Robust Trajectory Clustering for Motion Segmentation
(ICCV 2013 paper supplementary documentation)

Feng Shi[1], Zhong Zhou[1], Jiangjian Xiao[2], Wei Wu[1]

[1]State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, China

{shifeng,zz,wuwei}@vrlab.buaa.edu.cn

[2]Ningbo Industrial Technology Research Institute, CAS, China

xiaojj@nimte.ac.cn

1. Introduction

This documentation provides the mathematical derivation of Eq. (4) and Eq. (5) in Algorithm 2 of the main manuscript.

2. Derivation of Eq. (4) in Algorithm 2

Form $vec(\tilde{W}) = ((\tilde{w}^1)^T, ..., (\tilde{w}^{2P})^T)^T$, $vec(C) = ((c^1)^T, ..., (c^{2P})^T)^T$, and

$$\Psi = \begin{pmatrix} \Pi^1 \Omega_d X & & \\ & \ddots & \\ & & \Pi^{2P} \Omega_d X \end{pmatrix}$$

the error function in Eq. (3) can then be rewritten as:

$$f(X, C) = \frac{1}{2} \|\Psi vec(C) - vec(\tilde{W})\|_2^2 \qquad (s.1)$$

With the known X, computing C that minimizes Eq. $(s.1)$ is equivalent to computing the optimal solution of a linear programming problem in the least squares sense. To solve it, we calculate the first-order derivative of the error function in Eq. $(s.1)$ as:

$$\frac{df(vec(C))}{dvec(C)} = \Psi^T \Psi vec(C) - \Psi^T vec(\tilde{W}) \qquad (s.2)$$

Let $df(vec(C))/dvec(C) = 0$, the global minimum of Eq. $(s.2)$, $i.e.$ the least squares solution of C, is obtained as:

$$vec(C) = (\Psi^T \Psi)^{-1} \Psi^T vec(\tilde{W})$$

3. Derivation of Eq. (5) in Algorithm 2

Denote $r^p = \tilde{w}^p - \Pi^p \Omega_d X c^p$ as the residual between the measured and estimated trajectory of point p, the error function in Eq. (3) can be rewritten as:

$$f(X, C) = \frac{1}{2} \sum_{p=1}^{2P} (r^p)^T r^p \qquad (s.3)$$

Given the known trajectory coefficient matrix C, the first-order and second-order differentials of the error function in Eq. $(s.3)$ are:

$$\begin{aligned} df &= \frac{1}{2} \sum_{p=1}^{2P} ((dr^p)^T r^p + (r^p)^T dr^p) = \sum_{p=1}^{2P} (dr^p)^T r^p \\ d^2 f &\approx \sum_{p=1}^{2P} (dr^p)^T dr^p \end{aligned} \qquad (s.4)$$

where the second-order term $d^2 r^p$ is neglected in $d^2 f$.

Subsequently, the first-order differential of the residual r^p is derived as:

$$dr^p = -\Pi^p \Omega_d dX c^p \qquad (s.5)$$

By using the equality $vec(BXA) = (A^T \otimes B) vec(X)$ [1], we then vectorize both sides of Eq. $(s.5)$ into:

$$dr^p = -((c^p)^T \otimes (\Pi^p \Omega_d)) vec(dX) = -J^p vec(dX) \qquad (s.6)$$

where $J^p = (c^p)^T \otimes (\Pi^p \Omega_d)$.

According to the chain rule of differential calculus [1], we substitute Eq. $(s.6)$ to Eq. $(s.4)$, and obtain:

$$\begin{aligned} df &= vec(dX)^T (-\sum_{p=1}^{2P} (J^p)^T r^p) \\ d^2 f &\approx vec(dX)^T (\sum_{p=1}^{2P} (J^p)^T r^p) vec(dX) \end{aligned} \qquad (s.7)$$

So, Eq. (5) is derived:

$$g = -\sum_{p=1}^{2P} (J^p)^T r^p, H = \sum_{p=1}^{2P} (J^p)^T J^p$$

References

[1] J. R. Magnus and H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics, 2nd edition*. Wiley, 1999.