

三面世界における運動の解釈

浅田稔, 谷内田正彦, 辻三郎
(大阪大学 基礎工学部)

1 はじめに

近年, コンピュータ・ビジョンの新しい分野として動画像解析が注目され, 種々研究されている。^[1~3] 例えば, 物体の動きの検出, 分割や動物体の追跡,^[4] 心臓の動きの解析等が挙げられる。しかしそのほとんどが, 画像内の変化を2次元的な動きとして解釈している。Roach^[5]は凸多面体を発見, 追跡し, その直線運動を解析した。彼らの主眼は階層的なマッチング・プロセスを駆使して, 実際のシーンの画像から, 積み木を追跡する事であり, 回転運動を含む一般的な動きを対象にしていない。

Ullman^[6]は心理学の立場から, コンピュータが3次元運動を理解するための基本的な理論を提唱した。彼の理論や同様の数学的理論^[7,8]を用いれば, 対応点が既知であるという仮定の下に, 2次元画像列から物体の3次元構造が求められる。しかしこの事から, コンピュータが3次元運動を理解したとは言えない。即ち, これらの計算から, 各物体の位置と向きについて, 画面間の変化を記述できても, 我々人間と同じ様に, それらを平易で判りやすい形に表現できない。

我々は理論的立場から, 回転運動を含む一般的な3次元運動を理解するための研究を行い, 対応付けの問題や, 3次元構造の決定問題等^[9,10]を扱って来た。本文では, 三面世界で動く物体の向き, 位置を決定する3次元パラメータを入力として, 3次元運動を復元し, 人間と同じ様に簡潔かつ自然に運動を解釈する手法について述べる。

入力画像は, 人が腕を振って歩き, 角を曲がる様なシーンで, 人間にとって簡単に理解できるが, コンピュータ

にとって複雑な動きを示している。このような運動を解釈するために, 本方法は以下の3つのプロセスからなる。

1) ここでは物体の3次元パラメータが直交射影で与えられた画像から算出されたと仮定しているのど, 実行き方向の平行移動量を計算できない。そこで先ず各画面内における物体の大きさと向きから画面間の回転運動を決定する。2)次に関節運動の様な物体相互の複雑な動きを解釈するために, シーン中の物体から主部を発見する。主部とは, それに固定された座標系から他の物体の運動を観測した時, それが見がけの運動より簡単になるものを指し, ここでは画面間の回転軸の向きの変化を評価基準としている。即ち, 軸の向きの変化の少ないものを主部の候補として選び, その資格を試す。この様に, 主部を発見し, それに付随する他の物体の運動を記述し, シーンの階層的記述を得る。

3)最後に物体の平行移動を評価し, 3次元運動の簡潔で自然な解釈を試みる。画面間の動きの記述において, 回転軸の方向や回転角が一意的に決定されるのに対し, 平行移動成分は回転軸の位置によってその値が変る。逆に, 平行移動が滑らかに記述できる様に回転軸の位置を調節できる。そこで滑らかに記述するために平行移動の性質(等速直線運動等)を想定して, 運動を記述し, それが多い画面に適合できる時, その性質を用いて運動を表現する。もしどの性質も合致しなければ, それらの画面間で運動が急激に変化したと解釈し, 画像列をその前後に分割する。

この様にしてダイナミックシーンを理解するために, 各物体の運動を空間的

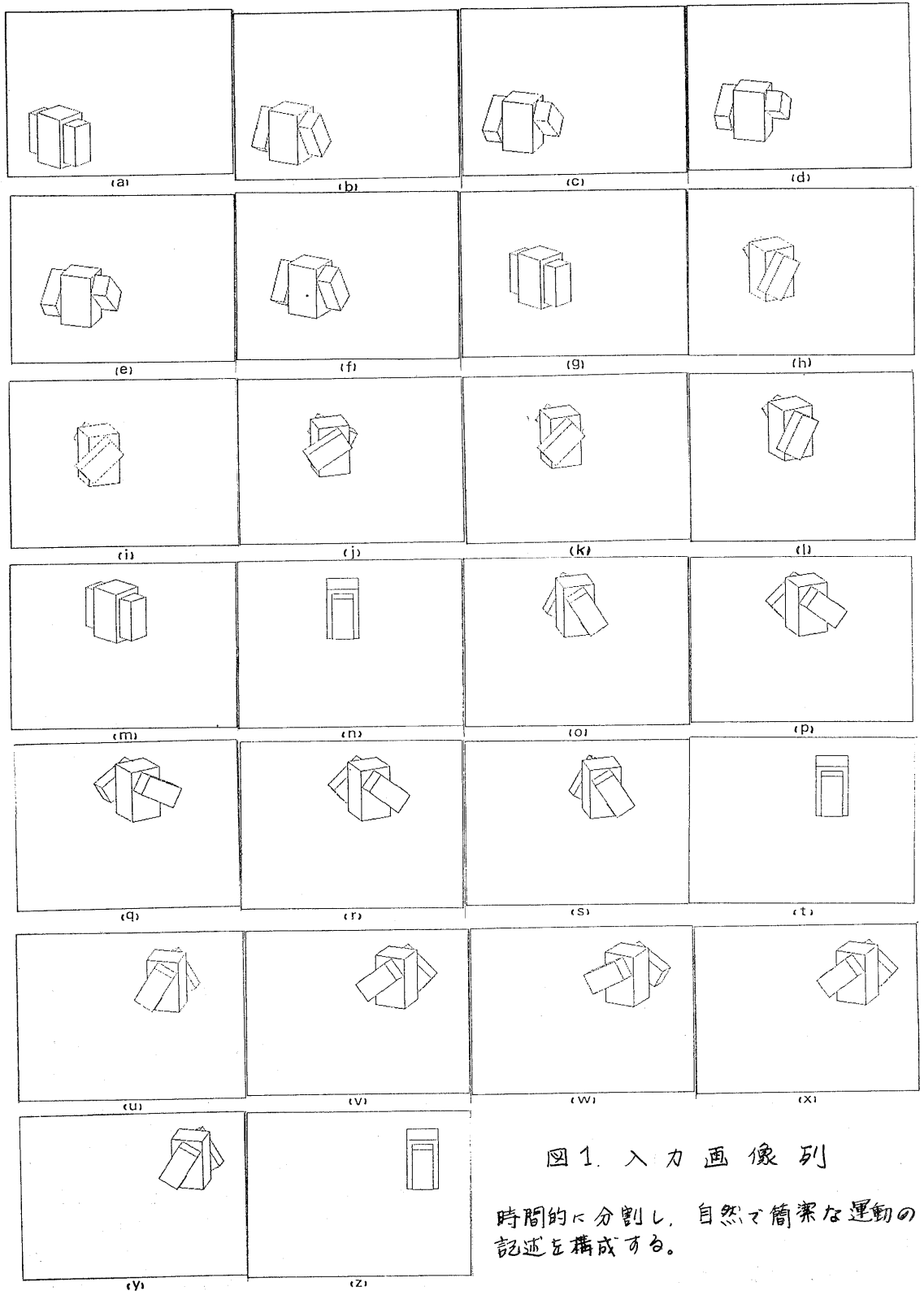


図1. 入力画像列

時間的に分割し、自然な簡潔な運動の記述を構成可る。

2. 画面間の動きの計測

2.1 3次元モデルの復元

各画面間で物体の対応点を与えられている時、我々は元のシーンの物体の3次元モデル、即ち、物体上の点の3次元座標を計算できる。Roach^[57]やMeiri^[58]は、中心射影の場合について、3次元パラメータを計算するために必要な点と画面の数について考察した。この場合、拡大係数が未知なので3つの座標の比が計算されている。Ullman^[61]やAsada^[62]らは直交射影の場合について研究し、3画面から3~4点の3次元座標を求める方法を提案している。

2.2 入力画像

Asadaらの方法を用いれば、物体上の各点の相対的な3次元座標が求まるので、本文ではそれら3次元座標をコンピュータへの入力としている。入力画像は26枚の画面からなり、図1(a)~(d)に示されている。このシーンでは、2つの小さな積み木(物体2,3)が他の運動する積み木(物体1)上で互いに逆方向に振り子状の運動を示し、あたかも人が両腕を振って歩き、角で曲がると言ったシーンを連想させる。上記の手法を用いれば、物体1,2について3次元パラメータが算出できるが、物体3については見えている部分が少く、正確な値が得られない。そこで今回の実験では物体3について3次元パラメータは入力されていない。また上記の手法は直交射影を想定しているので、奥行き方向の平行移動量は算出できず、それらの情報は使えないものとする。

2.3 画面間の回転運動の計測

一般に空間内での動きは1つの回転と1つの平行移動で表現される。ここでは先ず、回転運動の成分(回転軸の方向と回転角)の求め方について述べる。平行移動に関しては次章以降で触れる。

図2は積み木の動きの簡単な例を示

している。今、物体上の板の原点(物体上のどこでも構わない)の1枚目の位置 O_1 から2枚目の位置 O_2 へ平行移動した後、 O を通る軸の回りに物体が回転するものとする。実線と破線はそれぞれ、平行移動ベクトル、回転運動ベクトルを示す。図3は回転運動成分決定の原理図を描いたもので、平面 OAB_1 が O を通る軸 l の回りに回転して、平面 OAB_2 の位置に来た場合を示している。この図から点 A_2, A_3 (B_1, B_2)は l 上の点 O_A (O_B)を中心とする円周上にある事が判る。当然、線分 A_1A_2 (B_1B_2)が l となす角は直角である。今、これらの点の3次元座標が与えられているので、この条件から直線 l の方程式は容易に求められる。また直線 l の方程式を用いて回転角 $\angle A_1O_AA_2$ ($\angle B_1O_BB_2$)も計算できる。

この方法によって求めた各画面間の回転軸の向きの変化を図4(a)(物体1)、(b)(物体2)にそれぞれ示す。回転軸の方向は半球上の小円として表わされており、最初の画面間の回転軸が天頂に来る様に変換している。また時間順に小円の中心を結んでいる。この図から、物体1がシーン全体を通して回転軸の向きが一定しているのに対し、物体2は各画面間で軸方向がよく変化しているのが判る。

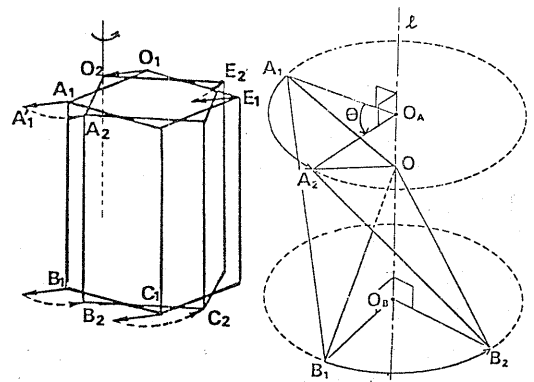


図2. 物体の動き 図3. 回転軸決定の原理図

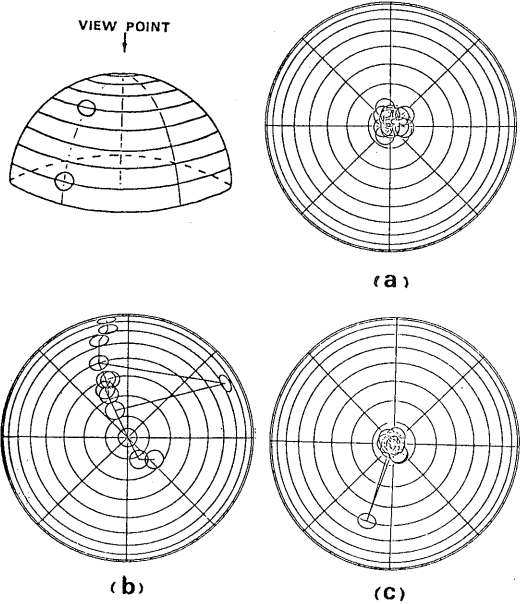


図4. 回転軸の向きの変化

3. ダイナミック・シーンの階層的記述 (空間的な分割)

前章で得られた各画面間の動きの記述をつなぎ合わせれば、一応シーン全体の動きを記述できる。しかしこの方法で複雑な動きを簡潔に表現できるか疑問である。例えばコンピュータに歩いている人間の腕を見せ、その運動を記述させる事を考える。腕の動きはカメラ座標系から観測され、記述されるので、非常に複雑な表現になり、余り有効な情報を与えてくれな。これに対し人間はその動きを「人が腕を振って歩いている」という様な簡単な記述で表現している。この場合、カメラ座標系も一般座標系も使われず、その代り動く人間自体に座標系が設定されていると考えられる。そこでこの様な表現法を我々のシーンに適用する事を考える。この時、重要な条件は、① 物体の運動を簡潔に表現するための座標系を発見する事、② その座標系から見た他の物体の動きも簡潔にできる事である。

我々は運動表現の簡潔さを以下の様に考えている。もし運動の性質が充分

長い間一定であれば、その運動は数少しいパラメータで記述できる。例えば、運動の2つの要素、回転運動と平行移動が一定であれば、その運動は簡潔に表現できる。そこで記述の簡潔性の目安として、本文では運動の2つの要素が変化しない事を採用する。以下では回転軸の方向が一定である時、簡潔な表現ができる例を示す。

図1に示す様に、ある物体が別の運動する物体上で回転運動する場合、前者の見えの動きは、後者より複雑になる。これはカメラ座標系で個々の動きを記述するからで、運動自体を階層的に記述すれば、より簡潔な表現ができると考えられる。物体相互の動きを解析するために物体間の関係(例えば「支えられている」とか「接触している」)を用いて表現する事も考えられるが、我々は各画面間の動きの情報を利用し物体間の関係を明らかにする。ここでは人工知能の分野で良く使われる手法、即ち、仮説を立ててそれを立証する手法を用いて物体の集りの中から主部を発見する。主部の候補としては、回転軸方向の変化の少ないものが選ばれ、その候補に対する他の物体の相対的な動きが計測される。この時、相対的な動きが簡潔に表現できれば、その候補は主部と断定され、物体間の関係が明らかになる。

図4(a)から物体1は回転軸がほぼ一定しているので、主部の候補として選ばれる。そこで物体2について物体1に対する相対的な動きを計測し、それが簡単に表現できるか試みる。物体2の相対的な動きを計測するために、ここでは、物体1が図1(a)から回転しない様に図1(b)へ(2)座標変換して物体2の回転運動を求めた。図4(c)はこの様に求めた物体2の相対的な回転運動の軸方向を示す。この図で1つの小円が中心から離れているのは、これら2つの画面間(図1(m)から(n))で物体2がほ

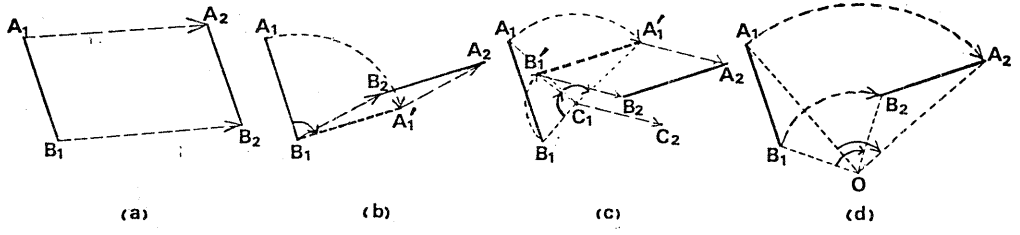


図5. 2画面間の2次元の動きの表現

とんど回転しながら、 T 為に生じたものである。この図から物体1に対する物体2の相対的回転運動の軸方向は一定であると判断される。

この様にして、物体相互の複雑な動きも、階層的に記述する事により、人間が知覚する形に近い、自然で簡潔な表現が可能になる。

4. より簡潔な記述の発見

(時間的分室)

前章では、回転運動の記述を基にしてダイナミック・シーンを階層的に記述した。本章では前章で扱われなかった平行移動について自然で簡潔な表現を試みる。

4.1 2次元運動の表現

3次元運動を記述する前に、2次元の場合について考察し、それを3次元に拡張する。

図5は2次元運動の例で、線分 AB が1枚目の位置 A_1B_1 から次の画面の位置 A_2B_2 に移動した場合を示す。図5(a)に示す様に $\vec{A_1B_1}$ が $\vec{A_2B_2}$ に等しければ、それらの間の動きは平行移動だけで表現される。等しくない場合は、図5(b)に示す様に、1つの回転と1つの平行移動に分割できる。この時、回転角は線分 A_1B_1 と A_2B_2 のなす角として一意に決定されるが、平行移動は一意に定まらず、回転中心の位置に左右される。(図5(b),(c))そこで平行移動量が最小になる様に回転中心を設定すれば、この2画面間においては最も簡潔な表現になると考えられる。実際には平行移動量を零にする(証明は簡単)事が可能で、図5(d)に示

す様に回転中心の位置が唯一に決定される。即ち2次元の運動で、2画面間の動きは、回転運動のみが、平行移動のみによって表現できる。

次にこの記述法が多くの画面を通して自然な表現となるか考える。図6は線分 AB の3画面における動きの例で、 A_1B_1, A_2B_2, A_3B_3 はそれぞれの画面の位置を示している。最初に各画面間に上記の方法を適用すると C_{12}, C_{23} がそれぞれの画面間の回転中心となる。この方法では回転中心が第2画面で突然移動し、非常に不自然であり、回転中心の平行移動が3画面を通して一定の性質を持った方がより自然である。3画面で規定される最も基本的な性質は等速運動であり、後で述べる方法でこの性質に従って運動を記述すると、 O_1, O_2, O_3 が各画面における回転中心の位置となる。即ち、線分 AB は O を中心に回転し、かつ等速直線運動をしているという簡潔で自然な記述が得られる。

では一般に n 画面とえられた時、ど

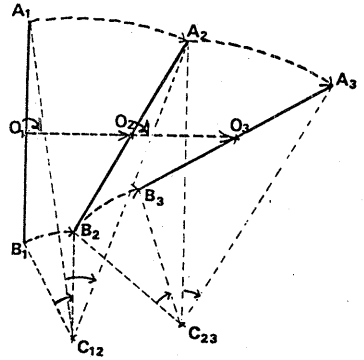


図6. 連続3画面の2次元の動きの表現

の様な性質が可能を考へる。今、 $A(I), O(I)$ を i 番目の画面における点 A と回転中心 O を表わす 2次元ベクトルとする。また $L(I) = (L_x(i), L_y(i))$ を i 番目の画面から次の画面までの平行移動を表わすベクトルとする。これから、 i 番目の画面から次の画面へ点 A と O を移す変換方程式は次の様に表わされる。

$$A(I+1) = T(I)(A(I) - O(I)) + O(I) + L(I) \quad (1)$$

$$O(I+1) = O(I) + L(I) \quad (2)$$

ここで

$$T(I) = \begin{pmatrix} \cos\theta(i) & -\sin\theta(i) \\ \sin\theta(i) & \cos\theta(i) \end{pmatrix}$$

また式(1),(2)から次式が得られる。

$$A(I+1) = T(I)A(I) +$$

$$(I - T(I)) \left(\sum_{K=1}^{I-1} L(K) + O(1) \right) + L(I) \quad (3)$$

即ち、 $2(n-1)$ 個の方程式に対し、 $2n$ 個の未知数 ($O(I)$ の 2成分と $2(n-1)$ 個の $L(I)$) が存在する。そこで未知数の間に 2つの関係式を付加する事により、運動を一意に決定できる。例えば 3画面の時は 4つの変換方程式に対し 6つの未知数 ($L_x(1), L_y(1), L_x(2), L_y(2)$) と回転中心の座標 (O_x, O_y) が存在する。そこで次の 2つの関係式を採用する。

$$L_x(1) = L_x(2), \quad L_y(1) = L_y(2) \quad (4), (5)$$

これらの関係式は平行移動が等速であることを示し、(3),(4),(5)の解がこれらのパラメータを与える。

4.2 3次元運動の表現

物体は各画面間で平行移動し、かつ 1軸回りで回転運動しており、その軸方向と回転角は既に求められている。そこで先ず以下の様に座標変換する。即ち、元の α, β, γ 座標系から、 Z 軸が回転軸と平行な (X 軸が γ - Z 平面に存する) X, Y, Z 座標系へ変換する。この変換により α, β, γ 空間で 1軸回りの回転(図7(a))が図7(b)に示す様に、 X - Y 平面上で点回りの回転に帰着でき、 Z 軸方向の動き

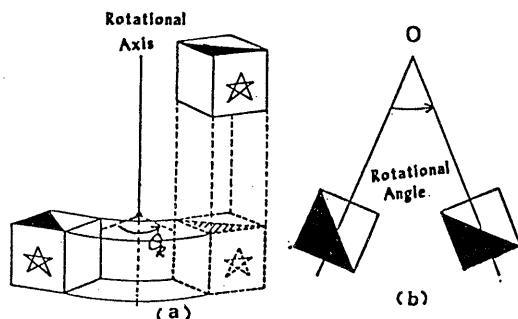


図7. 2画面間における3次元の動き

は平行移動によつてのみ表現される。2次元の場合と同様に平行移動成分は一意に定まらず回転軸の位置に依存するので移動量が最小になる様に決定する。その結果、平行移動は回転軸と平行になる。この様にして、2画面間における3次元運動は、ある軸回りの回転とその軸方向の平行移動によつて一意に決定される。

上記の方法は2画面間での物体の奥行き方向の関係 (Z 座標の差) が明らかでない時適用できるが、我々は直交射影を想定しているのだから、これらの関係は物体上の仮の原点の位置によつて異なる。そこで本文では奥行き方向 (Z 軸方向) の平行移動成分が零であると仮定して解析をすすめている。この条件は一見不自然に感じられるが、我々人間が時を経験する知覚と似ている。例えば、遠方から走って来る人間を望遠鏡で観た場合、実際に走者が近付いてくるにも関わらず、観測者がその様に感じない場合がこれに当る。

次に2次元の場合と同様に、一般に n 画面をえらべた時について考える。この場合、回転軸の向きが一定でない時の問題が生じる。前章で構成したダイナミックシーンの階層的記述から、その様な動きは単純な回転の組み合わせと考えられるので、以下の論議では回転軸の方向が一定である場合について述べる。今、 $(\alpha(i), \beta(i))$ を i 番目の画面における物体上の点の画像面上の座標

とすれば、そのX, Y, Z座標は次式で与えられる。

$$\begin{pmatrix} X(i) \\ Y(i) \\ Z(i) \end{pmatrix} = P \begin{pmatrix} x(i) \\ y(i) \\ z(i) \end{pmatrix} \quad (6)$$

ここでPはx, y, z座標系からX, Y, Z座標系へ変換する座標変換行列である。またO(I)をi番目の画面におけるX-Y平面上の回転軸の位置を示すベクトルとし、L(I)を回転軸の平行移動量を表すベクトルとすれば、2次元の場合と同様に以下の式が得られる。

$$\begin{pmatrix} X(i+1) \\ Y(i+1) \end{pmatrix} = T(I) \begin{pmatrix} X(i) \\ Y(i) \end{pmatrix} + (I-T(I)) \left(\sum_{k=1}^{N-1} L(k) + O(I) \right) + L(I) \quad (7)$$

一般にn画面与えられると、2次元の場合と全く同様に2(n-1)個の変換方程式に対し2n個の未知数(2(n-1)個の平行移動成分 $L_x(i), L_y(i)$ と最初の画面における回転軸のX, Y座標)が存在する。よって未知数間に2つの関係式を加えれば記述が一意的に定まる。(4), (5)式の条件を使えば、等速運動で記述され、図8に示す様に、等速運動する軸の回りに物体が回転しているという記述が得られる。4枚以上の画面が与えられれば、より一般的な性質(等加速度運動等)が考えられる。

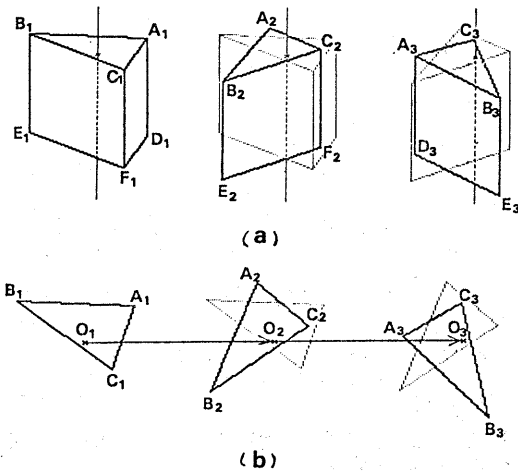


図8. 3画面における3次元運動

4.3 滑らかな性質を持つ運動の記述

前節の議論から、画面数が増えれば増える程、運動の性質はより一般的となり簡潔とは言えなくなる。その故、我々は3画面における平行移動の等速性を最も基本的で簡潔な性質と考えている。そこで4枚以上の画面が与えられた時は、先ず、連続する3画面毎に等速直線運動で回転軸の平行移動を記述する。もし等速運動の成分がシーン全体を通して一定ならば、その物体の運動は1軸回りの回転と求められた等速直線運動であると解釈される。また、等速運動の成分に変化が見られた時は、それらの画面に対してより一般的な性質(等加速度運動, 直線運動, 等曲率運動等)を適用して記述し、以降の画面がその性質を満足しないかを試す。満足されれば、その性質を用いて物体の運動を記述する。用意されたどの性質も満足されない時は、それらの画面間で運動が急激に変化したと解釈し、画像列をそれらの画面の前後に分割し、それぞれで運動を記述する。

図1にこれらの記述法を適用した結果を図9, 10に示す。図9では連続3画面毎に平行移動の等速性を適用して求めた回転軸の最初の画面における座標(X, Y, Z座標系のX, Y座標)を示す。ま

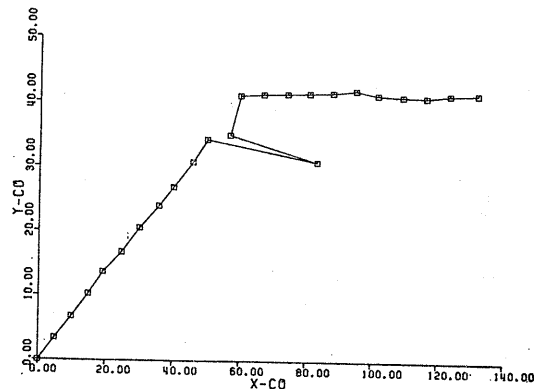


図9. 回転軸のX, Y座標の移り変り
(本文参照)

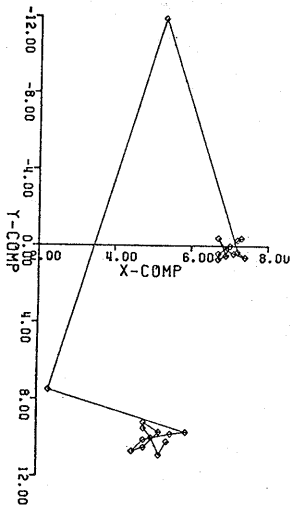


図10. 等速運動成分の構成

それぞれほぼ一定の値を示しており、等速直線運動していると考えられる。そこでより一般的な性質を適用せずに、(m)から(n)にかけて運動が急変したと解釈する。そして画像列を2つの部分(a)から(m)と(n)から(2)に分割し、運動の変化を解釈する。

このようにして、コンピュータは人間の知覚に類似した記述で、ガイナミックシーンの動きを分割し、それらを表現することができる。

5. おわりに

本文では、物体の3次元運動を空間的、時間的に分割し、人間の知覚に似た自然で簡潔な表現を得る手法について述べた。この手法は、理論的な観点から出発しており、実際のシーンを扱っていない。実際のシーンではかなりの雑音があると予想され、幾々の手法が常にうまく行くとは限らないだろう。それ故、2次元の画像から正しい3次元モデルを得る強力な手法の開発が望まれる。

本文では、主部発見の評価基準として、回転軸の方向があまり変化しない事を用いたが、一般にはそうとは限ら

ない。重要な事は各物体の運動がより簡潔に表現できる座標系を発見する事である。その様な座標系を発見する一般的な手法の一つとして、各物体の動きを階層的に記述した時、パラメータの数も最も少くなる様な主部を発見する事が考えられる。

本文では主部以外の物体の主部に対する平行移動や、3次元世界を2次元世界に射影するもう一つの手法、即ち中心射影の場合の平行移動について触れながら、たゞ、同様の議論が可能であり、別稿に譲る。

ここでは、幾つかの滑らかな性質を利用して、その種類やパラメータの値によって運動を分割した。次の段階は、これら分割された動きを、より広い視野から眺めて全体として意味ある記述に解釈する事であり、これからの研究課題である。

参考文献

- [1] W. N. Martin and J. K. Aggarwal, 'Dynamic scene analysis: A survey,' Comput. Graphics and Image Processing, vol.7, no.3, pp. 356-374, 1978.
- [2] H. H. Nagel, 'Analysis technique for image sequences,' in Proc. 4th Int. Joint Conf. On Pattern Recog., Kyoto, Japan: Nov. 1978, pp.186-211.
- [3] 'Special issue on motion and time-varying imagery,' IEEE Trans. Pattern Anal. Machin Intell., vol. PAMI-2, no. 6, 1980.
- [4] M. Yachida, M. Asada and S. Tsuji, 'Automatic motion analysis system of moving objects from the records of natural processes,' in Proc. 4th Int. Joint Conf. Pattern Recognition, Kyoto, Japan: Nov. 1978, pp. 726-730.
- [5] J. Roach and J. K. Aggarwal, 'Computer tracking of objects moving in space,' IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-1, no.2, pp.127-135, 1979.
- [6] S. Ullman, The Interpretation of Visual Motion, Cambridge, MA: MIT Press, 1979.
- [7] J. W. Roach and J. K. Aggarwal, 'Determining the movements of objects from a sequence of images,' IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-2, no.6, pp. 554-562, 1980.
- [8] A. Z. Meiri, 'On monocular perception of 3-D moving objects,' ibid., vol. PAMI-2, no.6, pp. 582-583, 1980.
- [9] M. Asada, M. Yachida and S. Tsuji, 'Three dimensional motion interpretation for the sequence of line drawings,' Proc. 5th Int. Joint Conf. Pattern Recog., pp.1266-1273, 1980
- [10] 浅田,谷田,辻 「三面世界で動く物体の線画の解釈」 信学会, PRL 研究 81-3A