

Shape from Motion を応用した什器類の立体データ作成

Making 3D Data of Utensils Applying “ Shape from Motion ”

中島重義、岡本次郎、濱裕光、細川省一
Shigeyoshi NAKAJIMA, Jiro OKAMOTO, Hiromitsu HAMA,
Shouichi HOSOKAWA

大阪市立大学工学部、〒 558 大阪市住吉区杉本 3-3-138
Faculty of Engineering, Osaka City Univ.,
3-3-138 Sugimoto Sumiyoshi-ku Osaka City, 558

キーワード: Shape from Motion, 立体データ, 回転運動, 動画画像での手法を静止画に応用する。

Keyword : Shape from Motion, 3-D data, rotation

あらまし:

人間の生活で使用する什器類は、民族学、考古学などの貴重な資料とされるものである。什器類の簡便で可搬性のある立体データ作成システムがあれば、人文科学の立体データベース構築に大きな貢献が出来るものと思われる。

筆者らは動画画像から回転剛体の運動パラメータと特徴点の3次元的位置の復元する手法を発表している。これを応用して、市販の一般的な器材を利用した什器類の立体データ作成システムを提案する。普通の動画画像とは、通常は

例えばビデオテープに撮った映像などを指すものであるが、本発表では静止画の写真を用いる。近年、アクティブ・ビジョンと呼ばれる分野では、機械によって撮像装置の位置や方向を能動的に操作してシーンの3次元立体

構造を理解する研究が進んでいる。しかし、広範なデータベースを構築するには、特殊な装置を必要とするシステムよりも、汎用器材を人力で操作するシステムが適当であると思われる。人力操作による誤差は、多くのフレ

ームショットを利用することで改善される。静止画を連続する動画画像のフレームと等価に扱うこ

Summary : Utensils used in human daily life are important materials in ethnology or archeology or other fields. If there is a simple portable system to make 3-D data of them, it might be contributive to culture science 3-D data base. Recently we proposed a method to recover 3-D rotation parameters of the rigid bodies and 3-D positions of feature points from a moving image. We propose a system applying it to make 3-D data of utensils using ordinary machines. For example, a video tape image is one of moving images. But in this paper, we use still images. Recently in active vision field, there are advanced works which recognize 3-D structures controlling camera calibrations in active manner. But the system in which a man control ordinary machines are assumed better than the one which needs special machines. The error from man control are corrected by the number of frame shots. We apply the method for moving images to still images dealing them as consequent frames of moving ones.

1 はじめに

人間が使用する什器などの器物の3次元形状を得てデータベースを作成するために、本論文では「動きからの形状回復 (Shape from motion)」を応用して、人力でアクティブに対象物を回転させて、通常の写真機で撮影する手法を提案する。

第2節では、コンピュータ・ビジョンの分野における3次元の立体形状回復の歴史的な流れを述べる。第3節では、什器などの器物の立体形状のデータベースの必要性について述べる。第4節では、提案する立体形状回復の手法について述べて、第5節において考察する。

2 3次元の立体形状

2.1 3次元の立体形状回復

コンピュータで、風景や写真から情報を得る分野はコンピュータ・ビジョン (computer vision, CV) と呼ばれる。特にその中でも、物体やランドスケープの3次元の形状を捉えることは、ロボットの動作や、自動走行車の開発からも重要視されてきた。「コンピュータビジョン」^[1]では以下の研究例があげられている。

1. 濃淡からの形状回復
2. 距離データからの形状回復
3. 形状のモデルを用いる手法
4. 動画像処理

濃淡からの形状再構成の例としては Shape from shading などがある。Shape from shading は Horn^[2] が提唱したもので、よくわかる事例として石膏塑像の写真が挙げられる。石膏塑像の表面は同じ色で、光に対する方向が違うだけである。石膏塑像の表面が滑らかであるとき、そのまわりの反射の変化などでその表面の向きがわかり、同時にその表面の相対的な奥行が計算できる。

距離データからの形状再構成とは、例えば、両眼立体視法がある。人間を代表として、顔の前面に複数の眼を持つ類人猿、フクロウなどの猛禽類のような動物は、自分の両眼の距離と、対象物が両眼で異なった角度に写る両眼視差とを手掛りに

して対象物と自分との絶対的距離を測る。対象物毎の距離の違いから、対象物間の相対的距離なども判定する。距離データからの形状再構成には、その他に投光法がある。これは、ピンスポット光、スリット光等を当てて、光の方向から斜めに見ると、光があたっている所の見え方から光のあたった部分の絶対的な奥行が測定できる。

形状のモデルを用いる手法は古くは Roberts^[3]の研究などがある。彼は既知の直方体、三角柱等からなる輪郭線でできた映像を解釈させることに成功した。Kanade^[4]は厚さの無い折り紙世界の映像の解釈を行った。

動画像を処理することで動きからの形状復元ができる。動きからの形状復元は Shape from motion と呼ばれる。これは、上で述べた手法が静止画を対象とするのに対して、動画像を対象とする。対象画像から、この分野では動いている間に物体の形状が変化しないという、「剛体仮説」を前提とするのが普通である。

2.2 動きからの形状復元

テレビ、ビデオ、映画等を見ていると、スチル写真を見ている状態に比べてより奥行感が得られることに気付く。これは、物体が動くことによって、静止画よりも動画がより多くの情報を含むことになるからである。例えば、見ている人に近づいて来るものは見かけの大きさがどんどん大きくなっていく。反対に遠ざかるものは見かけの大きさがどんどん小さくなっていく。このように、回転や変形などせずに一定の方向に物体が進む場合は、その物体上の各点の動きは遠近法では一点に集まり、その点は Focus of Expansion(FOE) と呼ばれる。

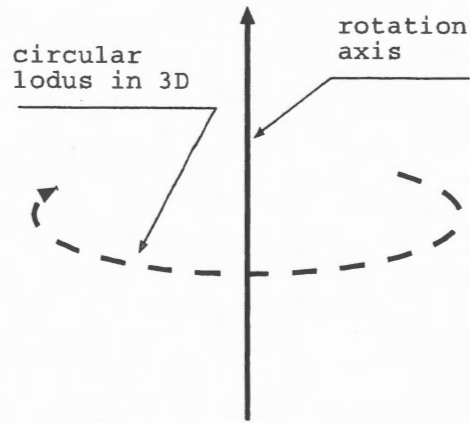


図1 3次元空間での回転運動

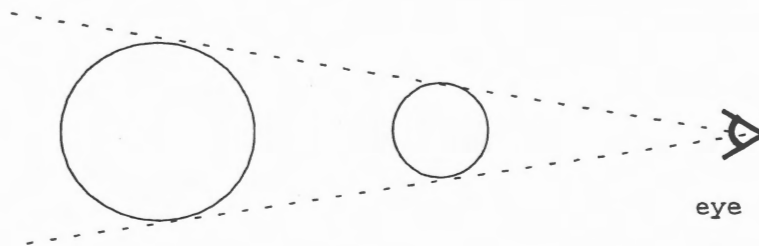


図2 単眼視での見かけの大きさ

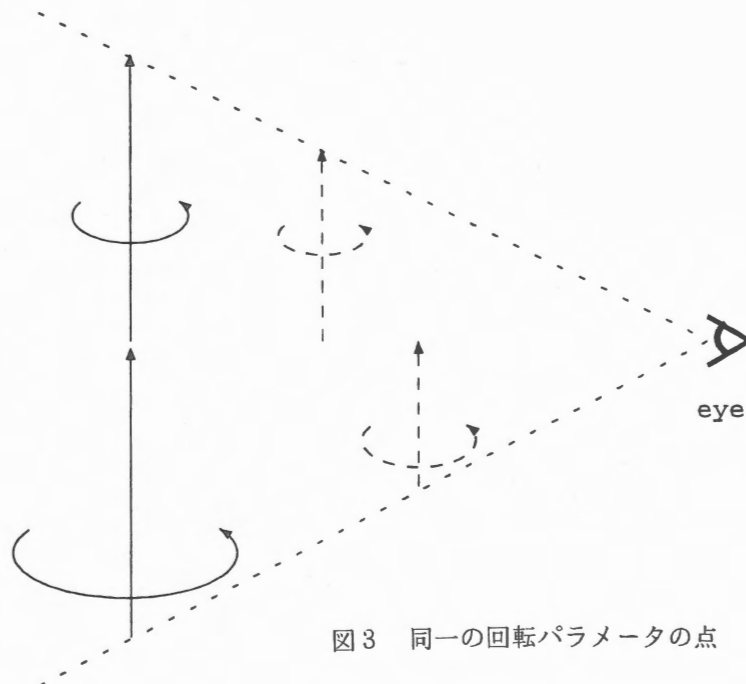


図3 同一の回転パラメータの点

物体が回転している場合については Sawhney ら^[5]が画像平面上の円錐曲線を利用した研究をしている。Sawhney らによれば、一定の軸のまわりを回転する点の軌跡は円を描く。この円形の軌跡は、遠近法によって画像に投影されるとその位置関係によって円錐曲線である、楕円、放物線、双曲線のいずれかになる。このとき、楕円になるのは、円形の軌跡が全て観察者の前方にある場合であり、放物線となるのは、円形の軌跡の一点が観察者の真横に来る場合であり、双曲線になるのは、円形の軌跡の一部が観察者の背後にある場合である。Sawhney らの実験では、回転する立方体の表面に黒い正方形が描かれ、その正方形の縁が直線として検出された。検出された直線から「角検出」と呼ばれる処理がなされ、正方形の4隅が1枚の画像上で得られる。このような画像が連続して何枚も動画から得られ、それらの位置を $x-y$ 画像平面上にプロットして、円錐曲線を形成する点列をグループ化する。グループ化された点列から得られた円錐曲線のパラメータは回転軸の向き、回転角速度などを与え、観測された点の相対的位置関係を与える。

ここで、観測された点の位置には大きさ因子 (scale factor) と呼ばれる値が定まっていないことに注意する必要がある。これは、画像が一台のカメラから撮影されたことに原因がある。片目で物を見ると遠近感が無くなるという経験が誰しもある。これは 1m 離れた所にある 20cm のボールと 2m 離れた所にある 40cm のボールは、普通、片目で見ると区別がつかない。現実では、置かれている場所などで判断がつく場合もあるが、そのような都合のいい場合ばかりとは限らない。これは、片目を覆ってテニスなどの球技をしてみるとよくわかる。それでは、観測された点の相対的位置関係はどのようにして得られるのかということ、同じ回転軸のパラメータを持った複数の軌跡を1個の物体に属するものとしてさらにグループ化して、回転軸までの絶対的距離が同じであると仮定するのである。回転軸までの絶対的距離が決まれば、各点と回転軸との位置関係、および各点間の位置関係が決まり、物体の形状が回復される。

筆者らは、Sawhney の円錐曲線による回転運動パラメータの回復をさらに発展させて、時空間内の螺旋軌跡による回転運動パラメータの回復を

行った^[6]。画像上の円錐曲線のかわりに時空間内の螺旋軌跡を用いることで軌跡の分離がより容易になった。また、この螺旋軌跡を途中で低次元の図形に近似することで、パラメータ回復の速度が向上した。

3 器物の立体形状データベース

3.1 立体形状作成

近年、計算機やネットワークの資源が豊富になってきた。美術館、博物館などでは、所蔵物の画像をホームページで公開するところもある。しかし、現在のところ、2次元的な画像を利用したものは多いが、3次元的なデータを利用したものが見当たらない。

日常生活で使用される器物、什器などのデータは、2次元的な画像だけではなく、3次元的な立体形状のデータを得て利用することは有用なことと思われる。しかし、3次元形状のデータを広範な対象から得ることは、現在では非常に困難なことである。それには、3次元形状を測定するための手間、コストがかかることが原因であると考えられる。

例えば、距離データからの形状回復を3次元形状の測定に応用することを考える。両眼を持つ生物には容易なことでも、機械で自動化するときには困難であるのは、右眼で見た風景のある部分が左眼で見た風景のどこに対応するかという対応問題である。この対応問題が正確に解けるならば、距離データからの形状回復の手法は人間並み、あるいはそれ以上の能力を持って対象物の3次元構造を得ることができる。しかし、実際には、画像の対応問題を自動的に正確に解くには人工知能の常識への対応能力の限界がある。また、対応問題の解を人間が与えるのは手間がかかる。

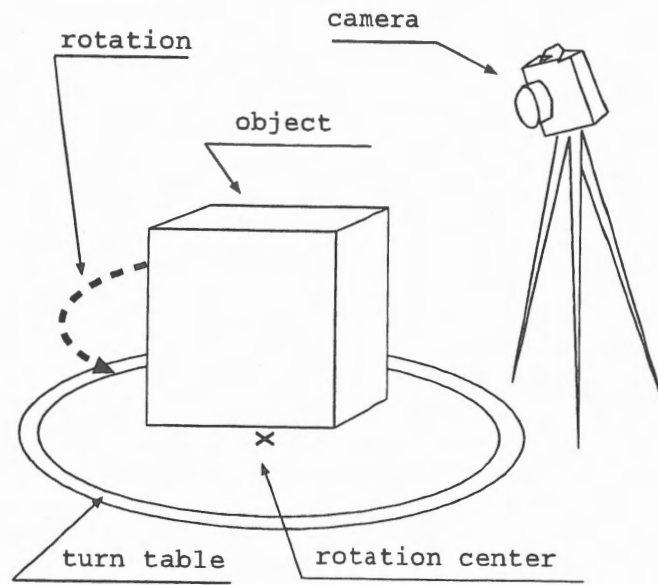


図4 ターンテーブルとカメラ

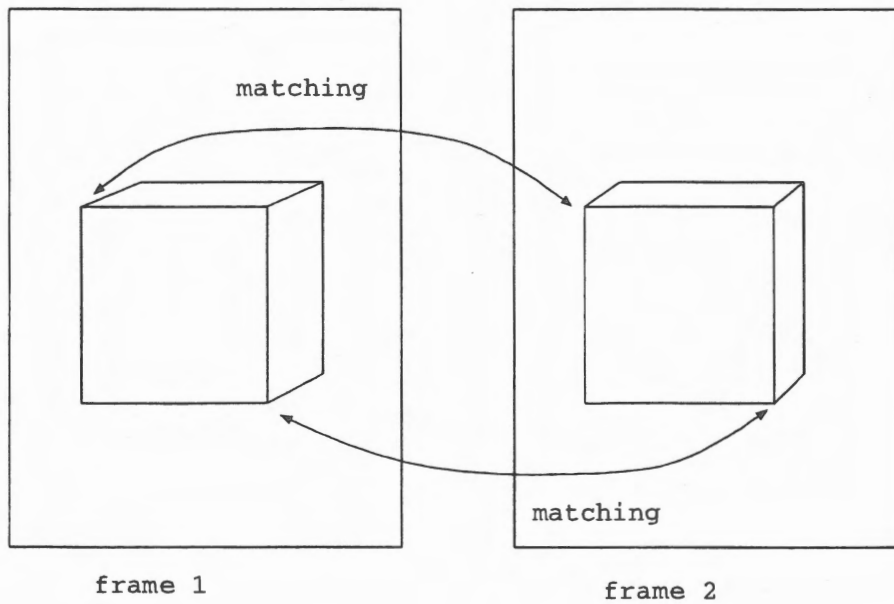


図5 連続フレームの対応点

動きからの形状回復のときも、対応問題に関わるが、動画像の場合、対応点が近接しているということ、多くの特徴点の動きを満足させる運動パラメータがクラスタリングなどで得られることなどが利点である。物体の動きをカメラが受動的に受ける場合と、カメラが能動的に動く場合がある。後者の場合がアクティヴビジョンと呼ばれるものである。本論文での発表は、物体の動きを回転運動に限定してカメラが受動的に画像を受け取るものであるが、物体とカメラの位置関係を人間の手で制御するもので、人手によるアクティヴビジョンとも考えることができる。

3.2 立体形状データベースの利用

立体形状データベースの利用用途としては、立体形状のマッチングによって、形状の類似した什器、器物の分類が考えられる。また、立体形状データに2次元的な表面を張り付けて3次元画像を作成し、ネットワーク等で展示することも考えられる。

4 立体画像作成手順

4.1 装置

装置は、回転するターンテーブルと通常の静止画写真カメラを用いる。ターンテーブルは、人力で回転し、回転角度が決められるようになっている。ターンテーブルには撮影対象を置く。ターンテーブルの詳細は撮影対象の大きさ、重量によって異なる。

4.2 静止画撮影

写真カメラは固定しておく。ターンテーブルを等しい角度間隔で回転し、カメラで対象のほぼ全体を撮影する。次のような注意点がある。

1. 角度間隔は任意であるが、大きすぎるとはいけない。 $\pi/8$ 程度を推奨する。
2. ターンテーブル上の撮影対象の位置は任意であるが、回転しても対象の全体がフレームに入ること。
3. 上の条件を満たしながら対象が大きく写る方がいい。

4. カメラの左右は水平であること。
5. ターンテーブルの回転中心はカメラの視野の中心にあること。
6. カメラの俯角を測定しておくこと。
7. ターンテーブルの回転中心の位置がわかる写真を撮影しておくこと。
8. カメラの位置と回転中心の距離を測定しておくこと。

4.3 連続フレームからのマッチング検出

撮影された画像は連続したフレーム(続いて移した写真)の間で計算機処理によりマッチングがとられる。マッチングを取るとは、写真の間で位置の対応付けをすることであり、対応問題を解くことになる。Broidaら^[7]の例では、自動車のタイヤの側面に明るい色の小さなシールを張り付け、その位置を画像処理で検出した。Sawhneyら^[5]の実験例では明るい色の直方体の側面に正方形の黒い紙片を貼り付け、その4隅を、「角検出器(corner detector)」と呼ばれるアルゴリズムで検出した。広範な対象に対してマッチングをとるためには、様々な什器に対する適切な画像処理とマッチングアルゴリズムがどのようなものであるかを検討しなければならないが、それは今後の課題である。本論文では、2次元相関を求めることで対応点を定める。2次元相関とは画面の小領域を重ね合わせて、明るいところどうし、暗いところ同士が重なれば高い数値を与えるような評価方法である。このようにして、計算機により自動的に隣り合うフレーム間で対応する点を見つける。しかし、このような対応点を見つけても、そこには誤りがおきる可能性がある。例えば、対応している点であっても、回転による見かけ上の変形があり、それほど相関が高くないことがある。また、見かけが似ている点と同じ画像の中であればその相関が高くなることもある。このように、相関の値だけを見ていると、正しい対応と誤った対応が混在することになるが、それらの対応の中で正しい螺旋軌跡を形成する対応の組を選べば、正しい対応点が得られる。誤った対応が偶然に螺旋軌跡を形成しないためにフレーム数は多いほうがいい。

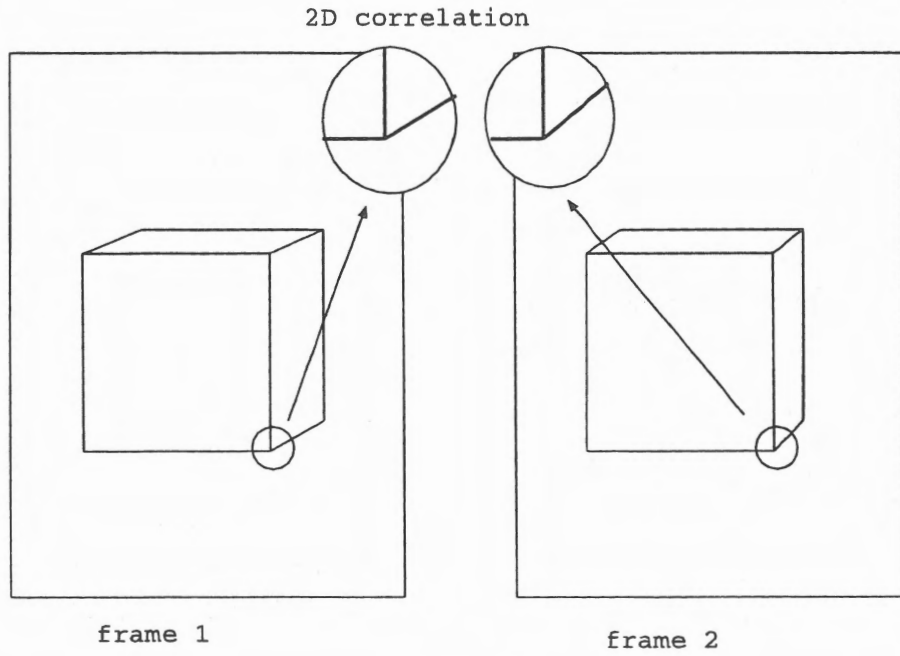


図6 2次元相関

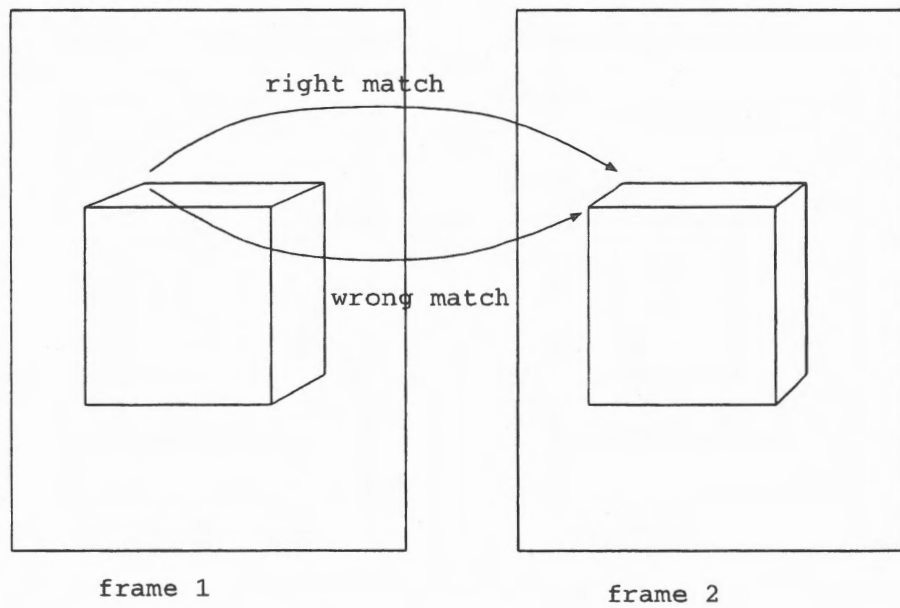


図7 正しい対応と誤った対応

4.4 螺旋軌跡の形成

対応する点は螺旋軌跡にグループ化されるが、低次元図形から段階的にグループ化される。対応する点列は、時空間内で部分的な直線で近似される。直線が最も低次元な近似図形にあたる。これらの部分的な直線は、さらに、時空間中の円、時空間中の楕円に近似される。低次元図形のパラメータから、高次元図形のパラメータが推定され、推定された値から漸近的にパラメータを修正して、時空間中の点列に適合するパラメータが求められる。時空間中の楕円から画像上の楕円のパラメータが推定され、それから螺旋軌跡が推定される。推定された螺旋軌跡をもとに漸近的にパラメータを変えていくと最も良い螺旋パラメータが得られる。こうして、各段階でグループ化とパラメータの推定と修正を経て、誤った対応が除外され、正しい対応によるパラメータが得られる。複数の点に対してそれぞれの螺旋のパラメータが得られたならば、それらの点間の相対的位置が回転軸を規準に与えられることになり、それらの点による形状が得られることになる。

5 考察

5.1 本手法の利点

人力アクティブビジョンともいえる本手法は、器材として回転角度がわかるターンテーブルと、通常の写真機を用意するだけなので、非常に簡便であり、広範囲の協力を得ることも可能である。また、人力の制御による誤差は計算機による漸近的な変化で修正される。

5.2 螺旋軌跡の形成の問題点

前に述べたように、画像処理をどうするかというのが問題である。単なる2次元相関でどこまでできるのか、角検出器などの処理がいいのか、実際に撮影された画像で試みる必要がある。また、撮像の諸条件である、角度変化はいくらぐらいが適当か、カメラの俯角はどうか、適切な撮影フレーム数なども検討課題である。今後、このシステムによるデータを集積して検討する必要がある。

また、本手法では画像上の2次元相関でマッチングをとるため、画像上に相関をとるにたる特徴

のない場合、例えば白磁の無地の部分などには適用できない。

ここまでは、ターンテーブルに対象物を載せて回転させる手法を説明したが、対象物が重すぎて動かせない、あるいはケースの中から取り出すことが許されないなどの理由でそれが不可能な場合は、対象のまわりでカメラの位置を操作して、回転運動を作り出す手法が考えられる。

6 おわりに

本論文では、什器などの器物の立体データを得るための手法を提案した。人力で操作することの簡便さと、人力による誤差を計算機で修正することによる利点があるが、問題点や課題も残る。今後はそのような問題点や課題の検討を行いたい。

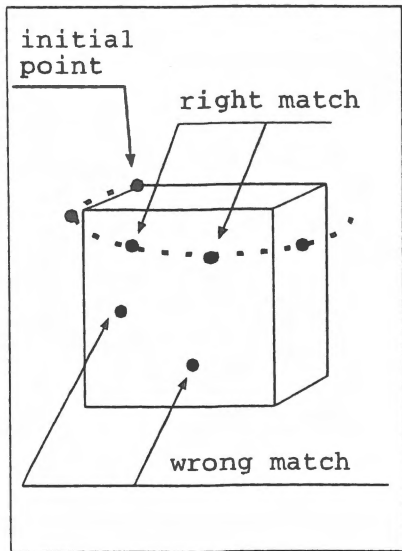


図8 回転と正しい対応点

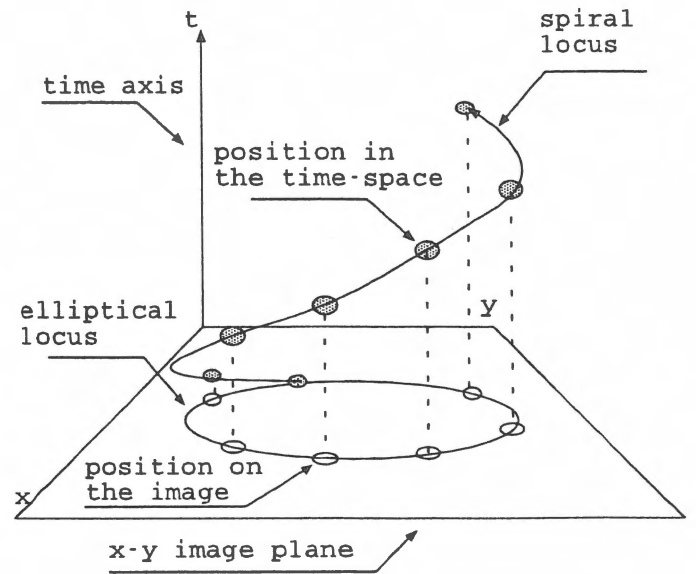


図9 時空間での螺旋軌跡

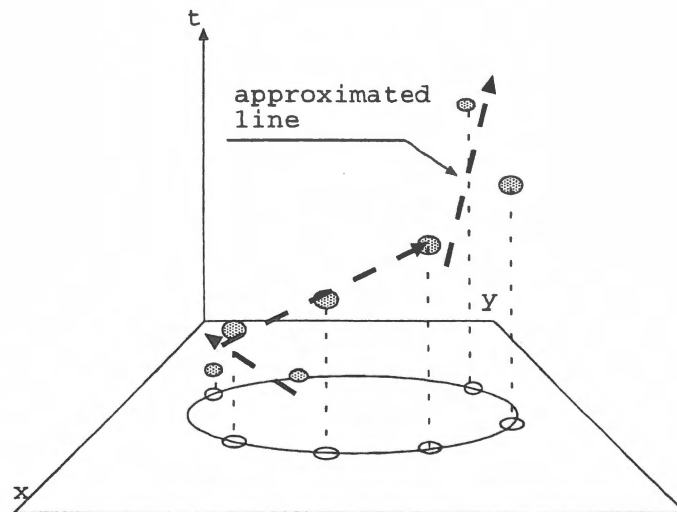


図10 時空間での直線近似

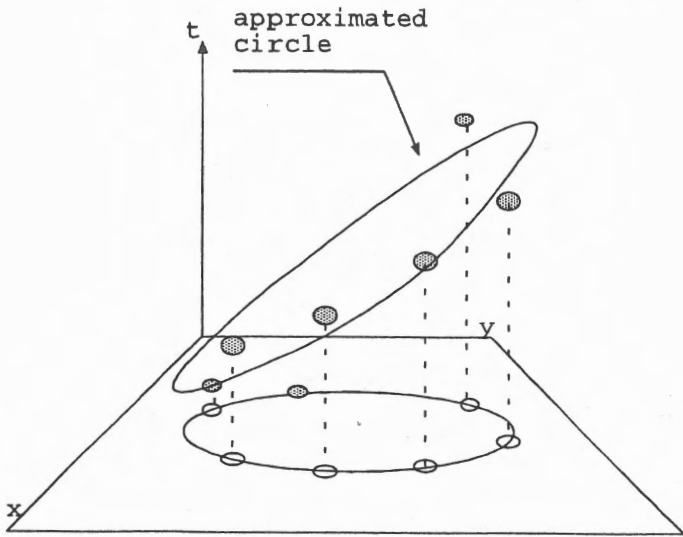


図 1 1 時空間での円近似

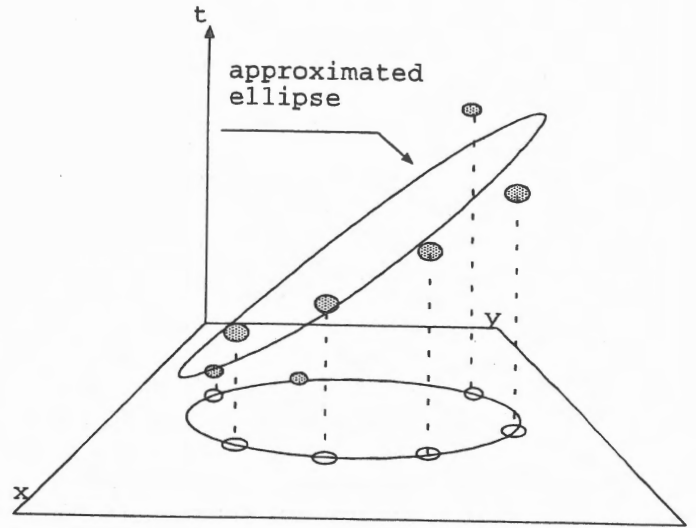


図 1 2 時空間での楕円近似

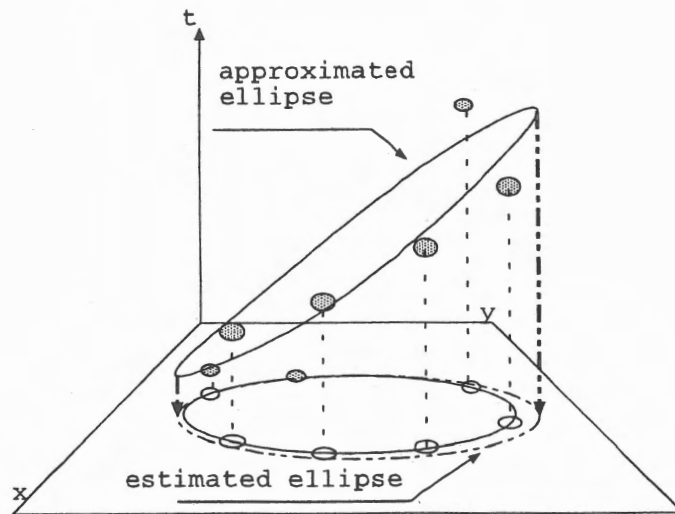


図 1 3 x-y 平面での楕円推定

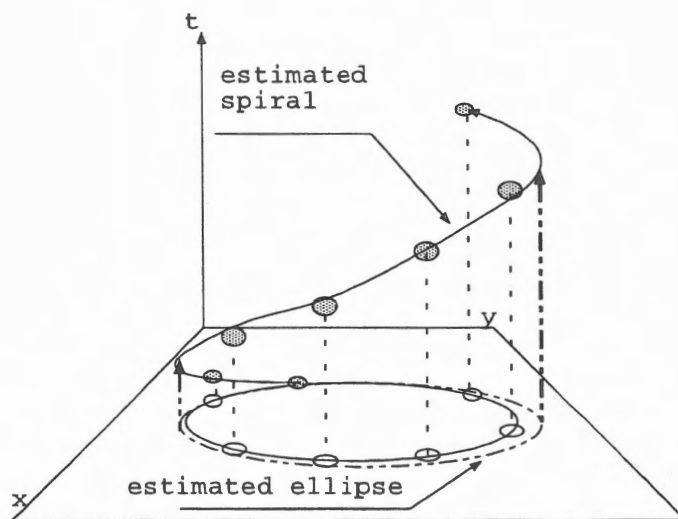


図14 時空間での螺旋推定

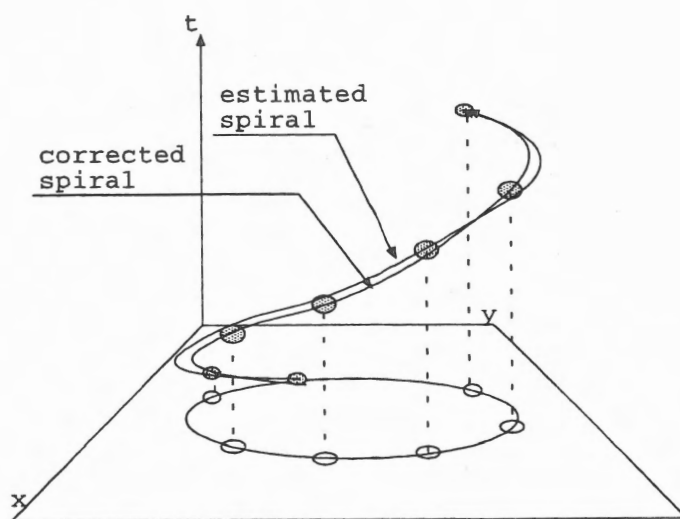


図15 時空間での螺旋修正

参考文献

- [1] 谷内田正彦 編：“コンピュータビジョン,” 丸善, Feb. 1990.
- [2] B. K. P. Horn: “ Obtaining shape from shading information, ” *The pshychology of computer vision*, P. H. Winston, Ed., New York: McGraw-hill, 1975.
- [3] L. Roberts : “ Machine perception of three-dimensional solids, ” *Optical and electro-optical information processing*, ”J. Tippett Ed., 1965.
- [4] T. Kanade : “ Recovery of the three-dimensional shape of an object from a single view, ” Rep. Np.CMU-CS-79-153, Comp@uter Science Dept., Carnegie- Melon Univ., 1979.
- [5] H. S. Sawhney et. al. : “Image Description and 3-D Reconstruction from Image Trajectories of Rotational Motion, ” *IEEE Trans. Pattern Analysis and Machine Intelligence*, 15 , 9, pp. 885-898(1993)
- [6] 中島重義, 濱裕光, 細川省一 : “ Hough 変換の多段階化による複数剛体の運動回復 , ” テレビジョン学会誌, 50,7,pp933-942,1996.
- [7] Broida, T. J. and Chellapa, R. : “Estimation of Object Motion Parameters from Noisy Images”, *IEEE Trans., PAMI-12*, 11, pp.1092-1098 (1990).