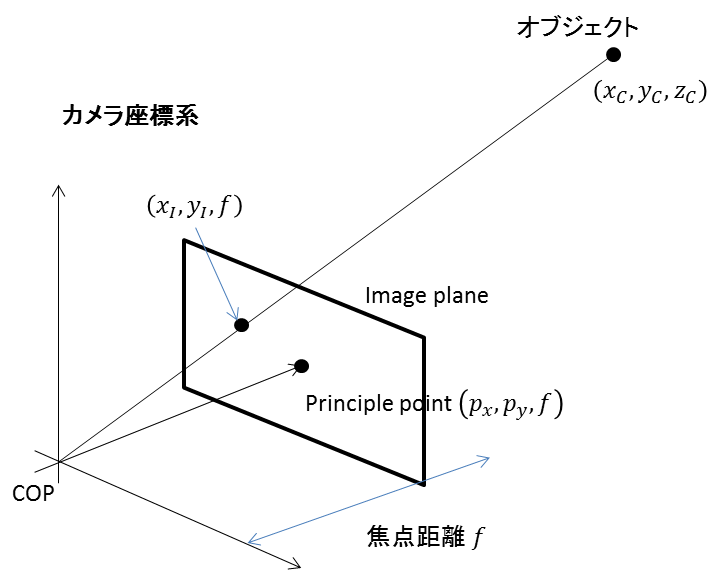
# Camera Calibration

## Intrinsic Calibration Matrix



まず、カメラのレンズの中心は**Center of Projection、つまりCOP**と呼ぶ。オブジェクトの、カメラ座標系での座標をとし、そのイメージプレーンへの写像点の、カメラ座標系での座標をとする。ただし、は焦点距離で、COPからイメージプレーンへの距離だ。また、**COP**からイメージプレーンへ下ろした垂線の足を、**principal point**と呼ぶ。Principal pointのカメラ座標系での座標をとする。この時、

よって、

これを、例によってHomogeneousなんとかの書き方にしてやると、

ここで、

が、カメラの**内部パラメータ行列**だ。しかし、レンズには一般的に歪みがあるので、もっと複雑になる。Radial distortionのみの歪を考える場合、半径方向に、半径の二乗に比例したズレが発生すると考えると、上の式は、次のように変わる。

よって、

同様に、homogeneousの書き方に変えると、

(1)

これが、より複雑な**内部パラメータ行列**で、スライドp.30の行列だ！

## Extrinsic calibration matrix

カメラ座標系とワールド座標系との関係を表す行列が、外部パラメータ行列だ。カメラが１個なら、カメラ座標系とワールド座標系を同じと見なしても良いんだけど、例えばカメラが２個ある場合などは、そうもいかないよね。

外部パラメータ行列は、単純に回転と平行移動の組合せで表される。つまり、ある点の、ワールド座標系での座標を、カメラ座標系での座標をとすると、

逆に言うと、カメラ座標が分かっている時、そのワールド座標系での座標は、

これを使えば、例えばprincipal point の、ワールド座標系での座標は、

で計算できる。

さて、先ほどの式は、例のhomegeneousの書き方だと、もっときれに書ける。

(2)

スライドp.31で、この外部パラメータ行列を

と呼んでいる。

## Optimization

内部パラメータ行列と外部パラメータ行列をまとめると、

(3)

まず、大きく分けて２ステップでoptimizationを行う。ステップ１では、、を画像の中心、とみなすことで、線型方程式に変換し、すばやく行列をestimateする。ステップ２では、元の非線型方程式を使って、estimateした行列を初期値として、非線形最小二乗法で最適解を探す。ポイントは、スライドp.49にある通り、**「非線形最小二乗法の収束は、初期値の精度に依存する」**だ。つまり、ステップ１で良い精度で初期値を求めることが重要となる。

**ステップ１（回転行列と、の推定）**

まず、とみなす。式(1)より、

ここで、、とおくと、

(4)

１行目から２行目を割ると、

パッド上の点の座標は0としても一般性は失わないので、。よって、

よって、

これを行列の形で表すと、

個の点があれば、この式を個積み上げていけば良い。この線型方程式の形は定数項がないので、**Homogeneous Linear System**と呼ばれ、左側の行列に対しての最小固有値に対応する固有ベクトルが、右側のベクトルの解となる。

スライドp.39では、**Cholesky分解**を使って解くと書いてあるが、Cholesky分解は正定値実対称行列にしか使えない。一般に、上の行列は縦に長い（overdetermined）ので、Cholesky分解は使えない。一般的に、**OverdeterminedのHomogeneous Linear Systemを解くには、SVD分解をし、行列の最後の列が解**となる。

ここで、ポイント！Homogeneous Linear Systemの解が得られたら、任意の実数でスケールしたやつも解だ。じゃあ、どうやってユニークな解を得るのか？それには、回転行列の各行がnormalizeされているという条件と、直交条件を使用すれば良い。つまり、SVDで得られた解に対して、

(5)

最初の２つの式を変形して、

(6)

上と下の式を掛けると、

(7)

また、３つ目の式を変形して、

両辺を二乗して、

(8)

式(7)から(8)を引くと、

(9)

ところで、式(6)より、

これを式(9)に代入して、

整理して、

(10)

SVDにより、、、は求まっているので、式(10)からを計算できる。

が求まったら、式(6)を使ってとを計算できる。後は、それぞれで割れば、回転行列が得られる。

なお、、については、式(5)を見ると分かるように二乗しているので、符号が＋か－か不明だ。式(5)の３つ目の式より、との積の符合は決定する。つまり、

しかし、これでもまだ２つの候補が残っていしまう。例えば、が正なら、とが共に正、または、共に負という２つの候補がある。これを見極めるためには、実際に3Dの座標を射影してみて、2Dの座標に近い方を選ぶしかない。なんだか、ヘタレだねぇ。。。具体的には、個の点について以下を計算し、正ならＯＫ。もし負なら、もう一方の符号にする。

ただし、は番目の観測された2D座標、は番目の射影された2D座標。

**※はまだ求まっていないので、適当な値を使うのか？**

以上で、回転行列の最初の２行のベクトルとが求まった。最後に、３つ目の行ベクトルは、単純にcross productで計算できる。つまり、

以上から、回転行列が得られる。なお、スライドでは、ステップ２以降の説明がない！？

**ステップ２（との推定）**

次に、とを計算する。まず、式(4)より、

式(2)を代入して、、を消去すると、

整理して、

思い出して欲しい。既に回転行列は得られているので、右辺は既知だ。つまり、との線形方程式になっている。一応、行列の形にしてあげると、

個の点があれば、この式を積み上げれば良い。の形になるので、ヤコビ法なり、ガウス・サイデル法でとを計算できる。

ここまでのステップにより、焦点距離と外部パラメータ行列が求まった。

**ステップ３（principal pointとの推定）**

最後に、estimateした行列と焦点距離を初期値とし、、を画像の中心、を初期値とし、以下の非線形方程式を使って、最小二乗法で最適解を探す。

なお、は番目の射影された2D座標で、式(3)により計算できる。

※この最適化は、たぶん、Minpackのlmdifを使って解くんだと思うけど。。。

## Zhangのアルゴリズム

まず、以下のモデルを使用する。

なお、

(1)

は、カメラ内部パラメータ行列だ。

**ステップ１（Homography行列を解く）**

パッドは、ワールド座標系の平面上にあると見なしても一般性を失わないので、上のモデルでとおく。すると、

つまり、

(2)

よって、

となる。ここで、を**ホモグラフィ行列**と呼ぶ。ホモグラフィとは、２次元の座標から２次元の座標への変換のことだ。あとは、一般的なホモグラフィの解法を使って、ホモグラフィ行列を求めれば良い。具体的には、まず、上の式を以下のように分解する。

そして、３番目の式を上の２つの式に代入してを消去してやると、

これを、以下のように行列の形に変形してやる。

思い出して欲しい。は、パッド上の点だったよね。１つの点につき、上のような式が１つできる。パッド上の個の点について、上のような式を積み重ねてやると、以下のようになる。

これは、の線形方程式だ。このような形を**Homogeneous Linear System**と呼ぶ。Overdeterminedの場合は、左の行列に対して、の最小固有値に対応する固有ベクトルが、右のの解となる。実際には、**SVD（特異値分解）**を行った時に、行列の最後の列が解となる。（assignmentでは、これを実装するとextra creditだと言っている。）

※ちなみに、このアルゴリズムを**DLT(Direct Linear Transform)**とも呼ぶらしい。

**ステップ２（行列を解く）**

まず、式(2)で、、は互いに直交し、ノルム（長さ）が同じだということから、制約条件を生成したい。式(2)で、とおいたよね。なので、

とおくと、

よって、、は互いに直交し、ノルム（長さ）が同じだということから、

式を変形して、

(3)

(4)

ここで、とおくと、行列は式(1)で定義しているので、実際に力ずくで計算できて、以下のように求まる。

(5)

これを見ると分かるとおり、行列は対称行列だ。なので、を以下のようにおく。

また、式(3)、(4)にを代入すれば、

ここで、一般に、について考えて見る。

これは、さらに以下のように変形できる。

(6)

左側のベクトルを、右側をとおくと、

と表される。これを、式(3), (4)に適用すれば、

行列の形にしてやると、

(8)

式(6)を思い出して欲しい。ベクトルは、homography行列の要素で構築されている。そして、それはステップ１で計算済みだ。つまり、上の式で、ベクトルだけが未知なのだ。

一応、行列の形にしておくと、

各画像について１つのhomography行列が求まる。つまり、上の式が１つできる。個の画像があれば、上の式(8)を個積み上げれば良い。そして、SVDを使ってベクトルを解く。ただし、画像が１個しかない場合は、式が２つということ。なので、２つの未知数しか求めることができない。一般的に、、とは画像の中心とみなし、残りのパラメータとを求めることになる。一方、画像が２個しかない場合は、式が４つということなので、４つの未知数しか求めることができない。一般的に、とみなし、残りの４つのパラメータを求めることになる。（assignmentでは画像を最低２個と言っているので、と見なして良いということだ）。ということは、ということ。さらに、、を画像の中心とみなし、以下の式を追加してあげることで、式が７個となり、行列を求めることが出来る。

この式(8)も**Homogeneous Linear System**なので、**SVD**を使って解くことが出来る。行列が求まれば、式(5)を使って、、、、、を計算すれば良い。つまり、

こうして、カメラ内部パラメータ行列が求まった。

次に、外部パラメータ行列を求めよう。思い出してくれ。Homography行列は、

よって、

つまり、

ただし、

もも既に求まっているので、、、は簡単に求まるよね。また、回転行列の３つ目の列については、

で、普通に解ける。

**ステップ３（推定値の改善）**

ここまでで、基本的には内部パラメータ行列、外部パラメータ行列は得られたが、iterativeな最適化により、更に推定値を改善できる。

個の画像を使用し、各画像内に個の点がある時、最適化に使用する目的関数は、

(10)

ただし、は番目の画像内での番目の点の座標、は3D座標から写像により得られた二次元座標だ。これを、**Minpack**の**lmdif関数**で解けば良い。

**ステップ４（レンズ歪みの推定）**

まず、レンズ歪みのモデルは、

ただし、は、歪みなしの二次元座標、は、歪みより実際に観測された二次元座標。また、

ただし、は正規化された座標で、以下で求められる。

後は、**Minpack**を使って、以下の目的関数を最小化すれば良い。

(14) (10)

## Bundle Adjustment

Wikipediaによると、Camera calibrationやら3D Reconstructionやらにおいて、必ず最後のステップとして実施されるプロセスとのこと。コンピュータビジョンの世界ではめちゃくちゃ有名らしい。

基本的には、3D座標を射影した2D座標と、実際に観測された2D座標との差を最小化する処理だ。Levenberg-Marquardtアルゴリズムが、この処理によく使用される。簡単に説明すると、現在の値の近傍について、目的関数がほぼ線形であるという前提に基づき、勾配を使って極小値へ向かう。**（Gradient Descentと何が違うの？）**

## Algebraic Error vs Geometric Error

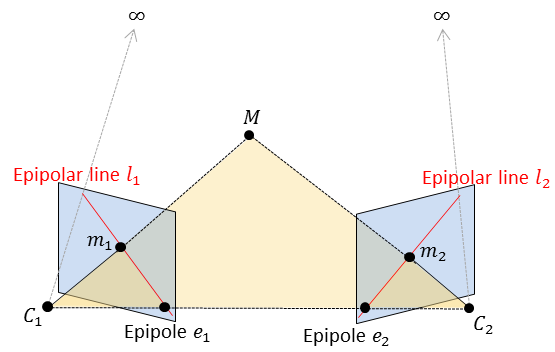
例えばZhangのアルゴリズムで、Homogeneous Linear Systemを解くところはAlgebraicアプローチと呼ばれる。なぜなら、というHomogeneous Linear Systemを解く際に、は最小化したい何らかの距離ではない。つまり、特に意味が無く、ただ制約条件に基づいて機械的に代数を使って解いているだけなのだ。

一方、最後のnon-linear optimizationのステップは、意味のある何らかの距離を最小化するよう最適化している。

一般的に、**「まずは線型方程式を使ってalgebraic errorを最小化し、得られたパラメータを初期値として、次にnon-linear optimizationを使って目的の距離を最小化する」**というのがよく使われる流れだ。

# Stereo Reconstruction

## Epipolar Geometry



左のカメラにとって、点はへ写像される。一方、右のカメラにとっては、点はへ写像される。この時、左のカメラについて、

※とは、Calibrationの時のそれとは違うような気がする。。。

この式を変形して、

(2.1)

この式が意味しているのは、この式を満たす点は、全てに写像されるということだ。例えば、上図の線上の点は、全てこの式を満たす。同様に、右のカメラについても、

(2.2)

(2.1)、(2.2)からを消去すると、

ここで、右辺の第二項は、カメラ１の、カメラ２への写像を表す。つまり、Epipole だ。スケールパラメータをとすれば、と表せる。また、とおくと、

となり、、、の線形結合となった。つまり、線形依存しているということだ。なので、これらを列とする行列の行列式は０となるはずだ。つまり、

よって、

ここで、

とおくと、

*、*は観測データなので、これを使えばを求めることが出来る。さらに、が求まれば、よりが求まり、さらに、より、も求まる。

後は、、を計算し、が求まる。

ポイントは、カメラの内部パラメータ行列、外部パラメータ行列が事前に分からなくても良いということだ！

は**Fundamental Matrix**と呼ばれる。

## 以下の資料のまとめ

<http://airglow.csl.illinois.edu/~jessie/CV_project_FinalVersion.pdf>