但し、

$$a = \begin{bmatrix} E_X \\ E_Y \\ -XE_X - YE_Y \end{bmatrix}$$

$$b = \begin{bmatrix} -Y(XE_X + YE_Y) - E_Y \\ X(XE_X + YE_Y) + E_X \\ XE_Y - YE_X \end{bmatrix}$$

$$b_c = \begin{bmatrix} -y_c(XE_X + YE_Y) - z_c E_Y \\ x_c(XE_X + YE_Y) + z_c E_X \\ x_c E_Y - y_c E_X \end{bmatrix}$$

$$c = \frac{1}{h} \begin{bmatrix} f_x \\ f_y \\ -1 \end{bmatrix}$$

$$d = \frac{z^{(n)}}{h} \begin{bmatrix} Y + f_y \\ -X - f_x \\ Yf_x - Xf_y \end{bmatrix}$$

$$d_c = \frac{1}{h} \begin{bmatrix} y_c + z_c f_y \\ -x_c - z_c f_x \\ y_c f_x - x_c f_y \end{bmatrix}$$

$$h = 1 - Xf_x - Yf_y - (Y + f_y)Q_x^{(n)} + (X + f_x)Q_y^{(n)} + (Xf_y - Yf_x)Q_z^{(n)}$$

ここで、モデル面の傾き (f_x, f_y) は

$$x = Xz^{(n)} + Q_z^{(n)}(Yz^{(n)} - y_c) - Q_y^{(n)}(z^{(n)} - z_c) - S_x^{(n)}$$

$$y = Yz^{(n)} - Q_z^{(n)}(Xz^{(n)} - x_c) + Q_x^{(n)}(z^{(n)} - z_c) - S_y^{(n)}$$

のときの $f(x, y)$ の偏微分値である。

反復計算をするたびに、得られる修正量を使ってモデルを逐次修正すれば、上記の繰返し推定式は次のように簡単になる。モデルの修正が行われているわけであるから、 $n+1$ 回目の繰返し計算は、 n 回目のモデルの位置を初期状態として行われる。従って、ヤコビ行列の要素は、常に $S^{(n)}=0, Q^{(n)}=0$ とみなして計算してよく、 (f_x, f_y) は投影面上の点 (X, Y) に対応する修正されたモデル面の傾き (g_x, g_y) となる。従って、式(46)の係数のうち c, d, d_c, h は次のように簡単になる。

$$c = \frac{1}{h} \begin{bmatrix} g_x \\ g_y \\ -1 \end{bmatrix}$$

$$d = \frac{z^{(n)}}{h} \begin{bmatrix} Y + g_y \\ -X - g_x \\ Yg_x - Xg_y \end{bmatrix}$$

$$d_c = \frac{1}{h} \begin{bmatrix} y_c + z_c g_y \\ -x_c - z_c g_x \\ y_c g_x - x_c g_y \end{bmatrix}$$

$$h = 1 - Xg_x - Yg_y$$

この係数には、モデル面までの距離 $z^{(n)}$ とその面の傾き (g_x, g_y) が含まれている。これらの量は式(35)あるいは式(38)を g について解く必要はなく、モデルの修正のたびに幾何モデラー⁽³⁾ によって直接与えられる。

線形推定式(45)はその係数行列 $J(T, R, S, Q)$ が正則なときに限り一意な最適解

$$[J^T J]^{-1} J^T e^{(n)}$$

を得る。繰返しの停止条件は、評価関数(6)が減少しなくなるときとする。繰返しが終了したとき、モデルは正しい位置・姿勢におかれ、3次元運動パラメータが推定されている。

4. ユニークな解をもつための条件

対象物体の形状やその像パターン、あるいはその運動によっては、運動パラメータやずれの修正量が一意に決定できない場合がある。運動パラメータとずれの修正量は推定式(45)を解くことにより得られるが、その係数行列 $J(T, R, S, Q)$ が正則でない場合、解は一意に決定できない。この係数行列が正則であるか否かのための必要十分条件を求めることは困難である。ここでは、次の三つの代表的な場合にユニークな解が得られないことを示す。

[条件1] $J(T, R)$ が正則でなければ、解は一意に決まらない。 $J(T, R)$ は $J(T, R, S, Q)$ の第1列から第6列までからなる小行列である。従って、 $J(T, R)$ が正則でなければ $J(T, R, S, Q)$ も正則でない。 $J(T, R)$ は運動パラメータのみを推定したときの推定式(9)の係数行列である。運動パラメータがユニークに決まらない対象として、床屋のサインポールや表面模様のないバケツなどがある⁽⁹⁾。

[条件2] $J(S, Q)$ が正則でなければ解は一意に定まらない。 $J(S, Q)$ は $J(T, R, S, Q)$ の第7列から第12列までからなる小行列である。従って、 $J(S, Q)$ が正則でなければ $J(T, R, S, Q)$ も正則ではない。係数 $J(S, Q)$ は距離動画像から運動パラメータを推定した式の係数行列と全く同じである⁽¹⁰⁾。ずれの修正量がユニークに決まらない例として、平面、球、回転体などがある。

[条件3] 物体の運動をカメラの運動と解釈したとき、移動ステレオ視が成立しないならば、 $J(T, R, S, Q)$ は正則でなくなる。物体とカメラの運動は互いに相対的なものであり、物体の動きをカメラの動きとみなすことができる。式(1)は次のように書き換えられる。

$$\delta = R \times p + t$$

$$t = \begin{bmatrix} T_x + y_c R_z - z_c R_y \\ T_y - x_c R_z + z_c R_x \\ T_z + x_c R_y - y_c R_x \end{bmatrix}$$

この $-t$ が物体の動きをカメラの動きとみなしたときのカメラの並進運動成分である。このとき、次の関係式が成り立つ。

$$a \cdot t = a \cdot T - b_c \cdot R$$

右辺の $a \cdot T - b_c \cdot R$ は式(46)の ΔS と ΔQ の係数、す

なわち、 $J(T, R, S, Q)$ の第7列から第12列までのすべての要素に乗ぜられている量である。従って、カメラの並進移動成分 $-t$ が 0 ならば $J(T, R, S, Q)$ は正則とならない。

上記の3条件以外にも、解がユニークに決定できないことがある。3.2で述べた簡単な例を再び考える。このとき推定式(45)の係数行列は

$$\begin{bmatrix} \frac{E_{x_1}}{z_1' + S_z}, -\frac{E_{x_1} T_x}{(z_1' + S_z)^2} \\ \frac{E_{x_2}}{z_2' + S_z}, -\frac{E_{x_2} T_x}{(z_2' + S_z)^2} \end{bmatrix} \quad (47)$$

となり、この行列式は

$$\frac{E_{x_1} E_{x_2} T_x (z_2' - z_1')}{(z_1' + S_z)^2 (z_2' + S_z)^2} \quad (48)$$

となる。ここでは、反復回数を示す添字は省略した。

[条件1]～[条件3]より $E_{x_i} = 0$, $T_x = 0$ のとき行列式の値は0になる。しかし、これ以外にも、明らかに $z_1' = z_2'$ のとき行列式の値が0となり、ユニークな解をもたない。

動きが一意に決定できないことが予想される場合には、あらかじめ推定式の次数を下げておく。

5. 実 験

提案した手法の効果を実験により確かめる。

物体の3次元モデルは幾何モデラー⁽³⁾を使ってあらかじめ作成しておく。この幾何モデラー⁽³⁾はずれの修正量 S , Q の値が具体的に与えられているならば、

(1) 距離マップ(画素点に対応するモデル表面までの距離) $g(x, y, S, Q)$, $G(X, Y, S, Q)$

(2) 物体モデルの投影像(画像中に占めるモデルの領域、モデル間の隠れ処理も行っている)

(3) 物体モデル面の傾き (g_x, g_y)

などの情報を提供することができる。

まず、運動パラメータの推定とモデルの位置・姿勢の修正が同時に可能であることを、3.2で述べた簡単な例についてシミュレーション実験で確認した。シミュレーションで用いた数値は次のように与えた。図1において、 $z_1 = 5$, $z_2 = 4$ なる物体は x 軸の正方向に速度 $T_x = 0.01$ で進むとする。モデルは正しい位置から z 軸の正方向に3ほどずらせておいた。2枚の画像からなる動画は計算機内で生成した。推定結果を表1に示す。8回の繰返して速度は正しく推定された。モデルの位置はいったん正しい位置を大きく通り越した後、ほぼ正しく修正されている。モデルの最初の位置ずれ

表1 物体速度の推定とモデル位置の修正過程

Table 1 Estimating the object motion and correcting the model position.

Iteration	Velocity(T_x)	Depth(z_1)
Initial values	0.0112	8.0000
1	0.0149	3.0147
2	0.0086	2.8312
3	0.0083	3.3275
4	0.0085	3.8482
5	0.0091	4.3417
6	0.0096	4.6916
7	0.0099	4.8773
8	0.0100	4.9558
Real vales	0.0100	5.0000

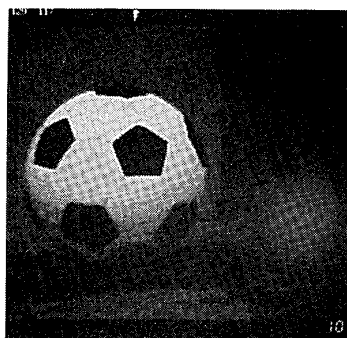


図2 追跡対象物体：第1フレーム

Fig. 2 A ball as tracking a target at the 1st frame.

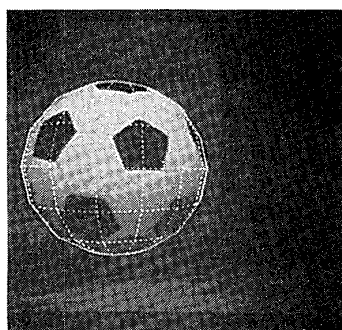


図3 モデルの初期位置

Fig. 3 3D model of the object at starting frame.

が3よりも大きくなると、修正の途中でモデルがカメラの内側に入り込み修正を続行することが不可能となった。物体の初期位置をカメラから離すにつれ、推定式の係数行列式(48)はその分母が次第に大きくなるため0

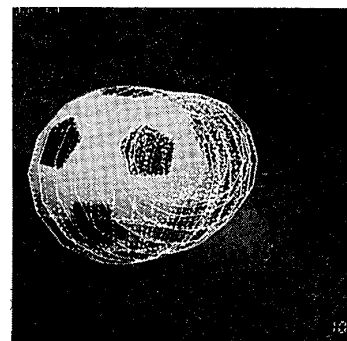


図4 10フレームごとのモデルの位置：ずれの修正を行わない追跡結果

Fig. 4 The 3D model at every 10 frames without correcting pose displacement.

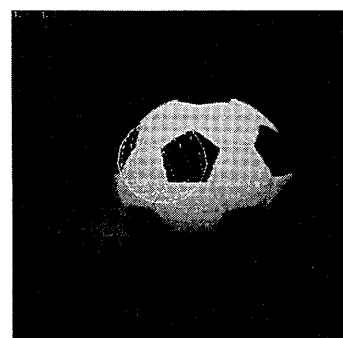


図5 ずれを修正しなかったときの、100フレーム目のモデルの推定位置

Fig. 5 The 3D model at 100th frame without correcting pose displacement.

に近づく。このとき推定式は悪条件となった。

実際に物体の追跡実験を行った。

動画像の1コマを図2に示す。画像のサイズは128×128で明るさの階調は256である。対象物体はサッカーボールで天井から細い糸でつり下げられている。糸はカメラの光軸にほぼ垂直な面に振り子運動し、ボール自身は糸を軸に回転運動をしている。背景は普通の実験室内である。このような複雑な動画像を100フレーム解析した。この例では、対象物体が球である。これは位置・姿勢の修正量がユニークに決定できない[条件2]を満たす。修正量がユニークに決定できないのは中心点周りの回転ずれである。そこで、修正量のうち並進成分のみを推定した。更に実際は、ずれの並進成分のうち、 z 軸方向(カメラの光軸方向)の成分は推定条件が悪条件に近いために意味のある解を得ることが難しい。これは物体の位置がその大きさに比べてカメラから遠かったためと考えられる。従って、推定した修正量は x と y 軸方向、すなわち S_x 、 S_y のみである。一方、運動パラメータは6成分すべて (T, R) 推定した。

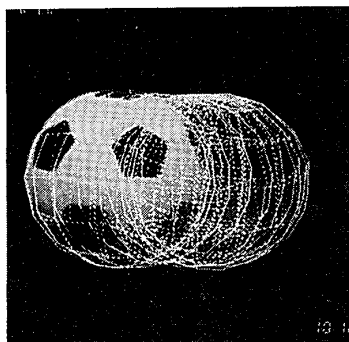


図6 ずれの修正を行ったときの10フレームごとの追跡結果

Fig. 6 3D model at every 10 frames by correcting pose displacement.

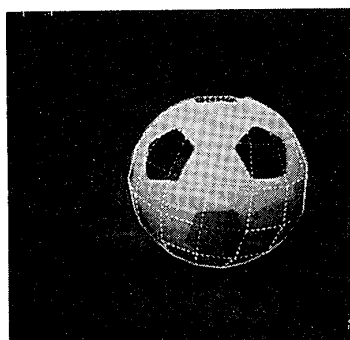


図7 ずれの修正を行ったときの64フレーム目のモデルの推定位置

Fig. 7 3D model at 64th frame by correcting pose displacement.

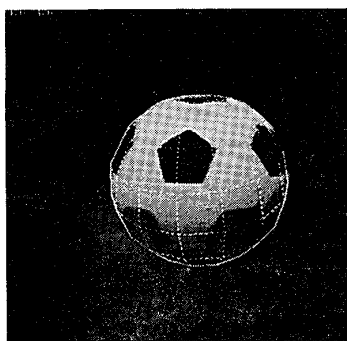


図8 ずれの修正を行ったときの100フレーム目のモデルの推定位置

Fig. 8 3D model at 100th frame by correcting pose displacement.

図3は追跡開始フレームにおけるモデルが示されている。物体とモデルの重合せは手で行った。物体上の特徴点(エッジ強度の大きい所)は約300点であった。モデルのずれの修正を行わないままに、追跡を行った結果が図4, 5で示されている。図4には10フレームごとのモデルの位置を第1フレームの画像に重ねて表

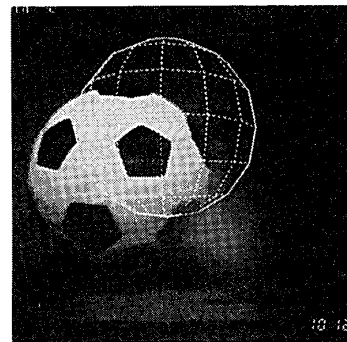


図9 ボールの半径分ずれたモデルの初期位置
Fig. 9 Initial position of the 3D model.

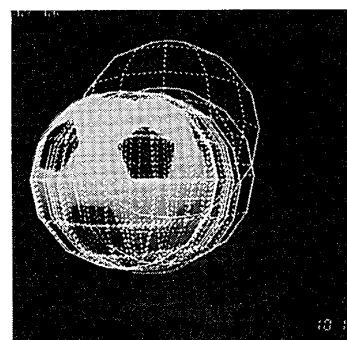


図10 モデルのずれの修正過程
Fig. 10 Correcting the pose of the 3D model at every iteration.

示した。図5は100フレーム目の画像とモデルを重ねて表示した。ずれが極端なまでに拡大したことがわかる。このずれのパターンはモデルの初期位置により変化する。

位置ずれを自動修正しながら追跡した結果を図6～8に示す。10フレームごとのモデルの推定位置を第1フレームの画像に重ね合わせた結果が図6に示されている。また、64フレーム目と100フレーム目のモデルの推定位置がそれぞれ図7, 8に示されている。ボールは自身の回転により自己オクルージョンを起こしている。それにもかかわらず追跡可能なのは3次元モデルにより対象の全形状が与えられているからである。各フレームでのパラメータ推定のための繰返し回数は3, 4回程度であった。また、計算時間は繰返し数にもよるが、1フレーム間当たり約5秒(Sparc Station 2)であった。図8のモデルの推定位置を、位置の修正を行わなかった場合(図5)と比較する。奥行き方向のずれの修正を行っていないにもかかわらず位置が正しく得られている。これは、今回の実験では物体がカメラ光軸方向にはほとんど移動しなかった。そのため、 x 軸、 y 軸方向の修正のみでモデルのずれが修正され、光軸方向の移

動量(0に近い)もほぼ正しく得られた。しかし、光軸方向の位置ずれの修正を行っていないため、更に長時間の追跡を実行した場合、光軸方向のずれが大きくなる可能性がある。

次に、修正できるずれの限界を調べた。第1フレームと第2フレームの動画像から位置の修正と運動パラメータの推定を行った結果を示す。図9には修正前のモデルが示されている。このモデルの位置は正しい位置からボールの半径分ほどずれている。図10はニュートン法の繰返しごとに、モデルのずれが修正される様子が示されている。この実験では7回の繰返し演算で正しい位置に収束した。モデルの初期位置をさまざまに変えて実験を行った。その結果、ボールの半径の半分程度のずれは、完全に修復されることが明らかになった。大きな初期位置ずれが与えられた場合には、背景を物体とみなす割合が多くなるためずれの修正はできなくなった。

6. む す び

物体の3次元形状が3次元モデルにより正しく与えられているとする。しかし、このモデルが正しい位置と姿勢からずれているとき、動画像から位置と姿勢のずれを修正し同時に運動パラメータを推定する手法を提案した。本手法は、モデルと物体像との照合も連続する画像間の照合をも必要とせず、物体の追跡に応用できるのが特徴である。

今回は手法の提案を行ったが本手法の評価は今後の課題である。用いた推定式が非線形連立方程式であるため、一般には複数の解が存在し、偽の解が求まる可能性がある。また、真の解の近傍で方程式系が悪条件となる場合には解が安定に収束しない。正しい解を得るための条件、解を安定に収束させるための条件を、対象物体の形状、物体表面の模様、カメラの特性、カメラと物体との相対位置、等について詳細に検討する必要がある。

本論文では明るさの不変性をより所に画像間の照合を行っている。しかし、物体の周辺部ではシェーディングのため明るさの不変性は成り立たなくなる。対処的には、エッジの位置はシェーディングの影響をあまり受けないので、画像間の明るさの差の代わりにフレーム間で対応するエッジ間の距離を用いることが考えられる。しかし、根本的には明るさの変化をも考慮した照合を行う必要がある。

今回提案した手法は単眼視を仮定しているため、光

軸方向の修正量の推定は悪条件に陥り困難であった。ステレオ動画像を使い光軸方向をも含むずれのロバストな推定法を別の機会に発表したい。

謝辞 本研究に、一部御支援頂きました新潟大学地域共同研究センターと(株)三協精機製作所に感謝致します。

文 献

- (1) 富士通編: "SSL II使用手引書 (科学用サブルーチンライブラリ)" (1980).
 - (2) Horn B. K. P. and Weldon Jr. E. J.: "Direct methods for recovering motion", International Journal of Computer Vision, 2, pp. 51-76 (1988).
 - (3) Koshikawa K. and Shirai Y.: "A 3-D modeler for vision research", ICAR, pp. 185-190 (1985).
 - (4) Lowe D. G.: "Integrated treatment of matching and measurement errors for robust model-based motion tracking", Proc. of 3rd ICCV, Osaka, pp. 436-440 (1990).
 - (5) Lowe D. G.: "Fitting parameterized three-dimensional models to images," IEEE, PAMI-13, 5, pp. 441-450 (1991).
 - (6) Negahdaripour S. and Horn B. K. P.: "Direct passive navigation", IEEE, PAMI-9, 1, pp. 168-176 (1987).
 - (7) Negahdaripour S.: "Direct methods for structure from motion", Ph. D Thesis, Mechanical Engineering, MIT (1986).
 - (8) Tsai R. Y. and Huang T. S.: "Estimating three-dimensional motion parameters of a rigid planar patch", IEEE, ASSP-29, 6, pp. 1147-1152 (1981).
 - (9) 山本正信: "動画像と距離情報の併用による3次元運動パラメータの直接的推定法", 信学論(D), J68-D, 4, pp. 562-569 (1985-04).
 - (10) 山本正信, Boulanger P., Beraldin J.-A., Rioux M. and Domey J.: "距離動画像を用いた非剛体運動パラメータの直接的推定法", 情処論, 32, 9, pp. 1129-1141 (1991).
- (平成5年3月29日受付, 9月6日再受付)

山本 正信



昭48九工大・工・制御卒。昭50東工大大学院修士課程了。同年、電子技術総合研究所入所。平4より新潟大工学部情報工学科教授。工博。動画像処理、コンピュータビジョン、知能ロボットなどの研究に従事。平1～2カナダ国立研究協議会招へい研究員。昭62情報処理学会研究賞授賞。情報処理学会、計測自動制御学会、日本ロボット学会、IEEE各会員。