

## 回転運動するカメラにより撮影された動画象からの

## カメラ回転角の導出の検討

## Study of Estimating the Rotational Parameters of the Rotatable Camera from the Video Sequence

柄折 泰史†  
Yasufumi Tochiori

江畑 勝‡  
Masaru Ebata

大谷 淳†  
Jun Ohya

## 1. はじめに

映画やドラマなどの映像作品における実写とCGとの合成において、カメラの撮影方向・動きの変化を正確に推定することが必要とされている。本研究ではそのような推定を動画情報から行う手法について、パン・ティルトというカメラが回転運動を行うときに限定して考察・提案を行っている。

動画から撮影時のカメラの回転角を導出する手法として、(1)ロドリゲスの公式(Rodrigues' Rotation Formula)を用いる手法、(2)モーメント行列から回転行列を求める手法、(3)相関行列を極分解、または特異値分解することによって求める手法などがある[1]。(1)は幾何学的、(2)、(3)は代数学的な解法であり、我々の提案する手法は幾何学的な手法である。従来の手法では回転計算には関与しない移動ベクトルが含まれてしまう可能性があり、エラー特徴点があった場合は取り除かなければならなかった。また手法(2),(3)では誤差にはロバストであるが、エラーは別の手段で除くことが必要である。提案手法では初期段階で二つの特徴点間の長さがフレーム間で一致するという条件として計算を行うことで、誤差のみならずエラーにも強くすることを目指す。本論文ではこれを実現する新たな手法についてその原理と共に示す。そしてその手法で回転角を導出できるかについてシミュレーションを行う。

## 2. 提案手法の原理

この章ではカメラがパン・ティルトを行いつつ実世界を連続したフレームに撮像するまでの過程を数学的に捉え、提案手法を示す。

まず、カメラパラメータは事前にわかっているものとして、カメラの投影のモデルを透視投影とする。このとき、座標軸はカメラの光学中心  $O$  を原点とし、光軸方向を  $z$  軸の正の方向、フレームの水平方向を  $x$  軸、垂直方向を  $y$  軸とする左手座標系を考える。すると実世界中のある点  $P$  の座標  $(X, Y, Z)$  は光学中心  $O$  から  $d$  の距離にあるフレーム上の点に式(1)の通りに投影される。

$$x = d \frac{X}{Z} \quad y = d \frac{Y}{Z} \quad (1)$$

このとき、フレーム上での座標は画像中心  $S$  を原点とし、画像平面  $\Pi$  の横幅を  $w$ 、縦幅を  $h$  とする。カメラの画角を  $\alpha$  としたとき、 $\alpha$  は光学中心  $O$  からフレームの対角線へ向けた角度なので光学中心  $O$  から画像中心  $S$  までの

距離  $d$  は式(2)で表わされる。

$$d = \frac{\sqrt{w^2 + h^2} / 2}{\tan(\alpha/2)} \quad (2)$$

本研究ではパンは光学中心  $O$  を中心としてカメラを垂直方向に回転運動 ( $x$  軸を回転の軸として右回りに回転する運動) するときの撮影とし、ティルトは光学中心  $O$  を中心としてカメラを水平方向に回転運動 ( $z$  軸を回転の軸として右回りに回転する運動) するときの撮影と定義し、それぞれの回転角度を  $\theta$ 、 $\phi$  とおく。またパン・ティルト前のカメラの撮影方向は  $z$  軸の正の方向にあるとする。パン・ティルト時の撮影モデルは図1のように表され、パン・ティルト前のフレームを基準フレーム  $f_1$ 、パン・ティルト後のフレームを比較フレーム  $f_2$  とする。

$f_1$ 、 $f_2$  からパン・ティルトの回転角度を求めるのにあたって、まず  $f_1$ 、 $f_2$  に共通して映る点  $P$  について考える。ここで、 $f_1$ 、 $f_2$  での座標系の情報を扱いやすくするためにカメラが固定され、世界座標が光学中心  $O$  を中心に回転するローカル座標について相対的に考える (図2)。このカメラのローカル座標を考えることで、カメラの光軸、フレームは固定され、カメラが固定される代わりに点  $P$  の座標は光学中心  $O$  を中心に実際のパン・ティルトとは反対方向に回転される。この座標系の元、パン・ティルト後に点  $P$  が移動した点を  $P'$  とすると、 $P'$  の座標  $(x', y', z')$  は次の式(3)で表わされる。

$$\begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} = \begin{pmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (3)$$

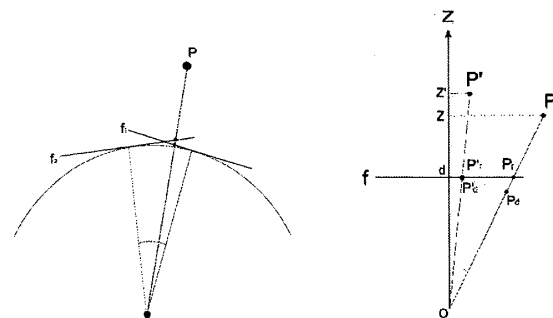


図1: フレームの推移 図2: カメラの相対的關係

図2のように点  $P$  は点  $P'$  に移動することでフレーム上にはそれぞれ点  $P_f$ 、点  $P'_f$  に投影される。ここで  $d'$  を  $OP$  あるいは  $OP'$  の長さとするればパン・ティルト時は光学中心  $O$  と点  $P$ ・点  $P'$  の距離は変わらないので  $d'$  は

† 早稲田大学理工学術院国際情報通信研究科 Graduate School of Global Information and Telecommunication Studies, WASEDA University

‡ (株) エム・ソフト M.SOFT CO.,LTD.

一定である。これを利用して Z とフレーム上の点の座標の関係は三角形の相似を用いて次の式(4)で表される。

$$Z = \frac{dd'}{\sqrt{x^2 + y^2 + d^2}}, d' = \sqrt{X^2 + Y^2 + Z^2} \quad (4)$$

パン・ティルトによって動くとき、Z の変化に伴って、式(4)に従って点  $P_f(x, y)$  は推移していくことがわかる。ここで、透視投影の式(1)の X, Y をフレームに投影する代わりに点 P の位置に関わらず常に一定の割合で投影するような点を式(5)によって考える。

$$\begin{pmatrix} x_d \\ y_d \\ -d \end{pmatrix} = \frac{d}{\sqrt{x^2 + y^2 + d^2}} \begin{pmatrix} x \\ y \\ d \end{pmatrix} = \frac{d}{d'} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (5)$$

点  $P_d(x_d, y_d, z_d)$  は式(1),(4),(5)から示されるようにフレーム上の点  $P_f(x, y)$  から求めることができ、3次元空間中にある点 P の座標を  $d/d'$  倍に投影した点である。図2に示すように点 P の位置に関わらず光学中心 O から d の距離にある。これは実世界の点 P を画像平面  $\Pi$  で投影するのではなく、光学中心 O を中心とし、半径を d とした仮想球体の曲面上に投影することに相当する。

ここで基準フレーム  $f_1$  と比較フレーム  $f_2$  の間で共通に映る3点を考え、それぞれ  $f_1$  での3点を S, A, B,  $f_2$  でのそれぞれの対応する点を  $S'$ ,  $A'$ ,  $B'$  とする。このとき  $\triangle SAB$  と  $\triangle S'A'B'$  は合同ではない。その理由はそれぞれの点はフレーム上に投影されるときに式(4)に従ってそれぞれの点の位置がずれるからである。それに対して、S, A, B,  $S'$ ,  $A'$ ,  $B'$  を光学中心 O から d の距離にある仮想球体上に投影し、それぞれを s, a, b,  $s'$ ,  $a'$ ,  $b'$  とすると  $\triangle sab$  と  $\triangle s'a'b'$  は合同となる(図3)。それは各点が式(5)のように元の3次元物理世界の点の座標を  $d/d'$  という一定の値で投影しているからである。

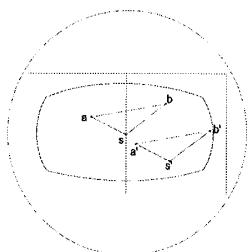


図3:仮想球体上に投影された三角形

カメラのパン・ティルトの回転角を求めるためにカメラの撮影方向を示す点の座標の推移を考える。つまり、基準フレーム  $f_1$  の画像中心(0, 0, d)を s とし、パン・ティルト後に仮想球体上で s が移動した点の座標  $s'$  ( $x_s, y_s, z_s$ ) とすれば点  $s'$  の座標を求めればよい。点  $s'$  の座標は他の s, a, b,  $a'$ ,  $b'$  の座標、 $\triangle sab$  と  $\triangle s'a'b'$  が合同であること、 $s'$  は仮想球体上にあることから求められる。求められた点  $s'$  の極座標を考えることで回転角  $\theta$ ・ $\phi$  は次の式(6)で表される。

$$\begin{aligned} \theta &= \arcsin(y_s / d) \quad (6) \\ \phi &= \arctan(-x_s / z_s) \end{aligned}$$

ただし、カメラを固定した相対的な座標を考えているので実際のカメラのパン・ティルトの角度は  $\theta$  と  $\phi$  の正負を逆にしたものである。以上の提案手法の流れをまとめると図4の通りとなる。

### 3. シミュレーション

本章では2章で示した提案手法についてのシミュレーションを行った。

まず、フレームのサイズに関して横幅  $w=1920$ , 縦幅  $h=1080$ , 画角を  $\alpha=96^\circ$  とした。d は式(2)より  $d=991.75$  と求まった。基準フレーム  $f_1$  と比較フレーム  $f_2$  に共通して映る3次元空間中の2点  $P_A, P_B$  の座標を  $P_A(-200, -100, 1800)$ ,  $P_B(1200, 400, 2000)$  とした。そしてパン・ティルトの角度を  $\theta=20^\circ$ ,  $\phi=40^\circ$  として回転移動を  $P_A, P_B$  に行った。するとその座標は  $(912.05, -709.61, 1398.08)$ ,  $(2215.24, -308.16, 773.15)$ , となった。

これらの座標に対して、まずフレームに投影した点 A, B,  $A'$ ,  $B'$  について求めた。これらは式(1)から導出され、それぞれ  $A(-110.19, -55.10)$ ,  $B(595.05, 198.35)$ ,  $A'(646.98, -503.37)$ ,  $B'(2841.59, -395.29)$  となった。シミュレーションではこれらフレームに映る点の元となる実世界の点の座標もあらかじめ知られているが、実画像で行う場合はこの A, B,  $A'$ ,  $B'$  の座標の情報のみから計算を始める。

次にフレーム上の点 A, B,  $A'$ ,  $B'$  に対して光学中心 O を中心にした半径 d 上の仮想球体へと投影する。このとき、A, B,  $A'$ ,  $B'$  の4点はフレーム上に投影された点であるので、z 座標は d である。式(5)を用いて投影した点 a, b,  $a'$ ,  $b'$  の座標を算出するとそれぞれ  $a(-109.35, -54.68, 984.19)$ ,  $b(502.91, 167.64, 838.18)$ ,  $a'(498.68, -387.99, 764.43)$ ,  $b'(928.39, -129.15, 324.02)$  となった。

これら4点 a, b,  $a'$ ,  $b'$ 、さらに  $f_1$  の画像中心  $s(0, 0, d)$ 、 $\triangle sab$  と  $\triangle s'a'b'$  が合同であること、 $s'$  は仮想球体上にあることを考えると、その座標は  $s'(599.04, -339.20, 713.91)$  となり、 $s'$  の座標は式(6)より  $\theta=20.00$ ,  $\phi=40.00$  となった。

### 4. まとめ

本研究ではカメラのパン・ティルトにより撮影された動画象から、その撮影時のパン・ティルトの角度を導出する方法について提案した。この手法ではフレームに映る3点を光学中心 O から等距離 d にある仮想球体上に投影することを考える。仮想球体上ではパン・ティルト後も3点の作る三角形の合同は保たれ、その性質を利用することでカメラの撮影方向を示す点の移動後の座標を求めることができ、回転角  $\theta$ ,  $\phi$  を求めることができる。本論文ではシミュレーションを行うことでその手法を理論的に妥当であることも示した。

今後の課題としては実写画像に対して提案手法を適用すること、2つのフレーム間に共通に映る点を抽出するのにどのような手法が最適であるかを検討すること、複数の点から得た結果をどう組み合わせていくのが最適であるかを検討することが考えられる。

#### 参考文献

- [1] 金谷健一, “画像理解-三次元認識の数理-”, 森北出版, (1990).
- [2] 佐藤淳, “コンピュータビジョン-視覚の幾何学-”, コロナ社, (1999).

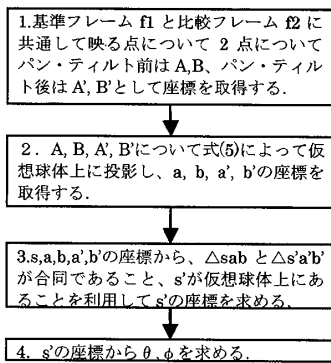


図4:フローチャート