

## 3D Motion Reconstruction for Real-World Camera Motion

西田研究室 17344219 二宮 悠二

本稿では論文「3D Motion Reconstruction for Real-World Camera Motion」の第 3 節「Adding 3D Body Keyframes」の後半部分についての説明を行う。内容としては、キーフレームの抽出箇所と再構成性、画像点からのカメラモーションの計算方法についてである。本稿では特に、論文内で参考にされた二つの手法について述べる。

## 3.1 Reconstructibility &amp; 3D Keyframes

前章までの発表にて、画像点から復元する方法と軌道から復元する方法の二つが挙げられた。その中で次の式

$$\arg \min_{\beta} \| \dot{\mathbf{X}} - \dot{\mathbf{\Theta}}\beta \|^2 + \lambda \| \beta \|^p \quad (3.1)$$

で与えられた目的関数の欠点は、コンピュータビジョン（例えば画像点の追跡）における他ツールから提供される追加情報を考慮しないという点である。逆に軌道生成に基づく NRSFM 法の欠点は、第 2 節で述べたように画像点の投影による再構成性が低いことである。具体的には、カメラの動きが点の動きと相関している、すなわち、カメラと点が同様な滑らかで緩やかな軌道を有する場合、再構成性は低下する。このときの目的関数は次式で与えられる。

$$\arg \min_{\beta} \| \mathbf{x} - \mathbf{P}\dot{\mathbf{\Theta}}\beta \|^2 + \lambda \| \beta \|^p \quad (3.2)$$

こうした欠点をカバーするために、新しく目的関数を次のように定義する。

$$\arg \min_{\beta} (1 - \gamma) \| \mathbf{x} - \mathbf{P}\dot{\mathbf{\Theta}}\beta \|^2 + \gamma \| \dot{\mathbf{X}} - \dot{\mathbf{\Theta}}\beta \|^2 + \lambda \| \beta \|^p \quad (3.3)$$

ここで、 $\gamma$  は二次元画像投影に関連した 3D キーフレームを制御するための重みで、 $\gamma = 1$  のとき、(3.1) 式で示した圧縮センシング問題に、 $\gamma = 0$  のとき、(3.2) 式で示した軌道基底問題が生じる。

## 3.1.1 Akhter の手法

この手法の中で 3.3 式における  $\mathbf{P}$  と  $\beta$  を同時に解く方法が提案された。本稿では、 $\mathbf{P}$  が体の剛体的な部分、すなわち運動中に不動であると考えられる胴体の部分から計算することができれば、より良い結果が得られることが分かったのだが、比較対象としてこの手法について記すことにする。

計測された 2 次元軌道は  $2F \times P$  の観測行列  $\mathbf{W}$  に含まれており、その中に  $F$  フレーム、 $P$  個の画像点を含んでいる。

この観測行列は  $\mathbf{W} = \mathbf{R}\mathbf{S}$  として正射影行列  $\mathbf{R}$  と構造行列  $\mathbf{S}$  に分解できる。ただし、 $\mathbf{R}$  は  $2F \times 3F$  の行列である。

さらに、 $\mathbf{R}_t$  は  $2 \times 3$  の正射影行列である。第 2 節にて、本稿では  $\mathbf{S} = \mathbf{\Theta}\mathbf{A}$  とおくことで次のように分解することができる。

$$\mathbf{W} = \mathbf{R}\mathbf{S} = \mathbf{R}\mathbf{\Theta}\mathbf{A} = \mathbf{\Lambda}\mathbf{A} \quad (3.4)$$

ここで、 $\mathbf{\Lambda} = \mathbf{R}\mathbf{\Theta}$  である。 $\mathbf{\Lambda}$  は  $3F \times 3k$  の行列であるので、 $\mathbf{W}$  のランクは  $3k$  である。

SVD ( Singular Value Decomposition ) を用いて特異値分解を行うと、 $\mathbf{W}$  は次のように表せる。

$$\mathbf{W} = \hat{\mathbf{\Lambda}}\hat{\mathbf{A}} \quad (3.5)$$

一般に、行列  $\hat{\mathbf{\Lambda}}$  と  $\hat{\mathbf{A}}$  はそれぞれ  $\mathbf{\Lambda}$  と  $\mathbf{A}$  には等しくならない。これは、因数分解の結果が一意的に決まらないためである。したがって、メートル構造体の復元のために必要であるのは次式で与えられるような可逆行列を算出することである。

$$\mathbf{\Lambda} = \hat{\mathbf{\Lambda}}\mathbf{Q} \quad , \quad \mathbf{A} = \mathbf{Q}^{-1}\hat{\mathbf{A}} \quad (3.6)$$

### 3.1.2 Wei & Chai の手法

また、本稿の考え方のベースとなっている Wei と Chai の手法について説明する。先行研究の論文では回転の制約だけでは曖昧さや無効な解が生じてしまうということが証明されている。ここでは形状のベースとなる部分を独自に決定し、それを基底制約として前述した曖昧さを取り除くことを目的としている。弱透視射影法において基底と回転の制約の両方を課することが非剛体形状と軌道復元の問題に解法をもたらすことにつながる。本稿での解法の正確性とロバスト性は合成データにおいて定量的に、現実のビデオ映像において定性的に評価される。これについてまとめる。

まず、定理 1 として次が挙げられる。

#### 定理 1

回転制約の一般的な解は  $\mathbf{G}\mathbf{H}\mathbf{G}^T$  で表される。ただし  $\mathbf{G}$  は訂正変換行列であり、 $\mathbf{H}$  は任意のブロック歪対称行列と任意のブロックスケール恒等行列の和である。

ここで、 $\mathbf{G}$  は  $K$  の三重列であり、それを  $g_k$  ( $k = 1, \dots, K$ ) とする。したがって  $g_k$  はそれぞれが  $3K \times 3$  の行列である。さらに  $g_k g_k^T$  を  $\mathbf{Q}_k$  とおく。 $\mathbf{Q}_k$  は  $3K \times 3K$  の対象行列である。 $g_k g_k^T$  は SVD 使うことで計算されるので、この  $\mathbf{Q}_k$  を計算することにより二つの制約、回転制約と基底制約を利用でき、制約として課することができるようになるのである。

さて、この定理 1 のために、 $\mathbf{Q}_k$  に回転制約を課して  $\mathbf{GHG}^T$  を曖昧な解にする。 $\mathbf{H}$  は  $K^2$  の  $3 \times 3$  のブロックを含み、 $\mathbf{H}_{mn}$  とおく ( $m, n = 1, \dots, K$ )。  $\mathbf{H}_{mn}$  は次のように 4 つの独立な項を含む。

$$\mathbf{H}_{mn} = \begin{bmatrix} h_1 & h_2 & h_3 \\ -h_2 & h_1 & h_4 \\ -h_3 & -h_4 & h_1 \end{bmatrix} \quad (3.7)$$

ここで補助的に、退縮していない状況では  $\mathbf{H}_{mn}$  はゼロ行列で

$$\mathbf{R}_i \mathbf{H}_{mn} \mathbf{R}_j^T = \mathbf{0}_{2 \times 2} \quad (3.8)$$

であるとする。ただし、 $\mathbf{R}_i$  と  $\mathbf{R}_j$  は  $2 \times 3$  の回転行列である。これを派生させることにより、次の定理 2 を得る。

## 定理 2

基底制約と回転制約の両方を課すと、 $\mathbf{Q}_k$  の適正な解が得られる。

この定理により、最小二乗法を用いて線形方程式を解くことで  $\mathbf{Q}_k = g_k g_k^T$  ( $k = 1, \dots, K$ ) を計算する。その後、SVD により  $\mathbf{Q}_k$  を分解して  $g_k$  を復元する。 $\mathbf{Q}_k$  の分解は任意の  $3 \times 3$  の正規直交変換  $\phi$  であり、 $(g_k \phi)(g_k \phi)^T$  もまた  $\mathbf{Q}_k$  に等しい。

この曖昧さは、 $g_1, \dots, g_K$  が異なる座標系の下で独立して推定されるという事実から生じている。これを解消するために、単一の基準座標系の下に変換を行う。すると、それぞれの回転の組が異なる  $3 \times 3$  までの正規直交変換となる。基準座標系を指定するために、一組の回転を選び、他の回転の組の符号をこれらの回転が対応する基準のものと一致するように選ぶ。最後に、回転の組から基準への正規直交変換を計算することにより、変換された  $g_1, \dots, g_K$  は望みの訂正行列  $\mathbf{G}$  を作り、正しい係数の計算および形状基底の復元がなされる。それらを組み合わせることで非剛性三次元形状が構成される。

## 参考文献

- [1] I.Akhter, Y.Sheikh, S.Khan, and T.Kanade, "Nonrigid structure from motion trajectory space", In *NIPS*, 2008.
- [2] J.Xiao, J.Chai, and T.Kanade, "A closed form solution to non-rigid shape and motion recovery", *International Journal of Computer Vision*.