

点対応を用いた 複数の 2 次元画像からの 3 次元形状復元 因子分解法の数理

産業技術総合研究所* 藤 木 淳

(受付 2000 年 10 月 18 日 ; 改訂 2001 年 1 月 18 日)

要 旨

複数の 2 次元画像からカメラ運動と対象物体の立体形状を同時に復元する問題はコンピュータビジョンにおいて基本的かつ重要な問題であり, その中でも点特徴の対応に基づいた複数の 2 次元画像からのカメラ運動と対象物体の 3 次元復元問題はもっとも基本的かつ重要な問題である. この問題を解決する手法の中でも因子分解法は実際のカメラモデルである透視射影をアフィン射影で近似することにより問題を簡略化することによって数値計算上安定でかつ比較的良好な結果を与える優れた手法である. 因子分解法は手法として優れているだけでなく, 複数のアフィン近似射影画像からのカメラ運動と対象物体の立体形状を同時に復元する問題を理解する上で非常に有用な方法である. 本稿では, 因子分解法を通して点対応を用いた複数のアフィン近似射影画像からのカメラ運動と立体形状の復元の数理について解説する. また, 透視射影による像からアフィン近似射影による像を推定することによって, 透視射影画像からのカメラ運動と立体形状を同時に復元する手法及び逐次型因子分解法についても紹介する.

キーワード : Structure from motion, 因子分解法, 逐次型, 計量アフィン射影, 透視射影.

1. はじめに

1.1 コンピュータビジョン

現在, コンピュータに人間と同等の機能を持たせるための研究が盛んに行なわれており, その中でも人間の視覚認識の機能をコンピュータを用いて実現しようとする研究はコンピュータビジョンと呼ばれている.

人間の視覚認識の機能の大きな特徴は, 3 次元空間を 2 次元画像の情報によって認識・理解していることであり, コンピュータビジョンの目的も 2 次元画像からコンピュータを用いて 3 次元空間を認識・理解するところにある.

2 次元画像の情報から 3 次元空間を認識・理解するときに問題となるのは, 3 次元空間を 2 次元画像に投影するときに 1 次元分の情報が不足することである. そのため, 同じ物体でも異なる視点から得られる画像では全く異なる形状が観測されたり, 実際に異なる物体でも同じ形

*脳神経情報研究部門 : 〒 305-8568 茨城県つくば市梅園 1-1-1 産総研つくば中央第 2.

状に投影されたりし、3次元空間の情報に不定性が生じてしまう。また、3次元空間においては遠く離れた点であっても2次元画像においては近くに投影されることがあることから、2次元画像におけるわずかな誤差が3次元空間における認識・理解において重大な影響を及ぼすこともありうる。このように不定性や誤差が存在する条件のもとで3次元空間を認識・理解することがコンピュータビジョンの課題である。

さて、2次元画像上から得られる3次元空間の認識・理解を助ける情報としては、点、線、面、輪郭、陰影、色彩などの特徴量が考えられる。そして動画像、つまり時間において連続な画像が得られれば、それらの2次元画像上の特徴量のみかけの運動情報が得られる。そこで、これらの2次元情報を利用して3次元空間を認識・理解するための様々な手法が提案されてきた。これらの手法は光学モデルなどの物理的知識や、カメラによって3次元空間がどのように2次元画像へと投影されるか、カメラの位置の変化によって物体の見え方がどのように変化するかなどの幾何学的知識を利用し、数理的最適化手法や統計的推定手法を駆使して3次元空間を認識・理解するというものであり、コンピュータビジョンによって初めて視覚認識研究が学問として確立されたという見方が定着した。

これらのコンピュータビジョンに関する手法は理論的な完成度が高かった反面、現実世界の問題解決には役に立たないという批判もあった。というのも、肝心の2次元画像から線、面、輪郭等の特徴量を抽出したり、動画像において対応している特徴量、つまり3次元空間において同一である特徴量を抽出する手法は、いまだに画像処理の中心的なテーマであり、必要な特徴量を正確に安定して抽出することのできる決定打とも言うべき手法はまだないと言っても過言ではないからである。そこで最近では多様な現実世界の状況においても頑健に動作するアルゴリズムも提案されてきている。しかしながら、多様な現実世界の状況において頑健に動作するアルゴリズムもその基本は多くの仮定を必要とする理想的な状況における理論的にしっかりとしたアルゴリズムにあることは間違いなく、理想的な状況においてでしか機能しないとしてもコンピュータビジョンに関する理論的研究を理解しておくことは必要である。

なお、コンピュータビジョンという用語は、用いられた当初は主に2次元画像から3次元空間の物体の形状や照明の位置などの3次元空間の構造の復元を指していたが、現在はより広い視覚情報処理研究も内包することとなる。例えば、2次元画像を利用した監視システムや人物特定などの知能システムの構成、知覚と行動の統合を目指すアクティブビジョンなどがある。アクティブビジョンとは、3次元空間を受動的に観測するのではなく、カメラや照明を能動的に制御し、撮像、認識・理解、環境の制御を繰り返して複雑な3次元空間を認識することである。他にも生の画像に基づいて仮想化された画像・映像を生成する Image-Based Rendering, Virtualized Reality, Augmented Reality(拡張現実感), Mixed Reality(複合現実感)など、マルチメディア情報処理の基本技術もコンピュータビジョンという分野に含まれ、そしてコンピュータビジョンにおける中心的な研究テーマとなっている。

1.2 点对応を用いた複数の2次元画像からの3次元形状の復元

コンピュータビジョン、すなわち2次元画像上の情報からの3次元空間の認識・理解において、複数の画像における物体の像からその物体の3次元形状及びカメラ位置を同時に復元する問題は structure from motion と呼ばれ、マンマシンインターフェイス、複合現実感を伴うシステム、自律歩行ロボットなど様々な応用が考えられるもっとも基本的かつ重要なテーマである。3次元空間の物体は、点、直線、面等のさまざまな特徴量から構成されているが、その中でもっとも基本的な特徴量は点であることから、structure from motion においてもっとも基本的な問題は複数の画像における点の対応の組から、それらの点及びカメラの3次元空間における位置関係を復元することである。

前節で述べたように、3次元空間においては遠く離れた点であっても2次元画像においては近くに投影されることがある。これは2次元画像におけるわずかな誤差が3次元空間における認識・理解において重大な影響を及ぼすこともあることを意味する。そのため、複数の画像における点対応、つまり3次元空間において同一の点であるという対応付けを精度良く行なうことが必要である。

そこで、複数の画像間の点特徴の対応付けに関するさまざまな研究が行なわれてきた。時間的に近接した画像間において対応する点は、画像における位置も近いと考えられることを利用して特徴点を追跡(トラッキング)することにより、画像間における点対応をつける手法が多く提案されている。その中でも Kanade-Lucas-Tomasi トラッカー (Shi and Tomasi (1994)) が優れた手法として有名である。時間的に離れた画像において事前知識なしに対応点を求めるのは難しいが、複数の点対応が既知であれば、エピポーラ幾何学 (Xu and Zhang (1996)) と呼ばれる、同一の3次元空間を撮影した視点の異なった複数の画像間に存在する幾何学的関係を用いて他の対応点の存在可能領域を絞り込むことが可能である。しかし、まだ高精度に安定して点特徴の対応付けを行なう手法はないと言っても過言ではなく、点特徴の対応付けを高精度に安定して行なう研究もコンピュータビジョンにおいて重要な研究テーマとなっている。

それでは複数の画像間における点対応が求められた後はどのようにすれば特徴点の3次元的位置関係が求められ、そしてそれら特徴点をどのように撮影すれば各画像が得られるだろうか。(2次元画像は対象物体とカメラの相対的な位置関係のみによって決定されるため、複数の画像を扱うときには問題を単純化するために対象物体を固定するかカメラを固定するかのいずれかを仮定しても一般性を失わない。本稿で解説する因子分解法は対象物体を固定するという立場をとり、このときに得られるカメラの位置の軌跡をカメラ運動と呼ぶ。)

それを理解するにはまず、3次元空間がどのように2次元画像に投影されるかというカメラモデルについて考慮する必要がある。

理想的なカメラモデルであるピンホールカメラは図1左のようになっている。ピンホールカメラにおいては対象物の上下が逆転して投影されるので、各画素にカラーまたはモノクロの輝度値を与えるカメラの受光部を構成する撮像素子が置かれる面(画像面; image plane)をピンホール(以降、カメラ中心; camera center と呼ぶ)の前に置くと図1右のように対象物の上下は逆転せずに投影される。以降、画像面はカメラ中心の前にあるものとする。

このカメラモデルを表現するためには3次元空間を表現する座標と、3次元空間を投影して得られる2次元画像の座標をどのように選べば良いだろうか。もちろん、コンピュータビジョンで扱う2次元画像は最終的にはコンピュータのディスプレイに表示される2次元に配列された画素から構成される座標(画素座標; pixel coordinate)によって表現される。しかし、カメラによって画素の数や配列の仕方が異なるため、カメラモデルを直接3次元から各画素への写像と考えると、カメラによって写像が異なることになり、すべてのカメラに対して統一的に議論することが難しくなる。そこで、すべてのカメラに対して統一的に議論が可能ないように、画素

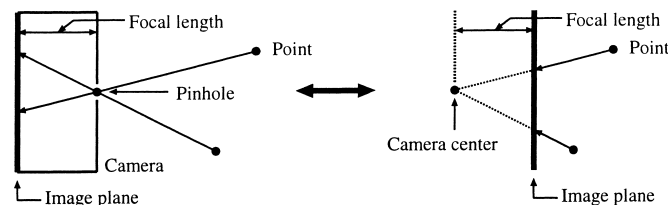


図1. ピンホールカメラ.

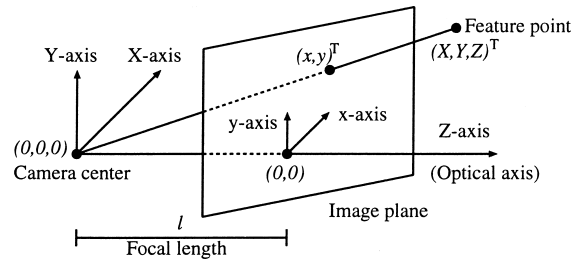


図 2. 座標系の設定.

座標よりも都合の良い座標系を用いることにする．それは，ピンホールカメラを表現する透視射影の表現が簡単になる座標系であり，具体的には，カメラを基準にした 3 次元空間を表現する座標であるカメラ座標 (camera coordinate) 及び 2 次元画像を表現する座標である画像座標 (image coordinate) を次のように設定する(図 2)ことによって得られる．

まず，カメラ座標系 (XYZ 座標系とする) の原点をピンホールに相当するカメラ中心とし，画像面へカメラ中心から下した垂線 (光軸; optical axis) 方向が Z 軸となるように正規直交座標系を定める．このとき， X 軸， Y 軸は画像面と平行になる．次に画像座標系 (xy 座標系とする) の原点を光軸と画像面の交点とし， x 軸， y 軸がそれぞれ X 軸， Y 軸と平行になるように定める．ここで x 軸， y 軸上の単位長は X 軸， Y 軸上の単位長と等しいものとする．

このようにカメラ座標系と画像座標系を設定するとカメラ座標が $(X, Y, Z)^T$ である 3 次元空間の点と，その透視射影として得られる画像座標 $(x, y)^T$ の間には

$$(1.1) \quad x = l \frac{X}{Z}, \quad y = l \frac{Y}{Z}$$

という簡単な関係式が成立する．ここで l はカメラの焦点距離 (focal length) である．

ここで，コンピュータのディスプレイに表示される画像を記述する画素座標と透視射影による 3 次元空間の像を記述する画像座標との間には，個々のカメラに関して固有である一対一の関係が成立するので，個々のカメラに対してこの写像を求めることができれば，画素座標と画像座標を自在に変換することができるので，複数の画像における点対応からのカメラ運動と点の相対的位置関係 (以下，立体形状と呼ぶことにする) の復元を考えると，点対応が画像座標で与えられるとして良い．なお，画素座標と画像座標の間の一対一写像を求めることをカメラの校正またはカメラキャリブレーションと呼び，この写像をどのように求めるかについてもコンピュータビジョンにおいて盛んに研究されている．

さて，理想的なカメラモデルであるピンホールカメラを表現する透視射影の枠組の中で複数の画像における点対応からのカメラ運動と立体形状の復元に関する理論的な研究が行われてきたが，式 (1.1) から分かるように透視射影は非線型写像であるから，透視射影における複数の 2 次元画像からのカメラ運動と 3 次元形状を復元する問題は非線型写像の逆問題となり，非線型最適化問題に帰着される．一般に非線型最適化問題はノイズに敏感で初期値依存性が高く数値計算上不安定であるという問題があるため，理想的なカメラにおいて安定して 3 次元形状を復元するのは難しい．

そこで理想的なカメラモデルである透視射影をアフィン射影に近似した，正射影モデルなどのアフィン近似射影に基づいた複数の 2 次元画像からの 3 次元形状復元手法が提案されている．2 次元画像がアフィン近似射影で得られると仮定した場合，複数のアフィン近似射影画像における点対応からのカメラ運動と立体形状の復元問題は線型写像の逆問題となり，非線型写像の

逆問題に比べて数値計算上安定して解くことができるからである。もちろん、実際の画像は理想的なカメラモデルである透視射影で得られるため、アフィン近似射影と仮定して復元した場合の復元精度には限界があるが、高速かつ数値計算上安定して復元できるという利点がある。また、アフィン近似射影に基づいた復元解は透視射影をカメラモデルとした場合の非線型最適化問題の初期値として利用できる。そこで、アフィン近似射影における複数の 2 次元画像からの 3 次元形状復元に関する理論的研究が行なわれた。

透視射影における複数の 2 次元画像からの 3 次元形状復元に関する理論に関しては金谷 (1990) などに詳しい。本稿においては透視射影のアフィン近似射影のもとでの複数の画像における点対応からの 3 次元形状復元問題について解説する。アフィン近似射影のもとでの点対応からの 3 次元形状復元問題について考察するには Tomasi and Kanade (1992) によって正射影モデルの場合について提案された因子分解法に基づくのが理解しやすいので本稿では因子分解法を利用してアフィン近似射影のもとでの点対応からの 3 次元形状復元問題について考察する。

1.3 本稿のあらすじ

本稿では、点対応を用いた複数のアフィン近似射影画像からの 3 次元復元について因子分解法に基づいて解説する。本稿では複数の画像における点特徴の対応は既に求められており、それらは画像座標として与えられている、つまりカメラは校正されているものとする。

アフィン近似射影においては、カメラで撮影するという写像是 3 次元空間の対象物体から 2 次元画像へのカメラの位置と方向によって定まるアフィン射影となる。今、画像が F 枚、特徴点が P 個与えられたとき、 P 個の 3 次元座標の F 個のアフィン射影による FP 個の画像座標が得られる。因子分解法では、これら FP 個の条件を行列の形に並べることによって複数の 2 次元画像からの 3 次元形状復元問題を単純な形で表現した。具体的には、 FP 個の画像座標を並べてできる $2F \times P$ 行列である計測行列、 F 個のアフィン射影の表現行列を並べてできる $2F \times 3$ 行列である運動行列、 P 個の特徴点の 3 次元座標を並べてできる $3 \times P$ 行列である形状行列の間に

$$(\text{計測行列}) = (\text{運動行列}) \times (\text{形状行列})$$

という関係式が成立することから、複数の 2 次元画像からの 3 次元形状復元問題を計測行列の分解に帰着させた。

このとき、実際に得られる計測行列の各成分は、アフィン近似射影によって投影された座標ではなく、透視射影によって投影された座標であることや、点対応をとるために立体上の特徴点を追跡した際に生じる追跡ミスによる誤差などを含むことから、この関係式の等号は厳密には成立しない。そこで、実際に得られた計測行列、すなわち関係式の左辺から、関係式の等号が成立するように誤差を取り除かなくてはならない。この誤差を取り除く最も簡便な方法は、関係式において、右辺のランクは 3(以下)であり、左辺のランクは誤差の影響から一般的に 4 以上となることを利用して、関係式の左辺、すなわち透視射影によって投影された画像座標を並べた計測行列を一番良く近似するランクが 3(以下)の行列を求めることである。

最小 2 乗誤差基準で透視射影によって投影された画像座標を並べた計測行列を一番良く近似するランクが 3(以下)の行列は、主成分分析を行ない(その表現は特異値分解(Singular Value Decomposition; SVD)によって与えられる)、第 3 主成分までを用いて、第 4 主成分以下を誤差とみなして無視する(0 とおく)ことによって得られる。このとき、第 3 主成分までの寄与率は透視射影をアフィン近似したときの妥当性を測る一つの尺度と考えることができる。なぜなら、第 3 主成分までの寄与率が大きいことは、計測行列のランク 3(以下)らしさが大きいということであり、それはすなわち、立体形状の 3 次元空間座標からカメラで撮影することによ

て得られた 2 次元座標への写像が線型写像に近いことを意味すると考えることができるからである。

このように、透視射影によって得られた計測行列の成分からアフィン近似射影によって投影されたときに得られる計測行列を推定することができれば、あとは計測行列を運動行列と形状行列の積に分解するだけでカメラ運動と立体形状を復元することができる。ここで画像座標が正規直交基底による表現であることから、正しい復元解を得るためには、画像座標の基底が正規直交基底となるように分解しなければならない。第 3 章で見るとように因子分解法はそのアルゴリズムを与える。このとき画像座標の基底が正規直交基底となるような計測行列の分解は 1 通りではなく、形状行列を表現する正規直交基底の自由度だけ計測行列の分解の自由度が残されている。しかし、複数の 2 次元画像からの 3 次元形状復元問題は対象物体とカメラとの相対的な位置の復元しかできないため、正規直交基底で表現される自由度は本質的に正規直交基底が右手系であるか左手系であるかの 2 つに集約される。なお、この 2 種類の復元解のうちいずれの解が真の復元解であるかは、点特徴のみからは区別することができない。これは Necker Reversal と呼ばれる現象(3.2 節参照)であり、いずれの解も実在可能であるからである。

結局、アフィン近似射影のもとでの、点対応を用いた複数の 2 次元画像からの 3 次元形状復元は互いに Necker Reversal の関係にある 2 組の実在可能な解の復元となる。

なお、得られた画像の組によっては、第 3 主成分までの寄与率が不十分であることがある。このような場合、因子分解法によって得られたカメラ運動と立体のユークリッド形状の精度は一般的に不十分である。そこで透視射影による像とアフィン近似射影による像とのずれを利用し、透視射影の画像、すなわち実際に得られた画像から、アフィン近似射影による画像を因子分解法を利用しながら反復的に推定することによって、因子分解法によるカメラ運動と立体形状を高精度に復元する手法も提案されている (Christy and Horaud (1996))。

本稿では、第 2 章において透視射影とそのアフィン近似射影について説明し、第 3 章にて因子分解法のアルゴリズムについて説明する。第 4 章では 3 次元形状を一意に定めるためには何枚のアフィン射影画像が必要であるかについて説明し、第 5 章において計測行列の特異値と復元解の安定性の関わりについて説明する。第 6 章において透視射影モデルの因子分解法について説明し、第 7 章において、実時間処理に対応するための逐次型因子分解法について説明する。

2. カメラモデル

本章では、ピンホールカメラを表現したモデルである透視射影モデルとそのアフィン近似モデルについて解説する。これらカメラモデルは、3 次元空間の点からカメラの画像面(2 次元平面)への変換として捉えることができる。

2.1 透視射影モデル

透視射影モデルはピンホールカメラを表現したモデルであり、3 次元座標から 2 次元座標への変換として捉えることができる。

第 f カメラ座標系における第 p 特徴点の座標を $\mathbf{X}_{fp} = (X_{fp}, Y_{fp}, Z_{fp})^T$ 、第 f 画像面における第 p 特徴点の画像座標を $\mathbf{x}_{fp} = (x_{fp}, y_{fp})^T$ とし、カメラの焦点距離を l とすると、透視射影モデルは

$$(2.1) \quad \mathbf{x}_{fp} = \frac{l}{Z_{fp}} \begin{pmatrix} X_{fp} \\ Y_{fp} \end{pmatrix} \iff \begin{pmatrix} \mathbf{x}_{fp} \\ l \end{pmatrix} = \frac{l}{Z_{fp}} \mathbf{X}_{fp}$$

として表現される(図 3)。

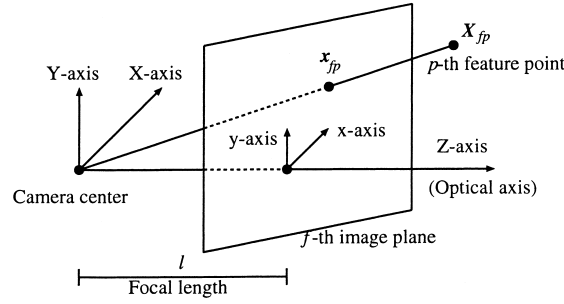


図 3. 透視射影モデル.

しかし、透視射影は非線型写像であるから、複数の透視射影画像からの物体の3次元形状の復元問題は非線型写像の逆問題となり、非線型最適化問題に帰着される。一般に非線型最適化問題はノイズに敏感で初期値依存性が高く数値計算上不安定であるという問題があるため、理想的なカメラにおいて安定して3次元形状を復元するのは難しい。この困難を乗り越えるために、理想的なカメラモデルである透視射影を正射影モデルなどのアフィン射影に近似することによって、複数の2次元画像からの点対応を用いたカメラ運動と立体形状の復元問題を線型写像の逆問題に帰着させるという試みが行われてきた。非線型写像の逆問題を線型写像の逆問題に近似することにより、数値計算上安定して復元解が得られるからである。もちろん、実際の画像は理想的なカメラモデルである透視射影によって得られるため、透視射影をアフィン近似したときの近似誤差が避けられないために復元精度に限界があるという欠点はあるが、数値計算上安定して復元解が得られるということは重要であるし、また透視射影の場合に帰着される非線型最適化問題の良い初期値として利用できるという利点もある。

そこで、次に透視射影のアフィン近似射影としての paraperspective 射影モデル (Poelman and Kanade (1997))、scaled orthographic 射影モデル、正射影モデルについて説明する。

2.2 Paraperspective 射影モデル

透視射影の関係式 (2.1) を、ある特徴点 (第 * 特徴点と呼ぶ) X_{f*} とその画像座標 x_{f*} を中心としてアフィン近似する。なお、アフィン射影においては、特徴点の重心が画像座標における特徴点の重心に変換されるので、特徴点の重心をアフィン近似の中心としても良く、このときは特徴点の重心を第 (G) 特徴点と考えれば良い。

このとき透視射影のアフィン近似射影は、アフィン近似の中心からの相対座標 $X_{f_p}^* = X_{f_p} - X_{f*}$ 、 $x_{f_p}^* = x_{f_p} - x_{f*}$ を用いると簡潔に記述できる。式 (2.1) により

$$(2.2) \quad \begin{aligned} \begin{pmatrix} x_{f_p} \\ l \end{pmatrix} &= \frac{l}{Z_{f*} + Z_{f_p}^*} (X_{f*} + X_{f_p}^*) \\ &= \frac{l}{Z_{f*}} \left\{ I_3 - \frac{1}{Z_{f*}} X_{f*} (0, 0, 1) \right\} X_{f_p} + \frac{l}{Z_{f*}} X_{f*} + O(\|X_{f_p}^*\|^2) \end{aligned}$$

が成立するので $O(\|X_{f_p}^*\|^2)$ の項を無視すると、カメラ座標 X_{f_p} から画像座標 x_{f_p} への射影は、アフィン射影

$$(2.3) \quad x_{f_p} = \frac{l}{Z_{f*}^2} \begin{pmatrix} Z_{f*} & 0 & -X_{f*} \\ 0 & Z_{f*} & -Y_{f*} \end{pmatrix} X_{f_p} + \frac{l}{Z_{f*}} \begin{pmatrix} X_{f*} \\ Y_{f*} \end{pmatrix} = A_f^{\text{para}} X_{f_p} + x_{f*}$$

によって近似される．このような近似モデルを paraperspective 射影と呼び， A_f^{para} を paraperspective 射影行列と呼ぶ．この近似が成立するのは， Z_{fp}^* が Z_{f*} に比べて十分小さいとき，すなわち，3次元物体がその厚みに比べて，十分カメラから遠くにある場合に成立する．そして，paraperspective 射影は具体的には次のようなアフィン射影となる(図4参照)．

1. X_{fp} は画像面と平行な平面 $Z = Z_{f*}$ に X_{f*} 方向に射影される．
2. この射影された点が画像面に透視射影される(画像としては1.の画像を原点中心に l/Z_{f*} 倍拡大したものとなる)．

2.3 Scaled orthographic 射影モデル

Paraperspective 射影モデルにおいて，アフィン近似の中心が光軸に十分近いとき，つまり $X_{f*}/Z_{f*} \approx 0$, $Y_{f*}/Z_{f*} \approx 0$ が成立するとき，paraperspective 射影モデル(2.3)は

$$(2.4) \quad x_{fp} = \frac{l}{Z_{f*}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} X_{fp} = A_f^{\text{scaled}} X_{fp}$$

のように，さらに簡単なモデルに近似することができる．このような近似モデルを scaled orthographic 射影モデル(weak perspective 射影モデル)と言い， A_f^{scaled} を scaled orthographic 射影行列と呼ぶ．そして，scaled orthographic 射影は具体的には次のようなアフィン射影となる(図4参照)．

1. X_{fp} は画像面と平行な平面 $Z = Z_{f*}$ に光軸方向に射影，すなわち正射影される．
2. この射影された点が画像面に透視射影される(画像としては1.の画像を原点中心に l/Z_{f*} 倍拡大したものとなる)．

2.4 正射影モデル

Scaled orthographic 射影モデルにおいて，カメラ中心と，参照した特徴点の距離が全画像を通してほぼ一定であるとき，つまり $Z_{1*} \approx Z$, $Z_{2*} \approx Z, \dots, Z_{F*} \approx Z$ が成立しているとき，式(2.4)は

$$(2.5) \quad x_{fp} = \frac{l}{Z} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} X_{fp}$$

のように近似することができる．ここで対象物体を l/Z 倍に拡大したときの特徴点のカメラ座標を Y_{fp} とおく，すなわち

$$(2.6) \quad Y_{fp} = \frac{l}{Z} X_{fp}$$

とおくと式(2.5)は

$$(2.7) \quad x_{fp} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} Y_{fp} = A_f^{\text{ortho}} Y_{fp}$$

となる．これは正射影であり， A_f^{ortho} を正射影行列と呼ぶ．

透視射影，paraperspective 射影モデル，scaled orthographic 射影モデル及び正射影モデルによる像の位置関係は図4のようになる．

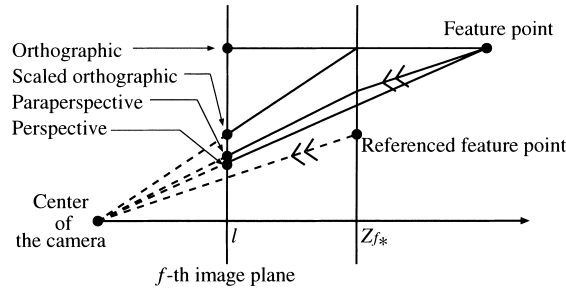


図 4. 透視射影とそのアフィン射影近似モデル.

2.5 計量アフィン射影モデル

前節で述べた，正射影，scaled orthographic 射影，paraperspective 射影モデルを統一的に扱うために，本節では，これらモデルを含む計量アフィン射影 (MAP; Metric Affine Projection) モデル(以前は一般アフィン射影 (Generalized Affine Projection) モデルという名称を用いていたが，Mundy and Zisserman (1992) によって提案されたアフィンモデルと区別しにくいために名称を変更した)について説明する．

Mundy and Zisserman (1992) によって提案されたアフィンモデルは校正されていないカメラに対するモデルとして定義され， X_{fp} から x_{fp} への変換が

$$(2.8) \quad x_{fp} = A_f X_{fp} + u_f \quad (2 \times 3)$$

の型のアフィン射影によって得られるモデルのことを言う．ここで A_f 及び u_f は未知パラメータである．アフィンモデルでは， A_f には何の仮定もなされていないために対象物体のアフィン復元，すなわちアフィン空間における位置関係を知ることは可能であっても，対象物体のユークリッド復元，すなわち長さや角度の計量情報を知ることはできない．そこで対象物体のユークリッド復元を行なうためには A_f の成分に条件を仮定しなければならない (u_f は画像座標における平行移動成分であるためにカメラ位置の復元にのみ影響を及ぼし，ユークリッド形状の復元には影響を及ぼさない)．そこで A_f に対してユークリッド復元を行なうための条件を仮定したモデルとして MAP モデルが提案された(藤木・蔵田(1997b))．

例えば scaled orthographic 射影モデルの場合に

$$(2.9) \quad x_{fp} = \frac{l}{Z_{f*}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} X_{fp} \iff x_{fp} = \frac{l}{tZ_{f*}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} (tX_{fp}) \quad (t > 0)$$

が成立するように一般に，対象物体の大きさが t 倍の物体を t 倍離れた距離から観測しても同じ画像が得られる(正射影モデルの場合は正射影モデルを仮定した時点で 3 次元空間全体のスケールを固定しているためこのようなことは起こらない)ため，複数の 2 次元画像からのカメラ運動と 3 次元形状復元において，3 次元空間全体のスケールの不定性が生じる．この t で表される不定性はグローバルスケールパラメータと呼ばれるパラメータであり， Z_f と分離できないと考えるのが自然である．つまり tZ_{f*} のように t と Z_{f*} をひとまとめとしたグローバルスケールパラメータを含むパラメータ $\lambda_{f*} = tZ_{f*}$ を考えるのが自然である．このパラメータを奥行きパラメータと呼ぶ．ここで $\lambda_{f*} = tZ_{f*}$ は奥行きパラメータだけでなく，グローバルスケールパラメータも兼ねることから A_f から奥行きパラメータ λ_{f*} をくり出すことができると考えるのが自然である．また，対象物体のユークリッド復元を行なうには A_f から奥行き

パラメータ λ_{f*} をくくり出した残りである行列 B_f の成分が既知である(画像座標から計算できる)必要がある. つまり

$$(2.10) \quad A_f = \frac{1}{\lambda_{f*}} B_f$$

(B_f は既知)という型をしている必要がある(正射影モデルのように奥行きパラメータを考慮しない場合は A_f 自体が既知であるとする. なお, 因子分解法を参照すればわかるように, 対象物体のユークリッド復元を行なうには $B_f B_f^T$ が既知であれば十分であるが, 本稿ではカメラ運動を復元するために B_f が既知であるとする).

さらにカメラの位置の復元を行なうためには u_f が既知であることが必要である. 通常はアフィン近似の中心である第 * 特徴点が透視射影に従うように選んでおく. このとき $\begin{pmatrix} x_{f*} \\ l \end{pmatrix} = \frac{l}{\lambda_{f*}} X_f$ である.

アフィンモデル以上に述べた仮定を加えたモデルを計量アフィン射影 (Metric Affine Projection; MAP) モデルと呼ぶ. 以上の仮定により MAP モデルは

$$(2.11) \quad x_{fp}^* = A_f X_{fp}^* = \frac{1}{\lambda_{f*}} B_f X_{fp}^*, \quad X_f^* = \frac{\lambda_{f*}}{l} \begin{pmatrix} x_{f*} \\ l \end{pmatrix}$$

と表現される(正射影モデルのように A_f が奥行きパラメータを含まない場合はカメラ位置は復元できない). ここで A_f を MAP 行列と呼ぶ.

2.6 MAP モデルと仮想画像面

前節の B_f を特異値分解した結果を

$$(2.12) \quad B_f = R_f \Sigma_f D_f, \quad R_f^T R_f = D_f D_f^T = I_2, \quad \Sigma_f = \text{diag}\{p_f, q_f\}$$

とすると B_f は既知であるから, 直交行列 R_f , 対角行列 Σ_f , 各行が正規直交である行列 D_f は既知である. このとき式 (2.11) により MAP モデルの相対座標による表現は

$$(2.13) \quad x_{fp}^* = A_f X_{fp}^* = \frac{1}{\lambda_{f*}} R_f \Sigma_f D_f X_{fp}^*$$

となる. ここで $D_f = (c_f, d_f)^T$ は各行が正規直交である行列であり,

$$(2.14) \quad X_{fp}^* = D_f X_{fp}^* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \widetilde{D}_f X_{fp}^*$$

とおくと, X_{fp}^* は X_{fp}^* を平面 $\text{span}\{c_f, d_f\}$ (c_f, d_f により張られる線型空間であり, これは A_f^T の列空間に等しい)へ正射影した画像の基底 $\{c_f, d_f\}$ による表現であることがわかる. ここで, \widetilde{Y} は 2×3 行列 $Y = (Y_1, Y_2)^T$ から導かれる 3×3 行列 $(Y_1, Y_2, Y_1 \times Y_2)^T$ を表すものとする. このとき

$$(2.15) \quad x_{fp}^* = R_f \Sigma_f \left(\frac{1}{\lambda_{f*}} X_{fp}^* \right)$$

であり, $\frac{1}{\lambda_{f*}} X_{fp}^*$ は X_{fp}^* を平面 $\text{span}\{c_f, d_f\}$ へ scaled orthographic 射影した画像の基底 $\{c_f, d_f\}$ による表現(\widetilde{D}_f が 3 次元回転行列であることに注意)であるから, 画像平面に MAP された画像は, 平面 $\text{span}\{c_f, d_f\}$ へ scaled orthographic 射影した画像を, この平面上で基底 $\{c_f, d_f\}$ における表現が $R_f \Sigma_f$ であるような線型変換を施したものとなっている. 逆に言えば,

$$(2.16) \quad \frac{1}{\lambda_{f*}} X_{fp}^* = \Sigma_f^{-1} R_f^T x_{fp}^*$$

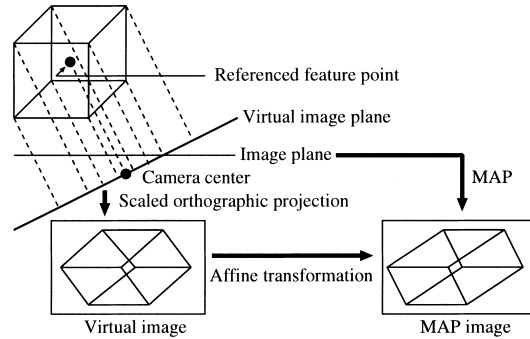


図 5. MAP モデルと仮想画像面.

が成立するので，MAP 画像は scaled orthographic 射影画像に帰着することができ，scaled orthographic 射影画像に関する任意の手法（例えば，Xu and Sugimoto (1998) など）は MAP 画像に適用できることがわかる．

ここでさらに λ_{f^*} の値がわかれば MAP 画像は正射影画像に帰着させることができるが，既に述べたようにグローバルスケールパラメータと奥行きパラメータは分離できないため，画像のみからでは λ_{f^*} の値は正確に求めることができない．しかし，複数の 2 次元画像においてグローバルスケールパラメータを 1 つ定めれば，それに応じて λ_{f^*} の値はそれぞれ 1 つに定まるので，結局， $\{\lambda_{f^*}\}_{f=1}^F$ の比の値は定めることができる． $\{\lambda_{f^*}\}_{f=1}^F$ の比の値の定め方についてはアフィンエピポーラ幾何学による方法 (Xu and Zhang (1996)) など複数提案されているが，本稿においては第 4 章で述べることにする．

ともかく，グローバルスケールパラメータを固定しても一般性は失われないので， $\{\lambda_{f^*}\}_{f=1}^F$ の比の値を求めることができれば $\{\lambda_{f^*}\}_{f=1}^F$ の値が求まったと考えることができる．よって，複数の 2 次元画像が与えられれば MAP 画像は正射影画像に帰着させることができると考えて良い．

このように，MAP 画像は平面 $\text{span}\{c_f, d_f\}$ への正射影（または scaled orthographic 射影）を経由して考えると非常に分りやすくなる．この平面 $\text{span}\{c_f, d_f\}$ を仮想画像面 (virtual image plane) と呼ぶことにする(図 5)．

3. 因子分解法

本章では，複数の 2 次元 MAP 画像からカメラ運動と立体形状を同時に復元する手法である因子分解法 (Tomasi and Kanade (1992)) について解説する．

2 次元画像はカメラと立体形状の相対的な位置関係によって定まるので(背景のない)複数の 2 次元画像からのカメラ運動と立体形状の復元問題はカメラと立体形状の相対的な位置の復元問題であると言える．つまり(背景のない)複数の 2 次元画像のみからでは，カメラが固定され立体が運動している，立体が固定されカメラが運動している，カメラも立体も運動している，の 3 つの状況を区別することができない．逆に言えば，複数の 2 次元画像におけるカメラと立体形状の相対的な位置関係を記述するときには立体が固定されカメラが運動していると仮定しても一般性を失わない．因子分解法は立体が固定されカメラが運動していると仮定してカメラ運動と立体形状を同時に復元する手法である．

前章で述べたカメラモデルはカメラに固定された座標系によって記述されているが，因子分

解法は立体が固定されていると仮定してカメラ運動と立体形状を同時に復元する手法なので、カメラモデルを立体に固定された座標系(世界座標系と呼ぶ)で記述する必要がある。ここで第 f 画像におけるカメラの位置の世界座標を t_f とし、第 f 画像面上の正規直交基底を $\{i_f, j_f\}$ 、カメラの光軸方向の単位ベクトルを k_f 、つまり、世界座標におけるカメラの向きを表す行列(カメラ基底行列)を $C_f = (i_f, j_f, k_f)^T$ とする。このとき第 p 特徴点の世界座標を s_p とし、第 f 画像のカメラ座標系における空間座標を X_{fp} とすると

$$(3.1) \quad s_p = t_f + C_f^T X_{fp}$$

が成立する。この表現をある特徴点 s_* からの相対座標 $s_p^* = s_p - s_*$, $t_f^* = t_f - s_*$ で表すと

$$(3.2) \quad s_p^* = t_f^* + C_f^T X_{fp} \iff X_{fp} = C_f(s_p^* - t_f^*)$$

となる。よって式 (2.11) により、MAP モデルを世界座標系で記述すると

$$(3.3) \quad x_{fp}^* = A_f X_{fp}^* = A_f C_f s_p^*, \quad t_f^* = -\frac{\lambda_{f*}}{l} C_f^T \begin{pmatrix} x_{fp}^* \\ l \end{pmatrix}$$

のようになる。このとき、各画像の仮想画像面は $(A_f C_f)^T$ の列空間となることに注意しておく。

3.1 因子分解法

P 個の点特徴の画像が F 枚得られたとき、複数の 2 次元画像からのカメラ運動と立体形状の復元問題は式 (3.3) $_{f=1; p=1}^F$ から $\{C_f\}_{f=1}^F$, $\{s_p^*\}_{p=1}^P$ (及び $\{\lambda_{f*}\}_{f=1}^F$) を求める問題となる。因子分解法は FP 個の式 (3.3) $_{f=1; p=1}^F$ から作られた行列を分解することによって $\{C_f\}_{f=1}^F$, $\{s_p^*\}_{p=1}^P$ (及び $\{\lambda_{f*}\}_{f=1}^F$) を求める手法である。

ここで計測行列 W^* 、運動行列 M 及び形状行列 S^* を

$$(3.4) \quad W^* = \begin{pmatrix} W_1^* \\ \vdots \\ W_F^{*T} \end{pmatrix}, \quad W_f^* = (x_{f1}^*, \dots, x_{fP}^*),$$

$$M = \begin{pmatrix} M_1 \\ \vdots \\ M_F \end{pmatrix}, \quad M_f = A_f C_f, \quad S^* = (s_1^*, \dots, s_P^*)$$

で定義すると

$$(3.5) \quad W^* = \underset{(2F \times 3)}{M} \underset{(3 \times P)}{S^*}$$

が成立する。式 (3.5) が式 (3.3) $_{f=1; p=1}^F$ と同値な条件であることは明らかであろう。ここで M にはカメラ運動に関する未知量 $\{C_f\}_{f=1}^F$ (および $\{\lambda_{f*}\}_{f=1}^F$) のみが、 S^* には立体形状に関する未知量 $\{s_p^*\}_{p=1}^P$ のみが含まれていることから、 W^* を M と S^* の積に分解することができれば直ちにカメラ運動と立体形状を復元することができる。

ここで W^* の M と S^* の積への分解においてみたすべき条件がある。それは C_f が 3 次元回転行列であることから導かれる計量拘束と呼ばれる

$$(3.6) \quad M_f M_f^T = A_f A_f^T = \frac{1}{\lambda_{f*}^2} B_f B_f^T$$

という条件式である(右辺は λ_{f*} を除いて既知であることに注意)。よって W^* を式 (3.6) $_{f=1}^F$ が成立するように M と S^* の積へ分解することができればカメラ運動と立体形状を復元することができる。この分解を以下の手順で行なうのが因子分解法である。

まず，計測行列 W^* のランクは式 (3.5) により 3(以下)となるので特異値分解やランク標準形などを利用して

$$(3.7) \quad W^* = \underset{(2F \times 3)}{\hat{M}} \underset{(3 \times P)}{\hat{S}^*} = \begin{pmatrix} \hat{M}_1 \\ \vdots \\ \hat{M}_F \end{pmatrix} \hat{S}^*$$

の形に暫定的に分解する(実際に観測された計測行列では，モデルの近似誤差や観測誤差などのノイズの影響により計測行列のランクが 4 以上となることがある．その際，計測行列の第 4 特異値以下を 0 とおくことにより，ノイズを含む計測行列からの計測行列の最小 2 乗推定量が得られるので，特異値分解を利用することが多い(詳しくは第 5 章参照)．また，特異値分解以外の分解例は Yokoya et al. (1998) 参照)．

このとき， $M, S^*, \hat{M}, \hat{S}^*$ の間には

$$(3.8) \quad M = \hat{M}A, \quad S^* = A^{-1}\hat{S}^*$$

をみたく 3×3 可逆行列 A が存在する．つまり暫定的な分解 \hat{M}, \hat{S}^* によって得られた運動行列と形状行列は可逆なアフィン変換によって真の復元解に移ることができる．よって暫定的な分解によって運動と形状がアフィン復元されていることがわかる．

このように暫定的な分解によって得られたアフィン復元解からユークリッド復元解を求める問題は 3×3 可逆行列 A を求める問題に帰着され， A は以下のようにすれば求めることができる．

A のみたすべき条件は， $Q = AA^T$ とおくと，計量拘束 (3.6) により

$$(3.9) \quad \hat{M}_f Q \hat{M}_f^T = A_f A_f^T = \frac{1}{\lambda_{f*}^2} B_f B_f^T$$

となる．ここで式 (3.9) $_{f=1}^F$ における未知量は $\{\lambda_{f*}\}_{f=1}^F$ の比，及び Q である．

今， B_f は既知であるから， B_f の特異値分解 $R_f \Sigma_f D_f$ において R_f, Σ_f は既知であり，

$$(3.10) \quad \hat{P}_f = (\hat{p}_f, \hat{q}_f)^T = R_f^T \hat{M}_f, \quad P_f = R_f^T M_f$$

とおくと拘束条件 (3.9) は

$$(3.11) \quad \hat{P}_f Q \hat{P}_f^T = \frac{1}{\lambda_{f*}^2} \Sigma_f^2$$

のようにより単純になる．このとき式 (3.11) は

$$(3.12) \quad \hat{p}_f^T Q \hat{p}_f = \frac{p_f^2}{\lambda_{f*}^2}, \quad \hat{p}_f^T Q \hat{q}_f = 0, \quad \hat{q}_f^T Q \hat{q}_f = \frac{q_f^2}{\lambda_{f*}^2}$$

$$\iff \frac{\hat{p}_f^T Q \hat{p}_f}{p_f^2} = \frac{\hat{q}_f^T Q \hat{q}_f}{q_f^2} = \frac{1}{\lambda_{f*}^2}, \quad \hat{p}_f^T Q \hat{q}_f = 0$$

となる．よって λ_{f*} を含まない

$$(3.13) \quad q_f^2 \hat{p}_f^T Q \hat{p}_f - p_f^2 \hat{q}_f^T Q \hat{q}_f = 0, \quad \hat{p}_f^T Q \hat{q}_f = 0$$

という方程式が得られる．この方程式は 3×3 正値対称行列 Q に関する線型同次連立方程式であり，この連立方程式を解くことによってグローバルスケールパラメータと等価である定数倍の不定性を除いて Q を一意に定めることができる．

グローバルスケールパラメータを適当な値に固定して Q を一意に定めた後、第 * 特徴点の奥行きパラメータ (* = (G), つまりアフィン近似の中心を特徴点の重心に定めた場合は平均奥行き (average depth) と呼ぶ) $\{\lambda_{f*}\}_{f=1}^F$ の値は

$$(3.14) \quad \lambda_{f*} = \sqrt{\frac{p_f^2}{\hat{\mathbf{p}}_f^T Q \hat{\mathbf{p}}_f}} = \sqrt{\frac{q_f^2}{\hat{\mathbf{q}}_f^T Q \hat{\mathbf{q}}_f}}$$

によって定めることができる。

ここで理論的には Q は正値対称行列であるが、実データを用いた場合観測誤差や線型近似誤差などの影響から Q が正値対称行列ではない形で求まることがある。そのため以下のように $Q = (Q_{ij})$ を推定することが多い。

$\mathbf{q} = (Q_{11}, Q_{12}, Q_{13}, Q_{22}, Q_{23}, Q_{33})^T$ と定めると、任意の \mathbf{a}, \mathbf{b} に対して

$$(3.15) \quad \omega(\mathbf{a}, \mathbf{b})^T \mathbf{q} = \mathbf{a}^T Q \mathbf{b}$$

をみたすような 6 次元列ベクトル $\omega(\mathbf{a}, \mathbf{b})$ が一意に定まる。このとき拘束条件 (3.13) は

$$(3.16) \quad \begin{pmatrix} q_f^2 \omega(\hat{\mathbf{p}}_f, \hat{\mathbf{p}}_f)^T - p_f^2 \omega(\hat{\mathbf{q}}_f, \hat{\mathbf{q}}_f)^T \\ \omega(\hat{\mathbf{p}}_f, \hat{\mathbf{q}}_f)^T \end{pmatrix} \mathbf{q} = \hat{\omega}_f^T \mathbf{q} = \mathbf{0}_2$$

となる。よって

$$(3.17) \quad \underset{(3F \times 6)}{\Omega} = (\hat{\omega}_1, \dots, \hat{\omega}_F)^T$$

の固有値 0 に対する固有ベクトル (実際的には最小固有値に対する固有ベクトル) として \mathbf{q} を求めれば良い。もしくは、グローバルスケールパラメータを $\omega(\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_1) = 1$ となるように固定して

$$(3.18) \quad \begin{pmatrix} \omega(\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_1) \\ \Omega \end{pmatrix} \mathbf{q} = \begin{pmatrix} 1 \\ \mathbf{0}_{2F} \end{pmatrix}$$

とおき、

$$(3.19) \quad \mathbf{q} = \begin{pmatrix} \omega(\hat{\mathbf{p}}_1, \hat{\mathbf{p}}_1) \\ \Omega \end{pmatrix}^+ \begin{pmatrix} 1 \\ \mathbf{0}_{2F} \end{pmatrix}$$

によって \mathbf{q} の最小 2 乗推定量を求めても良い。ここで X^+ は行列 X のムーア・ペンローズ逆行列を表すものとする。このとき

$$(3.20) \quad \lambda_{f*} = \sqrt{\frac{p_f^2}{\omega(\hat{\mathbf{p}}_f, \hat{\mathbf{p}}_f)^T \mathbf{q}}} = \sqrt{\frac{q_f^2}{\omega(\hat{\mathbf{q}}_f, \hat{\mathbf{q}}_f)^T \mathbf{q}}}$$

となる。この手続きで得られる Q は必ず対称行列となる (正確には Q は正値対称行列であり、データによっては Q が正値とならない場合がある。そこで Q が正値対称行列となるように Q を推定する手法も提案されている (Quan (1996)) が、 Q が正値とならない場合は 3 次元復元のために用いるデータが信頼できないと考えるのが自然であると筆者は考える)。

これらの手法によって得られた (正値) 対称行列 Q のコレスキー分解を $Q = LL^T$ とすると A の一般解は $A = L^T U (\forall U \in O(3))$ となり、 M, S^* の一般解は

$$(3.21) \quad M = \hat{M} L^T U, \quad S^* = U^T L \hat{S}^*$$

となる。ここで直交行列 U で表される M 及び S^* の自由度は、世界座標の取り方の自由度に一致するので、 $\det U$ の正負 (世界座標系を右手系で取るか左手系で取るか) に応じて立体形状が鏡映対称な 2 組の復元解を得る。

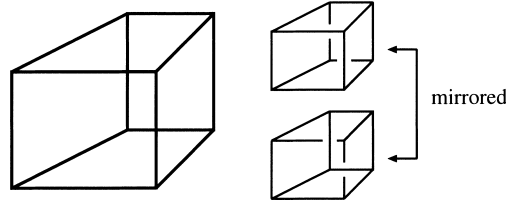


図 6. Necker reversal.

なお，第 f 運動行列 M_f ，または $V_f = \lambda_{f*} \Sigma_f^{-1} P_f$ から外部パラメータ C_f は

$$(3.22) \quad C_f = (\widetilde{A}_f)^{-1} \widetilde{M}_f = (\widetilde{D}_f)^{-1} \widetilde{V}_f$$

によって復元される．カメラ位置 t_f^* は式 (3.3) によって復元される．

3.2 2 組の復元解の関係

前節において， $\det U$ の正負に応じて立体形状が鏡映対称な 2 組の復元解を得ると述べたが，点対応を用いた複数の MAP 画像からの運動と形状の復元においてこれら 2 組の復元解のうちのどちらが真の復元解を表現しているか選び出すことができるだろうか．その答は否である．すなわち，MAP を仮定した場合に点対応のみからは 2 組の復元解のうちどちらが真の復元解を表現しているかを選び出すことはできない(透視射影を仮定した場合は，2 組の復元解のうちどちらが真の復元解を表現しているかを選び出すことができる．これについては第 6 章を参照)．この事実は Necker Reversal と呼ばれる現象と等価である．Necker Reversal とは，例えば，図 6 の立体を見ると，見方によって 2 通りに見ることができるような現象を言う．ここで互いに Necker Reversal となる 2 つの立体は奥行きが反転している，つまり鏡映対称な立体となっている．

点対応を用いた MAP 画像からの運動と形状の復元において Necker Reversal の不定性を除去できないことを示すには，ある立体形状を撮影した任意の画像に対して，その立体形状と鏡映対称な立体形状を，適当な方向から撮影すると全く同一の画像が得られることを示せば十分である．そこで，これらの位置関係を具体的に記述することを試みる．

ここで数学的には一方の解 S^* に対する鏡映対称な立体形状を $-S^*$ とおくのが簡明であり，視覚的には一方の解 S^* に対する鏡映対称な立体形状を $\text{diag}\{-1, 1, 1\}S^*$ とおくのが簡明である．本節では両方の場合についてカメラ運動がどのように対応しているかを明らかにする．

まず数学的に簡明な場合について考える．一方の解 S^* に対するカメラ基底行列 C_f ともう一方の解 $-S^*$ に対するカメラ基底行列 H_f との関係について述べる． C_f, H_f は 3 次元回転行列，すなわち行列式が 1 の直交行列であり，

$$(3.23) \quad A_f C_f = -A_f H_f \iff D_f C_f = -D_f H_f$$

をみたく．よって

$$(3.24) \quad H_f = (\widetilde{A}_f)^{-1} \text{diag}\{-1, -1, 1\} \widetilde{A}_f C_f = (\widetilde{D}_f)^T \text{diag}\{-1, -1, 1\} \widetilde{D}_f C_f$$

となる．なお，正射影，scaled orthographic，paraperspective 各モデルの場合

$$(3.25) \quad H_f = \begin{cases} \text{diag}\{-1, -1, 1\} C_f & (\text{正射影, scaled orthographic}) \\ \frac{1}{g_f^2} \begin{pmatrix} 2\mathbf{x}_{f*}^{\text{para}} \mathbf{x}_{f*}^{\text{para}T} - g_f^2 I_2 & 2l\mathbf{x}_{f*}^{\text{para}} \\ g_f^2 & g_f^2 - 2u_f^2 \end{pmatrix} C_f & (\text{paraperspective}) \end{cases}$$

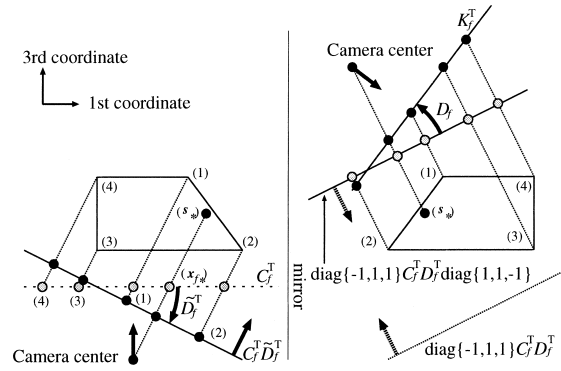


図 7. MAP の Necker reversal.

となる(藤木(2000)).

次に視覚的に簡明な場合について考える. 一方の解 S^* に対するカメラ基底行列 C_f ともう一方の解 $\text{diag}\{-1, 1, 1\}S^*$ に対するカメラ基底行列 K_f との関係について述べる. C_f, K_f は 3 次元回転行列, すなわち行列式が 1 の直交行列であり,

$$(3.26) \quad A_f C_f = A_f K_f \text{diag}\{-1, 1, 1\} \iff D_f C_f = D_f K_f \text{diag}\{-1, 1, 1\}$$

をみたく, よって

$$(3.27) \quad \begin{aligned} K_f &= -H_f \text{diag}\{-1, 1, 1\} \\ &= (\widetilde{A}_f)^{-1} \text{diag}\{1, 1, -1\} \widetilde{A}_f C_f \text{diag}\{-1, 1, 1\} \\ &= (\widetilde{D}_f)^T \text{diag}\{1, 1, -1\} \widetilde{D}_f C_f \text{diag}\{-1, 1, 1\} \end{aligned}$$

となる. ここで

$$(3.28) \quad K_f^T = \text{diag}\{-1, 1, 1\} C_f^T (\widetilde{D}_f)^T \text{diag}\{1, 1, -1\} \widetilde{D}_f$$

において, $\text{diag}\{-1, 1, 1\}X$ は X の各列を第 1 成分に関して対称移動させたものであり, $X \text{diag}\{1, 1, -1\}$ は X の第 3 列を反転させたものであることに気をつけると C_f^T と K_f^T の位置関係は図 7 のようになることがわかる. このとき, 図 7 の左側のように立体形状とカメラがあるときに撮影される画像は, 図 7 の右側のように立体形状とカメラがあるときに撮影される画像と全く同一のものとなるため, 画像座標からのみでは真の復元解が左側であるか右側であるかは区別することができない. もちろん, 実際の画像において不透明な立体を撮影したときには立体の裏側は隠れるので, どちらが真の形状かは判定できることもある.

4. 複数の MAP 画像からの復元解が一意に定まる条件

本章では復元解が一意に定まる条件について考えてみる. ここで, 点对応からの運動と形状の復元において Necker Reversal と等価な復元解の不定性は除いて考える. すなわち, 本章では互いに鏡映対象な復元解が一意に定まる条件について考えることにする. なお, 本章において観測している物体は直線や平面ではなく, 3 次元上に分布しているものとする. この条件は $\text{rank}S^* = 3$ と同値であり, 同一平面上にない 4 点があれば達成できる.

ここで $\text{rank}M = 2$ となるのは $\{M_f^T\}_{f=1}^F$ の列空間が全て一致する, すなわち F 枚の仮想画像面が全て一致するときのみである. このような場合が起こるのは, カメラが射影方向 (MAP

という仮定の下では透視射影と異なりすべての特徴点は同じ方向に射影される)に前後した場合のみであり、このとき得られる画像はすべて、観測している物体を同じ方向へ射影した画像となる。つまり、観測している物体を一方向からしか眺めていないことになるので 3 次元形状を復元することはできない。

これは透視射影の場合はカメラが一方向に前後した場合でも理論上物体の 3 次元形状を復元することが可能であることに矛盾しているように思われるが、MAP を仮定することにより、透視射影の下では各特徴点の射影される方向は一般に異なるのを、等しい方向であると近似したからである。逆に、透視射影の下ではカメラがカメラ中心を動かさずに回転した場合に物体の 3 次元形状を復元することは不可能であるが、MAP の下では理論上 3 次元形状を復元することが可能である。

このように、透視射影を MAP に近似したために理論的なずれが生じてしまうが、本稿では、画像が透視射影によってではなく、MAP によって得られたものであるとしてその復元理論について解説する。

$\text{rank}M = 2$ である場合は、今述べたように、画像が MAP によって得られた場合には 3 次元復元は不可能であるから以降においては断りのない限り $\text{rank}M = 3$ であるとする。つまり、以降においては断りのない限り $\text{rank}M = \text{rank}S^* = 3$ であるとする。

$\text{rank}M = \text{rank}S^* = 3$ が成立するとき、 $W^*_{(2F \times P)} = M_{(2F \times 3)} S^*_{(3 \times P)}$ の分解について、ある M に対する S^* が存在すれば一意(藤木・蔵田(1997a))であるから、因子分解法における計量行列 Q が一意に定まることと復元解が Necker Reversal の不定性を除いて一意に定まることは同値となる。

4.1 1 枚の画像から得られる条件

計量行列 Q を決定するために 1 枚の画像から得られる条件は式 (3.11) である。この式は計量行列が Q であるような 3 次元計量線型空間の 2 次元線型部分空間 $\text{span}\{\hat{p}_f, \hat{q}_f\}$ における部分計量を表現している。つまり、因子分解法において解が一意に定まるためには何枚の MAP 画像が必要であるかということと、3 次元計量線型空間における計量行列 Q が一意に定まるためには何枚の 2 次元線型部分空間における部分計量が必要であるかということは同値な命題となる。

ここで第 j 画像の仮想画像面が第 i 画像の仮想画像面と等しいとき、 $\text{span}\{\hat{p}_j, \hat{q}_j\}$ は $\text{span}\{\hat{p}_i, \hat{q}_i\}$ と同じ 2 次元線型部分空間となるので、第 j 画像から得られる Q を求めるための条件は第 i 画像から得られる Q を求めるための条件と同値となる。つまり、新しく得られた画像の仮想画像面が以前に得られた画像の仮想画像面と等しい場合、アフィン射影を仮定する限り何ら新しい 3 次元情報も得られないことに注意しておく。

4.2 2 枚の画像から得られる条件

本節においては、仮想画像面の異なる 2 枚の画像から計量行列 Q をどこまで特定できるかについて考える。1 枚の画像から得られる条件 (3.11) は 3 次元計量線型空間の 2 次元部分空間の計量を与える 3 つの独立した式であり、計量行列 Q は対称行列であるから自由度は 6 である。そのため、2 枚の画像が与えられれば一見 Q は一意に決定されそうであるが、2 枚の画像から与えられる Q を求めるための 6 つの条件式は互いに独立ではないため Q を一意に決定することはできない(藤木・蔵田(1997a))。そして以下に述べるように、 Q は 1 つの自由度を残して決定することができる。なお、2 枚の画像から運動と形状が 1 つの自由度を残して決定できることについては Koenderink and van Doorn (1991) によって知られることとなり、藤木・蔵田 (1997a) によって 1 つのパラメータで陽に表示された。

まず, 第 i, j 画像から与えられる 2 次元線型部分空間 $\text{span}\{\hat{p}_j, \hat{q}_j\}$ と $\text{span}\{\hat{p}_i, \hat{q}_i\}$ の交わりである 1 次元線型部分空間 $\text{span}\{\hat{p}_j, \hat{q}_j\} \cap \text{span}\{\hat{p}_i, \hat{q}_i\}$ における部分計量の無矛盾性から λ_{i*} と λ_{j*} の比を求めることができる. 何故なら $\hat{c}_{ij} \in \text{span}\{\hat{p}_i, \hat{q}_i\} \cap \text{span}\{\hat{p}_j, \hat{q}_j\}$ なる $\hat{c}_{ij} \neq \mathbf{0}_3$ は \hat{P}_i^T または \hat{P}_j^T の各列の線型結合で表現することができるので, 適当な 2 次元列ベクトル \hat{h}_i, \hat{h}_j を用いて

$$(4.1) \quad \hat{c}_{ij} = \hat{P}_i^T \hat{h}_i = \hat{P}_j^T \hat{h}_j$$

のように表すことができ, 式 (3.11) から

$$(4.2) \quad \lambda_{i*} : \lambda_{j*} = \|\Sigma_j \hat{h}_j\| : \|\Sigma_i \hat{h}_i\|$$

となるからである. 既に述べたように復元解にはグローバルスケールの不定性があるため, $\{\lambda_{f*}\}_{f=1}^F$ の具体的な値は求まらず, 比の値しか求まらないことに注意しておく. ここでグローバルスケールの不定性は取り除くことはできないが, 逆にグローバルスケールパラメータを具体的な値に固定しても一般性を失わない, つまり $\{\lambda_{f*}\}_{f=1}^F$ は既知であるとして良い. また, Σ_f も既知であるから, 拘束条件 (3.11) は

$$(4.3) \quad \hat{V}_f = (\hat{v}_f, \hat{w}_f)^T = \lambda_{f*} \Sigma_f^{-1} \hat{P}_f, \quad V_f = \lambda_{f*} \Sigma_f^{-1} P_f$$

とおくと

$$(4.4) \quad \hat{V}_f Q \hat{V}_f^T = I_2 \iff (\hat{V}_f \otimes \hat{V}_f) \text{cs } Q = \text{cs } I_2$$

が成立する. ここで cs は行列の列展開を表す記号であり, \otimes は行列のクロネッカー積を表す記号である. ここで式 (4.4) は正射影モデルの拘束条件であるから, 第 2 章で述べたように確かに MAP モデルは正射影モデルに帰着されることがわかる. つまり因子分解法において, 先に MAP モデルを正射影モデルに帰着し,

$$(4.5) \quad \hat{\mathcal{V}} = \left((\hat{V}_1 \otimes \hat{V}_1)^T, \dots, (\hat{V}_F \otimes \hat{V}_F)^T \right)^T, \quad \mathcal{I} = (\text{cs } I_2, \dots, \text{cs } I_2)^T$$

とおき $\text{cs } Q = \hat{\mathcal{V}}^+ \mathcal{I}$ によって計量行列 Q を求めても良いことがわかる.

さて, この \hat{c}_{ij} を用いて 2 枚の画像から得られる部分計量を表現する新しい基底を選び出すことにしよう. ここで 2 枚の画像は第 1, 2 画像であるとし, 繁雑にならないように \hat{c}_{12} を単に \hat{c} とする. また $\|\hat{c}\|_Q = 1$ としても一般性を失わない. このとき \hat{a}, \hat{b} を

$$(4.6) \quad \hat{a} = \langle \hat{c}, \hat{w}_1 \rangle_Q \hat{v}_1 - \langle \hat{c}, \hat{v}_1 \rangle_Q \hat{w}_1,$$

$$(4.7) \quad \hat{b} = \langle \hat{c}, \hat{w}_2 \rangle_Q \hat{v}_2 - \langle \hat{c}, \hat{v}_2 \rangle_Q \hat{w}_2$$

のように選ぶ(詳細は藤木・蔵田(1997a), 蔵田 他(1999 参照)と

$$(4.8) \quad \begin{aligned} \|\hat{a}\|_Q &= \|\hat{b}\|_Q = \|\hat{c}\|_Q = 1, \quad \langle \hat{a}, \hat{c} \rangle_Q = \langle \hat{b}, \hat{c} \rangle_Q = 0, \\ \hat{a} &\in \text{span}\{\hat{v}_1, \hat{w}_1\}, \quad \hat{b} \in \text{span}\{\hat{v}_2, \hat{w}_2\}, \\ \hat{c} &\in \text{span}\{\hat{v}_1, \hat{w}_1\} \cap \text{span}\{\hat{v}_2, \hat{w}_2\} \end{aligned}$$

が成立する($\text{span}\{\hat{v}_i, \hat{w}_i\} = \text{span}\{\hat{p}_i, \hat{q}_i\}$ ($i = 1, 2$) である). ここで \hat{a} は $\{\hat{v}_1, \hat{w}_1, \hat{v}_1 \times \hat{w}_1\}$ と $\{\hat{a}, \hat{c}, \hat{v}_1 \times \hat{w}_1\}$ がともに右手系か, とともに左手系であるように選ばれており, \hat{b} についても同様に選ばれている(図 8).

このとき, 式 (4.8) の 5 つの式は式 (4.4) $_{f=1,2}$ と同値であり, 5 つの式はそれぞれ独立な式である. よって $\langle \hat{a}, \hat{b} \rangle_Q$ の値が与えられることと計量行列 Q が定まることは同値となる. 一般に $\langle \hat{a}, \hat{b} \rangle_Q^2 \leq |\hat{a}|_Q^2 |\hat{b}|_Q^2 = 1$ であるから

$$(4.9) \quad \langle \hat{a}, \hat{b} \rangle_Q = \cos 2\theta, \quad 0 < 2\theta < \pi$$

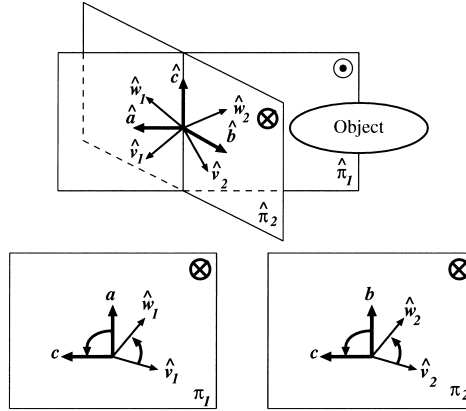


図 8. 新しい基底の選び方.

のようにパラメータ表示することができる(2枚の仮想画像面が異なるので $2\theta \neq 0, \pi$ である).
 このとき, $P = (\hat{a}, \hat{b}, \hat{c})$ とおくと

$$Q = P^{-T} \begin{pmatrix} 1 & \cos 2\theta & 0 \\ \cos 2\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} P^{-1} = P^{-T} K \text{diag}\{2 \cos^2 \theta, 2 \sin^2 \theta, 1\} K^T P^{-1}, \quad (4.10)$$

$$K = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

であるから,

$$A = P^{-T} K \text{diag}\{\sqrt{2} \cos \theta, \sqrt{2} \sin \theta, 1\} U_3, \quad U_3 \in O(3) \quad (4.11)$$

が成立する. ここで $X^{-T} = (X^T)^{-1}$ である. このように, 2枚の画像から得られる計量は1つのパラメータを含む. ここで, U_3 の与える影響は最終的に得られる解の世界座標の正規直交基底の選び方の自由度を与えるに過ぎないので, いま, U_3 として, $I_3 \in O(3)$ を選んでも一般性を失わない. そしてこのとき, $\{M, S^*\}$ は1つのパラメータ θ によって

$$\begin{aligned} M(\theta)^T &= \text{diag}\{\sqrt{2} \cos \theta, \sqrt{2} \sin \theta, 1\} K^T P^{-1} \hat{M}^T, \\ S^*(\theta) &= \text{diag}\{1/\sqrt{2} \cos \theta, 1/\sqrt{2} \sin \theta, 1\} K^{-1} P^T \hat{S}^* \end{aligned} \quad (4.12)$$

のようにパラメータ表示される. ここで

$$\begin{aligned} M(\theta)^T &= \text{diag}\{\sqrt{2} \cos \theta, \sqrt{2} \sin \theta, 1\} M(\pi/4)^T, \\ S^*(\theta) &= \text{diag}\{1/\sqrt{2} \cos \theta, 1/\sqrt{2} \sin \theta, 1\} S^*(\pi/4) \end{aligned} \quad (4.13)$$

が成立するので運動行列 $M(\theta)^T$ は $M(\pi/4)^T$ を $\text{diag}\{\sqrt{2} \cos \theta, \sqrt{2} \sin \theta, 1\}$ によってアフィン変換したものであり, 形状行列 $S^*(\theta)$ は $S^*(\pi/4)$ を $\text{diag}\{1/\sqrt{2} \cos \theta, 1/\sqrt{2} \sin \theta, 1\}$ によってアフィン変換したものであることがわかる.

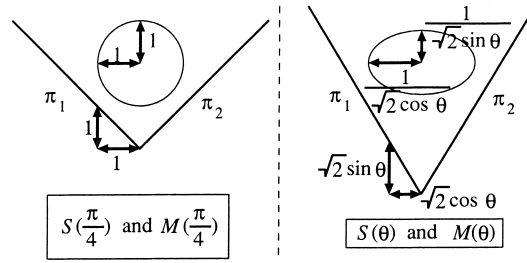


図9. 2画像からの復元解の一般解と $\theta = \pi/4$ の解.

ここで, $\hat{c} \in \text{span}\{\hat{v}_1, \hat{w}_1\} \cap \text{span}\{\hat{v}_2, \hat{w}_2\}$ は復元解において2枚の仮想画像面の交線方向のベクトル

$$(4.14) \quad c = \text{diag}\{\sqrt{2} \cos \theta, \sqrt{2} \sin \theta, 1\} K P^{-1} \hat{c} = (0, 0, 1)^T$$

となり, $a \in \text{span}\{\hat{v}_1, \hat{w}_1\}$, $b \in \text{span}\{\hat{v}_2, \hat{w}_2\}$ は復元解において, それぞれの仮想画像面上のベクトル

$$(4.15) \quad a = (-\cos \theta, \sin \theta, 0)^T, \quad b = (\cos \theta, \sin \theta, 0)^T$$

となる. 今 $|a| = |b|$, $a \perp c$, $b \perp c$ が成立することに注意すると, 2枚の仮想画像面のなす角の2等分面(2枚ある)上の c に垂直なベクトルの方向の単位ベクトルとして

$$(4.16) \quad u = (1, 0, 0)^T, \quad v = (0, 1, 0)^T$$

を選ぶことができる. つまり, 運動行列 $M(\theta)^T$ は $M(\pi/4)^T$ を u, v, c 方向にそれぞれ $\sqrt{2} \cos \theta$, $\sqrt{2} \sin \theta$, 1 倍拡大したものとなり, 形状行列 $S^*(\theta)$ は $S^*(\pi/4)$ を u, v, c 方向にそれぞれ $1/\sqrt{2} \cos \theta$, $1/\sqrt{2} \sin \theta$, 1 倍拡大したものとなる(図9). このとき, 2枚の仮想画像面のなす角は 2θ であり, c 方向の長さは θ によらず一定となることから, 真の解の縮尺と一致することに注意しておく.

4.3 複数のMAP画像からの復元解が一意に定まる条件

本章冒頭で述べたように, 点対応を用いた複数の2次元画像からの3次元形状復元問題において Necker Reversal の不定性を除いて解が一意に定まるためには, 計量行列 Q が一意に定まることが必要十分である. そして4.1で述べたように, 1枚の画像は3次元計量線型空間の2次元線型部分空間の部分計量を与えることから, 点対応を用いた複数の2次元画像からの3次元形状復元問題において Necker Reversal の不定性を除いて解が一意に定まることと, 複数の2次元線型部分空間の部分計量から3次元計量線型空間の計量行列を一意に決定できることは同値となる. さらにこの条件は, 計量行列の2次形式が表現する楕円体 $r^T Q r = 1$ を決定するためにはこの楕円体の(原点を通る)平面による切口が何枚必要か, という条件と同値となる.

楕円体を決定するために切口の楕円が何枚必要であるかという視点から見ると, 前章で示した λ_{f*} の比の値は, それぞれの楕円が交わるための条件として得られ, 2つの楕円が交わるようにそれぞれの楕円を拡大する倍率の比が λ_{f*} の比に対応することがわかる. また, 2つの交わる楕円を含む楕円体が一意に決まらない(図10(a))ことから2枚のMAP画像からカメラ運動と立体形状が一意に定まらないことも容易に理解することができる. さらに3つの交わる楕円が与えられれば楕円体が一意に決まる(図10(b))ことから3枚の(仮想画像面が異なる)MAP画像が与えられればカメラ運動と立体形状を一意に定めることができることがわかる. つまり, カ

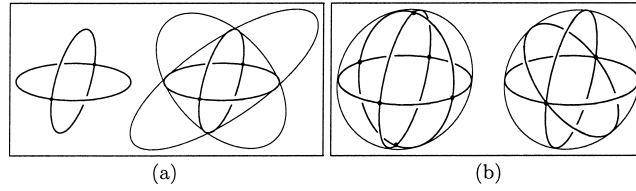


図 10. 形状復元には何枚の画像が必要か.

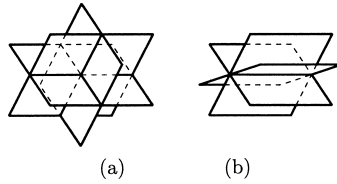


図 11. 互いに異なる 3 平面の位置関係.

メラ運動と立体形状の復元解が一意に定まるためには、仮想画像面の異なる 3 枚以上の MAP 画像が必要である(藤木・蔵田(1997b)). これは Ullman (1979) によって別の手法により証明された正射影の場合の拡張となっている. これを式で証明すると次のようになる.

互いに異なる 3 枚の仮想画像面 $\hat{\pi}_i (i = 1, 2, 3)$ の位置関係は図 11 の 2 通りに限られる.

(a) $\hat{\pi}_1 \cap \hat{\pi}_2 \cap \hat{\pi}_3$ が 1 点のとき

$\hat{e}_1 \in \hat{\pi}_2 \cap \hat{\pi}_3, \hat{e}_2 \in \hat{\pi}_3 \cap \hat{\pi}_1, \hat{e}_3 \in \hat{\pi}_1 \cap \hat{\pi}_2$ をみたく $\hat{e}_i (\neq 0) (i = 1, 2, 3)$ を選ぶことができ、このとき $|\hat{e}_1|_Q, |\hat{e}_2|_Q, |\hat{e}_3|_Q, \langle \hat{e}_1, \hat{e}_2 \rangle_Q, \langle \hat{e}_2, \hat{e}_3 \rangle_Q, \langle \hat{e}_3, \hat{e}_1 \rangle_Q$ の値が決まるので Q は一意に決定される.

(b) $\hat{\pi}_1 \cap \hat{\pi}_2 \cap \hat{\pi}_3$ が直線のとき

$\hat{e} \in \hat{\pi}_1 \cap \hat{\pi}_2 \cap \hat{\pi}_3, \hat{e}_i \in \hat{\pi}_i, \hat{e}_i$ は \hat{e} に平行でない、各 \hat{e}_i は同一平面上、をみたく $\hat{e}, \hat{e}_i (\neq 0) (i = 1, 2, 3)$ を選ぶことができる. 今、 $\hat{e}_3 = e_1 \hat{e}_1 + e_2 \hat{e}_2$ なる $e_1, e_2 \neq 0$ が存在するので、 $|\hat{e}_3|_Q^2 = e_1^2 |\hat{e}_1|_Q^2 + 2e_1 e_2 \langle \hat{e}_1, \hat{e}_2 \rangle_Q + e_2^2 |\hat{e}_2|_Q^2$ から $\langle \hat{e}_1, \hat{e}_2 \rangle_Q$ の値が求められ、 $|\hat{e}|_Q, |\hat{e}_1|_Q, |\hat{e}_2|_Q, \langle \hat{e}, \hat{e}_1 \rangle_Q, \langle \hat{e}, \hat{e}_2 \rangle_Q, \langle \hat{e}_1, \hat{e}_2 \rangle_Q$ の値が決まるので Q は一意に決定される(証明終).

5. 計測行列 W^* の特異値

因子分解法における計測行列の暫定的な分解において特異値分解が頻繁に用いられる. 画像が MAP によって得られ、観測時に誤差がない場合の計測行列 W^* のランクは 3 (以下)であり、この場合は $W^*_{(2F \times P)} = \hat{M}_{(2F \times 3)} \hat{S}_{(3 \times P)}$ の分解においてランク標準形を用いても、特異値分解を用いても全く同等な復元解が得られる. しかし実際に得られる画像は透視射影によって得られた画像であるために、透視射影によって得られた画像を MAP によって得られた画像であると仮定することによって生じるモデルの近似誤差、及びトラッキングや画像座標の量子化などによって生じる計測誤差が含まれ、計測行列 W^* のランクは一般に 4 以上となる. そこで因子分解法を実現するためには誤差を含むためにランクが 4 以上となった計測行列からランク 3 の真の計測行列を推定する必要がある.

この推定における基準として最小 2 乗基準、すなわち、観測された修正計測行列とのフロベ

二ウス距離が最小であるという基準を用いると、良く知られているように、求める推定量は、計測行列を特異値分解し、特異値を大きい順に並べ、4 番目以降の特異値を 0 とすることによって得られる。

特異値分解が主成分分析の表現を与えることに注意すると、この推定は、計測行列 W^* の行空間を主成分分析して、第 3 主成分までが真の修正計測行列の成分であり第 4 主成分以下をモデルの近似やトラッキングエラーなどに起因するノイズとみなして取り除いたことに相当する。

この視点からすると、計測行列の特異値は、因子分解法の復元結果の信頼性を考察する上で重要な値であるといえる。何故なら、透視射影を MAP で近似したときの近似誤差が小さければ計測行列 W^* のランクの 3 らしさは大きくなる、つまり計測行列の第 3 特異値までの寄与率が大きくなるのが期待でき、因子分解法を用いて得られた復元結果が良好であると期待できるからである。もちろん、透視射影を MAP で近似したときの近似誤差が大きい場合でも計測行列の第 3 特異値までの寄与率が大きくなる可能性があるが、少なくとも計測行列の第 3 特異値までの寄与率が小さい場合は透視射影を MAP で近似したことが妥当であるとは言えず、一般に良好な復元結果は期待できない。

計測行列の特異値と因子分解法の復元結果の関係に関する精密な理論は今後明らかにされるべき問題であるが、実際に因子分解法を適用している範囲においては、計測行列の第 3 特異値までの寄与率と復元精度は密接な関係があることが経験的に確かめられている。同様な理由により第 4 特異値と第 3 特異値の比も因子分解法による復元結果の信頼性の尺度となりうる。

このように、現状では理論的には不完全な面もあるが、特異値の評価は因子分解法の安定性を議論する上で非常に重要な量となる。そこで何が計測行列の特異値を定めるかについて考えることにする。

計測行列 W^* は $W^* = MS^*$ のように分解されることから、修正計測行列 W^* の特異値は M, S^* に依存することは明らかである。そこで本節では、計測行列 W^* の特異値が M, S^* のどの部分に依存するかについて明らかにする。

一般に行列の積 AB と BA が共に定義されるとき、 AB と BA の 0 でない固有値は重複度もこめて一致するので計測行列 $W^{*T}W^*$ の 0 でない固有値は $(M^T M)(S^{*T} S^*)$ の 0 でない固有値、つまり一般固有方程式

$$\Phi_{M^T M}^{(S^* S^{*T})^{-1}}(\mu) = \det(M^T M - \mu(S^* S^{*T})^{-1}) = 0$$

の 0 でない解に重複度もこめて一致する。

計測行列 W^* の 0 でない特異値は $W^{*T}W^*$ の 0 でない固有値の正の平方根であるから、計測行列 W^* の特異値は $M^T M, S^* S^{*T}$ に依存することがわかる。ここで特徴点の相対世界座標 s_p^* の共分散行列は

$$(5.1) \quad \text{Cov}(s_p^*) = \frac{1}{P} S^* S^{*T}$$

であるから、 $S^* S^{*T}$ は特徴点の相対世界座標の分布の共分散行列の定数倍となる。

同様にして $M^T M$ は射影方向の分布の共分散行列の定数倍となるが、 M の各行の長さは 1 とは限らないため、射影方向の重みつき分布の共分散行列の定数倍となる。具体的には、

$$(5.2) \quad M^T M = \sum_{f=1}^F \frac{1}{\lambda_{f^*}^2} \{p_f^2 v_f v_f^T + q_f^2 w_f w_f^T\}$$

が成立するので、 $M^T M$ は $\{v_f, w_f\}_{f=1}^F$ の方向の重みが $\{p_f/\lambda_{f^*}, q_f/\lambda_{f^*}\}_{f=1}^F$ であるような重みつき分布の共分散行列の定数倍となる。

よって計測行列の特異値は特徴点(の相対世界座標)の分布の共分散行列(の定数倍) $S^* S^{*T}$ の逆行列 $(S^* S^{*T})^{-1}$ に関する射影方向の(重みつき)分布の共分散行列(の定数倍) $M^T M$ の一

般固有値の正の平方根となる。すなわち、計測行列の特異値は特徴点の分布を正規化したときの射影方向の(重みつき)分布の主成分となることがわかる。これによって $W^{*T}W^*$ は特徴点(の相対世界座標)の分布を正規化したときの射影方向の(重みつき)分布の共分散行列(の定数倍)であることがわかる。

本章を終えるにあたって念のために強調しておくが、特異値分解によって最適な推定量が得られるというのは最小 2 乗誤差を最小にするという基準での話であり、この推定量が別の基準で最適な推定量になるとは限らない。例えば、最小 2 乗推定量がアウトライヤーに敏感であることから、トラッキングエラーが非常に大きい場合、最小 2 乗基準で推定することは良くないことが知られている。そのため、LMeds 推定などのロバスト推定により計測行列を推定する手法も提案されている(蔵田 他(1999b))。

6. 透視射影モデルの因子分解法

前章までで、MAP モデルの因子分解法とその数理について解説してきた。MAP モデルの因子分解法は線型手法により構成されているため数値計算上安定な優れた手法である。しかし、MAP モデルの因子分解法は透視射影の近似射影であるためにモデルの近似誤差が不可避であり、高精度の復元解を得るためには、計測行列がアフィン近似が成立する(第 2 章参照)ような画像により構成されている必要があるという欠点がある。この欠点を克服するために因子分解法を透視射影モデルの場合に拡張する手法が提案されている(Christy and Horaud (1996), Strum and Triggs (1996), Ueshiba and Tomita (1998), 金谷 (1998))。

MAP モデルの場合、奥行きパラメータはアフィン近似の中心に集約されたため、各画像につき 1 つであった。しかし、透視射影の場合は奥行きパラメータは各画像の各特徴点ごとに必要であり奥行きパラメータ(射影的奥行きと呼ばれる)の数は MAP モデルに比べて飛躍的に増大する。これが透視射影によるカメラ運動と立体形状の復元が難しくなる大きな理由である。それゆえ、透視射影によるカメラ運動と立体形状の復元を実現するためには、何らかの形で射影的奥行きを推定しなければならない。

透視射影によるカメラ運動と立体形状の復元には射影的奥行きの推定が必要であるという事実は、因子分解法の枠組では計測行列が射影的奥行きを含む形で定義されるというところに反映されており、透視射影によるカメラ運動と立体形状の復元を因子分解法で実現するためには、あらかじめ射影的奥行きを推定しておくか、射影的奥行きを反復的に推定しながら因子分解を行わなければならない。

Strum and Triggs (1996) はエピポーラ幾何学を利用して射影的奥行きの比をあらかじめ推定した後計測行列を求めることにより、反復計算を用いずに射影的復元を行なう因子分解法を提案した。しかし、エピポーラ幾何学から推定された射影的奥行きは特徴点の計測誤差に対して敏感であり、復元結果が計測誤差に対して安定とはいえないという欠点をもつ。これはすべての画像を均等に扱うことにより、特定の画像の信頼性が低くても、全体として安定な復元結果が得られるというアフィン射影モデルの因子分解法の利点を活かしているとは言えない。

Ueshiba and Tomita(1998) はすべての画像をほぼ均等に扱うような評価関数を利用し、射影的奥行きを含む計測行列を推定することにより射影的復元を行なう因子分解法を提案した。また、カメラがキャリブレーションされている場合に射影的復元形状からユークリッド復元形状を求める手法も提案した。しかし、この手法は射影的奥行きを含んだ計測行列の推定のための反復計算の収束が遅いという欠点がある。

金谷 (1998) も 2 枚の透視射影画像が得られれば射影的奥行きの比が推定できることから、あらかじめ射影的奥行きの比を推定して反復計算を用いずに射影的復元形状、ユークリッド復元形

状を求める手法を提案した．この手法は本人も指摘しているように，計測行列の特異値分解に相当する部分を固有値問題に書き直したものであり，この手法によって得られる結果は Strum and Triggs (1996) などの手法と同等なものである．

Christy and Horaud (1996) はユークリッド復元解と射影的奥行きを交互に反復推定することによって透視射影によるユークリッド復元解を推定する手法を提案した．この手法は透視射影モデルによる像から paraperspective モデルによる像を反復推定し，paraperspective モデルからの因子分解法によって透視射影によるユークリッド形状を推定することによって運動と形状を高精度で復元する手法であるとも言える．この手法は反復計算の収束が速いという特徴があるものの，反復計算において形状のユークリッド復元が前提となっているため，カメラがキャリブレートされている場合のみ有効な手法である．

本章では，Christy and Horaud (1996) によって提案された透視射影による像から paraperspective 射影による像を反復的に推定する手法を因子分解法の枠組で説明する．

第 f 画像面上に第 p 特徴点を透視射影した画像座標を x_{fp}^{per} ，第 f 画像面上に第 p 特徴点を paraperspective 射影した画像座標を x_{fp}^{para} とすると， $x_{fp}^{\text{per}} = x_{fp}^{\text{per}} - x_{f*}$ と $x_{fp}^{\text{para}} = x_{fp}^{\text{para}} - x_{f*}$ の間には図 12 により

$$(6.1) \quad x_{fp}^{\text{para}} = \mu_{fp}^* x_{fp}^{\text{per}}, \quad \mu_{fp}^* = \frac{\lambda_{fp}}{\lambda_{f*}}$$

という関係が成立する(数式による導出は Ueshiba and Tomita (1998) 参照)．ここで λ_{fp} は射影的奥行き (projective depth) と呼ばれる(グローバルスケールパラメータを含んだ)，第 f 画像における第 p 特徴点の奥行きであり，このとき μ_{fp}^* は第 f 画像における特徴点 $*$ を基準とした第 p 特徴点の相対射影的奥行きとなる．

Christy and Horaud (1996) によって提案された手法は式 (6.1) を用いて透視射影の画像座標から paraperspective 射影の画像座標を反復的に推定することによって paraperspective 射影モデルの因子分解法の運動と形状の復元精度を高める手法である．ここで射影的奥行きを含む計測行列 W^* を

$$(6.2) \quad W^* = \begin{pmatrix} \mu_{11}^* x_{11}^* & \dots & \mu_{1P}^* x_{1P}^* \\ \vdots & \ddots & \vdots \\ \mu_{F1}^* x_{F1}^* & \dots & \mu_{FP}^* x_{FP}^* \end{pmatrix}$$

とする．ここで $x_{fp}^* = x_{fp} - x_{f*}$ は実際に観測された値であり，透視射影に従うと仮定して良い．

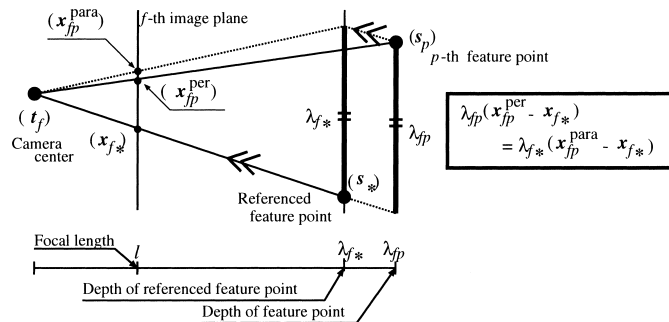


図 12. 透視射影と paraperspective 射影.

このとき Christy and Horaud の手法のアルゴリズムは以下のようにして実現される .

- (0) $\mu_{fp}^* = 1$ とおき, paraperspective 射影モデル因子分解法によりユークリッド復元を行なう . このとき, 一方の復元解を $M^{(+)}, S^{*(+)}$, その鏡映解を $M^{(-)}, S^{*(-)}$ とし, それぞれの復元解 $M^{(\pm)}, S^{*(\pm)}$ に対し, 式 (6.1) を用いて $\mu_{fp}^{*(\pm)}$ を計算する (以下, 複号同順).
- (1) 計測行列 $W^{*(\pm)}$ を更新し, paraperspective 射影モデルの因子分解法によるユークリッド復元を行なう .
- (2) それぞれの $W^{*(\pm)}$ から得られる 2 組の復元解のうち $M^{*(\pm)}, S^{*(\pm)}$ の構造が鏡映でないものを用いて $M^{(\pm)}, S^{*(\pm)}$ を更新する .
- (3) 更新された $M^{(\pm)}, S^{*(\pm)}$ に対し, 式 (6.1) を用いて $\mu_{fp}^{*(\pm)}$ を更新する .
- (4) $\mu_{fp}^{*(\pm)}$ の変化が適当に定めた閾値 ϵ 以下でなければ (1) に戻る .
- (5) $M^{(\pm)}$ から計算された外部パラメータを用いて透視射影による画像座標 \hat{x}_{fp}^* を再計算して $M^{(+)}, S^{*(+)}$ 及び $M^{(-)}, S^{*(-)}$ のうち, 観測された画像座標とのずれ

$$\sum_{f,p} \|\hat{x}_{fp}^* - x_{fp}^*\|^2$$

を小さくする方を復元解として選ぶ .

以上のアルゴリズムによって, 同一平面上にない 4 点の異なる 3 枚の画像が得られれば, カメラ運動と立体形状を 1 通りに定めることができる .

7. 逐次型因子分解法

これまでに述べた MAP 画像の因子分解法や透視射影の因子分解法は, 処理すべき全ての画像を同時に必要とするため, 画像枚数が増加するに従って計算時間が増大し, 実時間処理に適用するのは容易ではない . そこで, これらの因子分解法を実時間処理に適用するために逐次型因子分解法が提案された (Morita and Kanade (1997), 藤木 他 (1998)).

Morita and Kanade (1997) は, 形状行列の行空間が計測行列の行空間に等しいことを利用し, 計測行列から形状行列の行空間の正規直交基底の近似値を高速に求める手法を用いて運動と形状を復元した . しかし, この近似は初期段階では良くないため, 初期段階での復元結果の信頼性は低いという欠点がある .

藤木 他 (1998) は, 主成分分析を用いて運動行列の持つ運動情報を圧縮して扱う計測行列のサイズを小さくすることによって, すべての画像をなるべく均等に扱いながら計算時間を短縮することを実現した .

本章では, 藤木 他 (1998) の逐次型因子分解法について説明する . なお, 本章で説明する逐次型因子分解法と Christy and Horaud (1996) の手法を組み合わせることにより, 透視射影モデルの逐次型因子分解法のアルゴリズムを得ることができる (Fujiki and Kurata (2000)).

7.1 主成分分析による運動情報の圧縮

バッチ処理の因子分解法においては画像枚数の増加に比例して計算時間も増大するため, 因子分解法を実時間処理に適用するためには復元情報を圧縮する必要がある . 本章では, 本稿で提案する MAP モデルの逐次型因子分解法を実現するために不可欠な, 計算時間の短縮のための運動情報の圧縮について説明する .

既に述べたように計測行列 W^* を暫定的に分解した時点で運動と形状のアフィン復元解が得られるのでアフィン復元情報を圧縮するには計測行列 W^* を圧縮すれば良い . 一般に計測行列 W^* のランクは 3 であるから P 次元行ベクトル空間の $2F$ 個ベクトルが張る 3 次元部分空間の

基底をなすベクトルを選び出してやれば良い．例えば， W^* の互いに線型独立な 3 つのベクトルを選んだり， S^* をそのまま用いたり， W^* から主成分分析により第 3 主成分ベクトルまでを選んだりすることが考えられる．

しかし，ユークリッド復元情報を圧縮するためにはアフィン復元解からユークリッド復元解を導くために必要な計量拘束を考慮して計測行列の圧縮を行なわなければならない．ところで計量拘束は式 (3.6) のような運動行列の各行の内積の値(の一部)を列挙した条件であるから，運動行列を圧縮することにより計測行列の圧縮を行なえば良い．そしてそのためには，一般に運動行列 M のランクは 3 であるから 3 次元行ベクトル空間の基底をなすベクトルを選べば良い．

ここで選び出された 3 次元行ベクトル空間の基底をなすベクトルを l_1^T, l_2^T, l_3^T とし， $M = (l_1, l_2, l_3)^T$ とすると圧縮された計測行列 W^* 及び計量条件は

$$W^* = MS^*, \quad MM^T = \begin{pmatrix} l_1^T l_1 & l_2^T l_1 & l_3^T l_1 \\ l_1^T l_2 & l_2^T l_2 & l_3^T l_2 \\ l_1^T l_3 & l_2^T l_3 & l_3^T l_3 \end{pmatrix}$$

によって表現される．このような M の構成の方法としては，例えば， M の互いに線型独立な 3 つのベクトルを選んだり，単位行列 I_3 を選べば良い．しかし，本稿ではすべての画像を同等に扱うことにより特定の画像に誤差が含まれていても全体としては安定な復元解を得ることができるというバッチ処理の因子分解法の特徴を失わないよう，以下の理由により主成分分析を用いて第 3 主成分ベクトルまでを選ぶ方法を提案する．なお，3 次元行ベクトルを主成分分析するので主成分は高々 3 つしかないことからすべての主成分及び主成分ベクトルを選ぶことになる．

式 (3.5) により，計測行列の各行は形状行列を運動行列の各行の方向へ射影したものである．よって運動行列の各行の分布は，立体形状を射影した方向，すなわち立体形状をどの方向から観測したかの分布と考えることができる．一般に多く観測された方向の情報ほど信頼度が高いと考えることができるため，運動行列の各行の分布は，復元された立体形状をその方向から見たときの信頼度を表す一つの尺度と考えることができる．したがって，すべての画像を同等に扱うバッチ処理の因子分解法の特徴を失わないためには，運動行列の各行の分布を保存するように運動情報を圧縮すれば良い．その圧縮が主成分分析を用いることにより実現されることを以下に述べる．

ここで，運動行列の各行が期待値がゼロベクトルである多重正規分布(立体形状を射影する方向として任意の方向が同等に選ばれと仮定することにより運動行列の各行の 3 次元行ベクトルの期待値はゼロベクトルであるとすると)に従うことを仮定する．このとき運動行列 M の各行の分布は運動行列 M の共分散行列によって表現される．より詳しく言えば，運動行列 M の各行の分布から得られる共分散行列の定数倍である $M^T M$ によって運動行列 M から導かれるすべての統計的性質が記述できる．すなわち $M^T M$ は十分統計量 (sufficient statistic) であり，運動行列 M の各行のもつ標本情報のすべてが含まれる．

したがって運動行列 M を

$$(7.1) \quad M^T M = M^T M$$

をみたとすような M に置き換えても運動行列 M の各行のもつ標本情報は失なわれまいと考えることができる．

つまり，バッチ処理の因子分解法のもつ，すべての画像を同等に扱うことにより特定の画像に誤差が含まれていても全体としては安定な復元解を得ることができるという特色を失わないためには，式 (7.1) をみたとすような運動空間の基底を選び出せば良く，このような基底は運動行列 M を主成分分析することによって得られる．

主成分分析の表現は特異値分解によって得られる．運動行列 M の特異値分解を

$$(7.2) \quad M = FAE$$

とすると， E は主成分ベクトルを並べた行列， A は M の主成分を並べた対角行列を表す．ここで F は射影する方向の AE の各行を基底としたときの表現を表すが， F は実際に射影する方向を選び出すという操作に相当し，運動空間の統計的性質に対して本質的ではないので無視することができる．つまり運動空間の統計的性質は

$$(7.3) \quad M = AE$$

によって記述される．実際，主成分分析によって得られた M に対して明らかに $M^T M = M^T M$ が成立するので， M と M のもつ運動情報は同等であると考えることができる．つまり 3×3 行列 M によって $2F \times 3$ 行列 M の運動情報が圧縮されたと考えることができる．

なお，一般に 3×3 行列 X に対する方程式 $X^T X = M^T M$ の一般解は $X = U^T M$ ， $U \in O(3)$ であるから，運動空間の統計的性質を保存するような M の選び方は本質的に 1 通りであることがわかる．

この M を主成分運動行列と呼ぶ．このとき主成分運動行列に対する計測行列 (主成分計測行列と呼ぶ) 及び計量拘束は

$$(7.4) \quad W^* = MS^*, \quad MM^T = A^2$$

と表現される．

この圧縮により，バッチ処理の因子分解法では新たに第 $F+1$ 画像が計測された場合 $(2F+2) \times P$ 行列 $(W^{*T}, W_{F+1}^{*T})^T$ を因子分解しなければならないが，運動行列を圧縮することにより $5 \times P$ 行列 $(W^{*T}, W_{F+1}^{*T})^T$ を因子分解するだけで済む．このことにより大幅な計算時間の短縮が実現され，実時間処理への応用が期待できる．

また，第 5 章で述べたように MAP モデルの因子分解法が適用可能か，すなわち透視射影が MAP モデルと近似して良いかを判定する一つの基準として計測行列 W^* の第 3 特異値までの寄与率が挙げられる．計測行列の第 3 特異値までの寄与率は，いわば計測行列のランク 3らしさを表す．そして計測行列の第 3 特異値までの寄与率が高いほど透視射影によって得られた画像を MAP モデルによって得られた画像とみなすことの妥当性が保証され，カメラ運動と立体形状の復元精度も高いと考えることができる．

本稿で提案した運動情報の圧縮において得られる主成分計測行列 $W^* = MS^*$ の特異値は $W^* = MS^*$ の特異値と重複度をこめて一致する (実際は誤差の影響により完全には一致しない)．なぜなら

$$(7.5) \quad W^{*T} W^* = S^{*T} M^T M S^* = S^{*T} M^T M S^* = W^{*T} W^*$$

が成立し， W^* ， W^* の特異値は $W^{*T} W^*$ ， $W^{*T} W^*$ の固有値の正の平方根として得られるからである．よって，計測行列を主成分計測行列に置き換えても，MAP モデルの因子分解法が適用可能かどうかの度合は変化しないことがわかる．

なお，主成分分析を用いた運動情報の圧縮の幾何学的意味は，運動行列 M の共分散行列 (の定数倍) から導かれる確率密度が一定の等高面である楕円面を表現するために，運動行列の行空間の基底として楕円面の 3 つの軸方向の単位ベクトルと，それら軸の長さを選び出したことに相当し， E の各行は楕円面の軸方向の単位ベクトル， A の各対角成分は楕円面の軸の長さを表している．

7.2 MAP モデルの逐次型因子分解法

前章で述べたような運動情報の圧縮を利用することによって、因子分解法を逐次型に改良することができる(藤木 他(1998)). また, Christy and Horaud (1996) と組み合わせることにより, 透視射影を考慮した paraperspective 射影モデルの逐次型因子分解法を構成することができる (Fujiki and Kurata (2000)).

既に述べたように, 画像座標からカメラ運動と立体形状を復元するときはそれらの相対的な位置関係しか復元することができない. そのため, 同じ運動と形状の復元結果であっても座標系によって表現が異なる. そこで逐次型因子分解法においては運動と形状を固定した世界座標系で表現する必要がある. そこで, 2 つの形状行列を結び直交行列を最小 2 乗推定することによって, 異なる世界座標系で表現された運動と形状の復元解を同じ世界座標系で表現することにする.

ここで, 形状行列の, ある世界座標系における表現及び別の座標系における表現を, それぞれ $S_{\text{ref}}^*, S_{\text{ob}}^*$ とし, $S_{\text{ref}}^* S_{\text{ob}}^{*\text{T}}$ の特異値分解を $S_{\text{ref}}^* S_{\text{ob}}^{*\text{T}} = UDV^{\text{T}}$ とすると, S_{ob}^* を基準世界座標系に変換する直交行列, つまり $S_{\text{ref}}^* = \mathcal{E} S_{\text{ob}}^*$ をみたく直交行列 \mathcal{E} の最小 2 乗推定値は,

$$(7.6) \quad \mathcal{E} = UV^{\text{T}}$$

によって得られる.

逐次型因子分解法のアルゴリズムは (0) 3 枚以上の画像から運動と形状を復元(この復元解を記述する座標系に世界座標系を固定). (1) 運動情報を圧縮し, 主成分計測行列と新しく観測された画像から $5 \times P$ の計測行列(逐次計測行列と呼ぶことにする)を構成. (2) 固定した世界座標系における運動及び形状を復元し, (1) に戻る. の順に行なわれる. 以下に述べる詳細なアルゴリズムは, MAP 画像を正射影画像に帰着した形で書き下してある. 正射影画像に帰着していない形のアルゴリズムは藤木 他 (1998) を参照のこと.

- (0) $k(\geq 3)$ 枚の画像によって構成される計測行列 W^* から運動と形状をユークリッド復元する. なお, ここで得られた復元解を記述する座標系に世界座標系を固定する.
- (1) 運動情報の圧縮により第 $f(\geq k)$ 主成分運動行列 $\mathcal{M}_{[f]}$, 第 f 主成分計測行列 $\mathcal{W}_{[f]}$ 及び計量拘束 $\mathcal{M}_{[f]} \mathcal{M}_{[f]}^{\text{T}} = A_{[f]}^2$ を求め, 第 $f+1$ 画像から得られる計測行列 W_{f+1}^* 及び第 f 主成分計測行列 $\mathcal{W}_{[f]}^*$ から第 $f+1$ 逐次計測行列 $W_{[f+1]}^*$ を $W_{[f+1]}^* = (\mathcal{W}_{[f]}^{*\text{T}}, W_{f+1}^{*\text{T}})^{\text{T}}$ によって構成する. ここで第 f 画像のみから得られる諸量の添字と区別するために第 1 画像から第 f 画像までの情報を圧縮して得られる諸量には $[f]$ のような添字を用いている.
- (2) 因子分解法を用いてユークリッド復元し, 基準世界座標系における運動行列 $M_{[f+1]} = (\mathcal{M}_{[f]}^{\text{T}}, M_{f+1}^{\text{T}})^{\text{T}}$ 及び形状行列 $S_{[f+1]}^*$ を求める. このとき, 計量拘束は

$$\begin{aligned} \hat{\mathcal{M}}_{[f]} Q \hat{\mathcal{M}}_{[f]}^{\text{T}} &= \mathcal{M}_{[f]} \mathcal{M}_{[f]}^{\text{T}} = A_{[f]}^2 = \text{diag}\{\lambda_{[f],1}, \lambda_{[f],2}, \lambda_{[f],3}\}, \\ \hat{V}_{f+1} Q \hat{V}_{f+1}^{\text{T}} &= V_{f+1} V_{f+1}^{\text{T}} = I_2 \end{aligned}$$

であり, 計量拘束をみたく Q は

$$\hat{V}_{[f+1]} = \begin{pmatrix} \hat{\mathcal{M}}_{[f]} \otimes \hat{\mathcal{M}}_{[f]} \\ \hat{V}_{f+1} \otimes \hat{V}_{f+1} \end{pmatrix}, \quad \mathcal{I}_{[f+1]} = \begin{pmatrix} \text{cs}(A_{[f]}^2) \\ \text{cs } I_2 \end{pmatrix}$$

とおくと $\text{cs } Q = \hat{V}_{[f+1]}^+ \mathcal{I}_{[f+1]}$ によって求まる.

もしくは, $\hat{\mathcal{M}}_{[f]} = (\hat{\mathbf{i}}_{[f],1}, \hat{\mathbf{i}}_{[f],2}, \hat{\mathbf{i}}_{[f],3})^{\text{T}}$ とし, $\Omega_{[f+1]}$ と $\mathcal{L}_{[f+1]}$ を

$$\begin{aligned} \Omega_{[f+1]} &= (\omega_{[f],11}, \omega_{[f],22}, \omega_{[f],33}, \omega_{[f],12}, \omega_{[f],13}, \omega_{[f],23}, \hat{\omega}_{f+1})^{\text{T}}, \\ \omega_{[f],ij} &= \omega(\hat{\mathbf{i}}_{[f],i}, \hat{\mathbf{i}}_{[f],j}), \\ \mathcal{L}_{[f+1]} &= (\lambda_{[f],1}^2, \lambda_{[f],2}^2, \lambda_{[f],3}^2, \mathbf{0}_3^{\text{T}}, 1, 0, 1)^{\text{T}} \end{aligned}$$

とおくと $q = \Omega_{[f+1]}^+ \mathcal{L}_{[f+1]}$ によって求まる．ここで鏡映解の不定性は

$$\det S_{[f]} S_{[f+1]}^T > 0$$

をみだす組を選ぶことにより除くことができる． Q を用いて運動と形状を復元し (1) に戻る．

以上のアルゴリズムにより因子分解法を逐次型に改良することができた．因子分解法を逐次型に改良することにより，扱う画像枚数の増加に従って増大する計算時間を抑えることができる．このアルゴリズムはアフィン近似射影モデルの場合であるが，第 6 章で解説した手法と組み合わせることにより，透視射影の逐次型因子分解法を構成することもできる (Fujiki and Kurata (2000)) ．

8. おわりに

本稿では点対応を用いた複数の 2 次元画像からの 3 次元形状復元の理論について，主にアフィン近似射影の場合について因子分解法の立場から眺めてきた．本稿では述べることはできなかったが，点特徴を用いた複数物体の因子分解法 (Costeria and Kanade(1998), Gear(1998))，関節物体の因子分解法 (長崎 他(1998))，線特徴を用いた因子分解法 (Quan and Kanade (1997), Morris and Kanade (1998)) など提案されている．

参 考 文 献

- Christy, S. and Horaud, R. (1996). Euclidean reconstruction: From paraperspective to perspective, *Proceedings of 4th European Conference on Computer Vision*, **2**, 129–140.
- Costeria, J.P. and Kanade, T. (1998). A multibody factorization method for independently moving objects, *International Journal on Computer Vision*, **29**(3), 159–179.
- 藤木 淳 (2000). Paraperspective 射影画像からの運動と形状の 2 組の復元解の関係，電子情報通信学会 2000 年情報・システムソサエティ大会，D-12-50, p. 237.
- 藤木 淳，蔵田武志 (1997a). アフィンモデルにおける因子分解法の数理とその検証，電子情報通信学会技術研究報告，PRMU 97-22, 167–174.
- 藤木 淳，蔵田武志 (1997b). 一般アフィン射影モデルの因子分解法の数理，電子情報通信学会技術研究報告，PRMU 97-142, 101–108.
- Fujiki, J. and Kurata, T. (2000). Recursive factorization method for the paraperspective model based on the perspective projection, *Proceedings of 15th International Conference on Pattern Recognition*, **1**, 406–410.
- 藤木 淳，蔵田武志，田中 勝 (1998). 一般アフィン射影モデルの逐次型因子分解法，電子情報通信学会技術研究報告，PRMU 98-118, 45–52.
- Gear, C.W. (1998). Multibody grouping from motion images, *International Journal on Computer Vision*, **29**(2), 133–150.
- 金出武雄，コンラッドポールマン，森田俊彦 (1993). 因子分解法による物体形状とカメラ運動の復元，電子情報通信学会論文誌，**J76-D-II** (8), 1497–1505.
- 金谷健一 (1990). 『画像理解 — 3 次元認識の数理 —』，森北出版，東京．
- 金谷健一 (1998). 因子分解法を用いない因子分解法：平行投影から透視変換へ，電子情報通信学会技術研究報告，PRMU 98-26, 1–8.
- Koenderink, J. J. and van Doorn, A. J. (1991). Affine structure from motion, *J. Opt. Soc. Amer. A*, **8**(2), 377–395.

- 蔵田武志, 藤木 淳, 坂上勝彦 (1999a). アフィンエピポラ幾何の因子分解法による解釈, 情報処理学会論文誌, **40**(8), 3188-3197.
- 蔵田武志, 藤木 淳, 坂上勝彦 (1999b). ロバスト推定を用いた一般アフィン射影モデルの因子分解法, 電子情報通信学会技術研究報告, PRMU 98-194, 17-24.
- Morita, T. and Kanade, T. (1997). A sequential factorization method for recovering shape and motion from image streams, *IEEE Trans. Pattern Analysis and Machine Intelligence*, **19**(8), 858-867.
- Morris, D. D. and Kanade, T. (1998). A unified factorization algorithm for points, line segments and plane with uncertainty models, *Proceedings of International Conference on Computer Vision'98*, 696-702.
- Mundy, J. L. and Zisserman, A. (ed.) (1992). *Geometric Invariance in Computer Vision*, MIT press Cambridge, Massachusetts.
- 長崎 健, 川嶋稔夫, 青木由直 (1998). 因子分解法に基づく運動画像列解析による多関節物体の構造推定, 電子情報通信学会論文誌, **J81-D-II-3**, 483-491.
- Poelman, C. J. and Kanade, T. (1997). A paraperspective factorization method for shape and motion recovery, *IEEE Trans. Pattern Analysis and Machine Intelligence*, **19**(3), 206-218.
- Quan, L. (1996). Self-calibration of an affine camera from multiple views, *International Journal on Computer Vision*, **19**(1), 93-105.
- Quan, L. and Kanade, T. (1997). Affine structure from line correspondences with uncalibrated affine cameras, *IEEE Trans. Pattern Analysis and Machine Intelligence*, **19**(8), 834-845.
- Shi, J. and Tomasi, C. (1994). Good features to track, *Proceedings of Computer Vision and Pattern Recognition'94*, 593-600.
- Sturm, P. and Triggs, B. (1996). A factorization based algorithm for multi-image projective structure and motion, *Proceedings of 4th European Conference on Computer Vision*, **2**, 709-720.
- Tomasi, C. and Kanade, T. (1992). Shape and motion from image streams under orthography: A factorization method, *International Journal on Computer Vision*, **9**(2), 137-154.
- Ueshiba, T. and Tomita, F. (1998). A factorization method for projective and Euclidean reconstruction from multiple perspective views via iterative depth estimation, *Proceedings of 5th European Conference on Computer Vision*, 296-310.
- Ullman, S. (1979). *The interpretation of visual motion*, MIT Press, Cambridge, Massachusetts.
- Xu, G. and Sugimoto, N. (1998). A linear algorithm for motion from three weak perspective images using Euler angles, *Proceedings of Asian Conference on Computer Vision'98*, 543-550.
- Xu, G. and Zhang, Z. (1996). *Epipolar Geometry in Stereo, Motion and Object Recognition—A Unified Approach*, Kluwer, Dordrecht.
- Yokoya, N., Takemura, H. and Hwang, K. (1998). A factorization method using 3-d linear combination for shape and motion recovery, *Proceedings of International Conference on Pattern Recognition'98*, 959-963.

3D Reconstruction from Sequences of 2D Images
under Point Correspondences
— An Mathematical Analysis of the Factorization Method —

Jun Fujiki

(National Institute of Advanced Industrial Science and Technology)

Recovering the camera motion and the object shape from multiple images is a fundamental and important problem in the field of computer vision. Especially, the problem under point correspondences is the most fundamental and most important. To solve this problem, many methods are presented and among them, the factorization method is an excellent method because it is stable in numerical computation and it gives good reconstruction although it is based on the affine approximation of the perspective projection. The factorization method is useful not only for solving the problem, but also for understanding the mathematical meaning of the problem under affine approximated projection. In this paper, the mathematical analysis of recovering the camera motion and the object shape from multiple affine approximated projection images under point correspondences by the factorization method is considered. The way to recover the camera motion and the object shape from perspective images by estimating the affine approximated projection images from perspective images and the recursive factorization method are also considered.

Key words: Structure from motion, factorization method, recursive, Metric Affine Projection, perspective projection.