On 3-D Motion Estimation from Feature Tracks in 2-D FS Sonar Video

Sun Qinxuan

January 8, 2019

Title: On 3-D Motion Estimation from Feature Tracks in 2-D FS Sonar Video

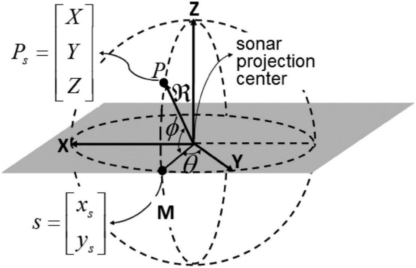
Published in: IEEE Transactions on Robotics (Volume: 29, Issue: 4, Aug. 2013)

Author: Shahriar Negahdaripour, Fellow, IEEE

The author is with the Department of Electrical and Computer Engineering,
University of Miami, Coral Gables, FL 33146 USA (e-mail: shahriar@miami.edu).

OUTLINE

NOTATION AND PRELIMINARIES

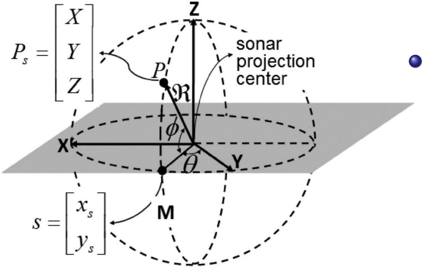


- 场景中的点P的直角坐标$\boldsymbol{P}_s = [X,Y,Z]^T$和极坐标$[\mathfrak{R},\theta,\varphi]$.

$$\boldsymbol{P}_s = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathfrak{R} \begin{bmatrix} \sin\theta\cos\varphi \\ \cos\theta\cos\varphi \\ \sin\varphi \end{bmatrix} \tag{1}$$

$$\begin{aligned} \mathfrak{R} &= \sqrt{X^2+Y^2+Z^2} \\ \theta &= \tan^{-1}(X/Y) \\ \varphi &= \sin^{-1}(Z/\mathfrak{R}) \end{aligned} \tag{2}$$

NOTATION AND PRELIMINARIES

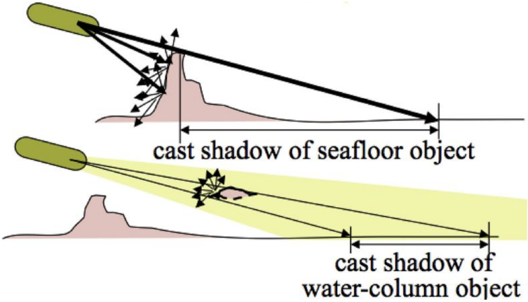


- 声纳图像(beam-bin image)$I(x_s, y_s)$.

$$\boldsymbol{s} = \begin{bmatrix} x_s \\ y_s \end{bmatrix} = \Re \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix} \tag{3}$$

NOTATION AND PRELIMINARIES

物体在海床上的投影



NOTATION AND PRELIMINARIES

令(形成阴影的)物体边缘一点$\boldsymbol{P}_s$对应的声纳图像上的坐标为$\boldsymbol{s}$.
$\boldsymbol{P}_s$在海床上的投影点为$\widehat{\boldsymbol{P}}_s$，$\widehat{\boldsymbol{P}}_s$对应声纳图像上点的坐标为$\widehat{\boldsymbol{s}}$，海床平面的法向量$\boldsymbol{n} = [n_x, n_y, n_z]^T$，则有

$$\boldsymbol{n} \cdot \widehat{\boldsymbol{P}}_s = -1 \tag{4}$$

$$\widehat{\boldsymbol{s}} = -(\boldsymbol{P}_s \cdot \boldsymbol{n})^{-1} \boldsymbol{s} \tag{5}$$

分析物体上的点及其阴影点在声纳图像上坐标随声纳运动(角速度$\boldsymbol{\omega} = [\omega_x, \omega_y, \omega_z]^T$线速度$\boldsymbol{t} = [t_x, t_y, t_z]^T$)的变化规律，进而通过跟踪物体上的点及其阴影点来恢复声纳的运动.

Image Motion Model of Stationary 3-D Object Point

场景中一个静止的点$\boldsymbol{P}_s$相对于声纳的移动速度为

$$\frac{d\boldsymbol{P}_s}{dt} = -\boldsymbol{\omega} \times \boldsymbol{P}_s - \boldsymbol{t} \tag{6}$$

$\boldsymbol{P}_s$对应声纳图像上的点$\boldsymbol{s}$的移动速度为

$$\frac{d\boldsymbol{s}}{dt} = \frac{1}{\cos\varphi}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}\frac{d\boldsymbol{P}_s}{dt} + \tan\varphi\,\boldsymbol{s}\frac{d\varphi}{dt} \tag{7}$$

化简得

$$\frac{d\boldsymbol{s}}{dt} = (\boldsymbol{u}\cdot\boldsymbol{\omega})\boldsymbol{s}_n + \frac{\sin\varphi}{\mathfrak{R}}(\boldsymbol{u}\cdot\boldsymbol{t})\boldsymbol{s} - \frac{1}{\cos\varphi}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}\boldsymbol{t} \tag{8}$$

其中$\boldsymbol{u} = [\sin\theta\tan\varphi, \cos\theta\tan\varphi, -1]^T$， $\boldsymbol{s}_n = [-y_x, x_s]^T$.
在声纳图像中仰角φ是不可测的，需要对其进行估计。

Image Motion Model of Stationary 3-D Object Point

把式(8)展开成声纳图像坐标的多项式形式

$$\begin{bmatrix} \frac{dx_s}{dt} \\ \frac{dy_s}{dt} \end{bmatrix} = \begin{bmatrix} -\frac{t_x}{\cos\varphi} - \left(\frac{t_z \sin\varphi}{\Re}\right) x_s + \omega_z y_s + \left(\frac{\sin\varphi \tan\varphi\, t_x}{\Re^2}\right) x_s^2 \\ + \left(\frac{\sin\varphi \tan\varphi\, t_y}{\Re^2} - \frac{\tan\varphi\, \omega_x}{\Re}\right) x_s y_s - \left(\frac{\tan\varphi\, \omega_y}{\Re}\right) y_s^2 \\ -\frac{t_y}{\cos\varphi} - \omega_z x_s - \left(\frac{t_z \sin\varphi}{\Re}\right) y_s + \left(\frac{\sin\varphi \tan\varphi\, t_y}{\Re^2}\right) y_s^2 \\ + \left(\frac{\sin\varphi \tan\varphi\, t_x}{\Re^2} + \frac{\tan\varphi\, \omega_y}{\Re}\right) x_s y_s + \left(\frac{\tan\varphi\, \omega_x}{\Re}\right) x_s^2 \end{bmatrix}$$

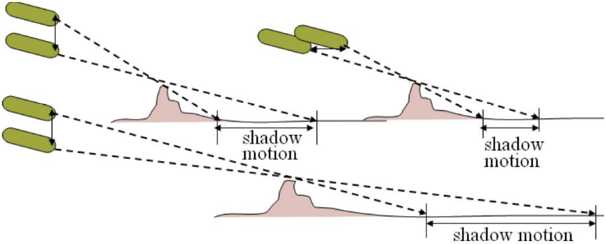
(忽略二阶项)其一阶声纳图像运动模型为

$$\begin{bmatrix} \frac{dx_x}{dt} \\ \frac{dy_x}{dt} \end{bmatrix} \approx \begin{bmatrix} -t_x - \left(\frac{t_z \sin\varphi}{\Re}\right) x_s + \omega_z y_s \\ -t_y - \omega_z x_s - \left(\frac{t_z \sin\varphi}{\Re}\right) y_s \end{bmatrix} \qquad (9)$$

Shadow Motion

阴影点$\widehat{\boldsymbol{P}}_s$对应声纳图像点$\widehat{\boldsymbol{s}}$的运动

$$\frac{d\widehat{\boldsymbol{s}}}{dt} = \frac{\widehat{\mathfrak{R}}}{\mathfrak{R}}\left(\frac{d\boldsymbol{s}}{dt} - \left(\frac{\widehat{\mathfrak{R}}-\mathfrak{R}}{\mathfrak{R}}\right)(\boldsymbol{n}\cdot\boldsymbol{t})\boldsymbol{s}\right) \tag{10}$$



Pure Rotation

当声纳只进行旋转运动时

$$\left(\frac{d\boldsymbol{s}}{dt}\right)_r = (\boldsymbol{u}\cdot\boldsymbol{\omega})\boldsymbol{s}_n \tag{11}$$

一个点只提供一个独立的约束，有四个未知量$(\boldsymbol{\omega},\varphi)$.
N个点提供N个约束，有$N+3$个未知量.
无法直接恢复运动和结构.
若可以估计出φ，则$N\geq 3$个点便可以恢复旋转运动.

Pure Translation

$$\left(\frac{d\boldsymbol{s}}{dt}\right)_t = \frac{\sin\varphi}{\mathfrak{R}}(\boldsymbol{u}\cdot\boldsymbol{t})\boldsymbol{s} - \frac{1}{\cos\varphi}\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}\boldsymbol{t} \tag{12}$$

已知$\cos\varphi \approx 1$，则

$$\left(\frac{d\boldsymbol{s}}{dt}\right)_t = \frac{\sin\varphi}{\mathfrak{R}}(\boldsymbol{u}\cdot\boldsymbol{t})\boldsymbol{s} - \begin{bmatrix}t_x\\ t_y\end{bmatrix} \tag{13}$$

如果φ已知，则(13)给出两个关于$\boldsymbol{t}$的线性方程.
如果无法得到φ的估计，可通过下式消去φ

$$\boldsymbol{s}_n\cdot\left(\frac{d\boldsymbol{s}}{dt}\right)_t + \boldsymbol{s}_n\cdot\begin{bmatrix}t_x\\ t_y\end{bmatrix} = 0 \tag{14}$$

$$\left[y_s, -x_s, y_d\left(\frac{dx_s}{dt}\right)_t, -x_s\left(\frac{dy_s}{dt}\right)_t\right]\cdot[t_x, t_y, 1] = 0 \tag{15}$$

至此可以通过$\left(\frac{ds}{dt}\right)_t$沿$\boldsymbol{s}_n$方向的分量(normal component of translational image motion)求解得$\{t_x, t_y\}$.

Pure Translation

再考虑$\left(\frac{ds}{dt}\right)_t$沿$\boldsymbol{s}$方向的分量(tangential component of translational image motion)可得

$$t_z = -\frac{1}{\Re \sin\varphi}\left(\left(\left(\frac{ds}{dt}\right)_t + \begin{bmatrix} t_x \\ t_y \end{bmatrix}\right)\cdot \boldsymbol{s}\right) \tag{16}$$

如果知道至少一个点的φ就可以确定最后一个平移分量t_z.

General Motion

$$\begin{bmatrix} -(1/\cos\varphi)\mathbf{s}^T + \Re\sin\varphi\,\mathbf{u}^T & \mathbf{0}_{1\times 3} & -(d\mathbf{s}/dt)_s \\ -(1/\cos\varphi)\,\mathbf{s}_n^T & \Re^2\mathbf{u}^T & -(d\mathbf{s}/dt)_n \end{bmatrix} \begin{bmatrix} \mathbf{t} \\ \boldsymbol{\omega} \\ 1 \end{bmatrix} = \mathbf{0}_{2\times 1}$$

需要额外的方法来提供φ的估计.

- *Method L1:* 先计算$\boldsymbol{t}$，再用得到的$\boldsymbol{t}$计算$\boldsymbol{\omega}$.
- *Method L2:* 同时计算$\boldsymbol{t}$和$\boldsymbol{\omega}$.

Partial Motion Set

如果忽略声纳图像运动中的二阶项，只保留一阶项，则

$$\begin{bmatrix} 1 & 0 & x_s \sin\varphi/\mathfrak{R} & -y_s & dx_s/dt \\ 0 & 1 & y_s \sin\varphi/\mathfrak{R} & x_s & dy_s/dt \end{bmatrix} \begin{bmatrix} \boldsymbol{t} \\ \omega_z \\ 1 \end{bmatrix} = 0 \tag{17}$$

若消去φ项

$$\left[y_s, -x_s, -\mathfrak{R}^2, y_s \frac{dx_s}{dt}, -x_s \frac{dy_s}{dt} \right] \cdot \left[t_x, t_y, \omega_z, 1 \right] = 0 \tag{18}$$

Object-Shadow Pairs

$$\frac{d\widehat{\boldsymbol{s}}}{dt} = \frac{\widehat{\mathfrak{R}}}{\mathfrak{R}}\left(\frac{d\boldsymbol{s}}{dt} - \left(\frac{\widehat{\mathfrak{R}}-\mathfrak{R}}{\mathfrak{R}}\right)(\boldsymbol{n}\cdot\boldsymbol{t})\boldsymbol{s}\right) \tag{19}$$

$$(\boldsymbol{s}\boldsymbol{n}^T)\boldsymbol{t} = \left(\frac{\mathfrak{R}}{\widehat{\mathfrak{R}}-\mathfrak{R}}\right)\left(\frac{d\boldsymbol{s}}{dt} - \frac{\mathfrak{R}}{\widehat{\mathfrak{R}}}\frac{d\widehat{\boldsymbol{s}}}{dt}\right) \tag{20}$$

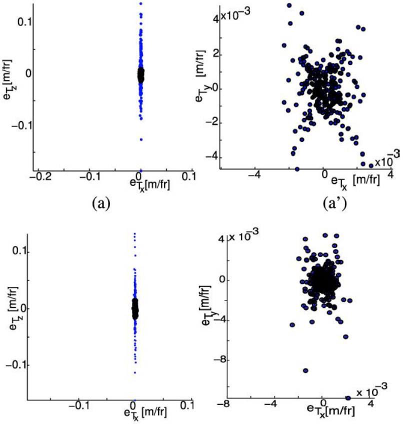
对于一个object-shadow pair，(18)和(20)为四个运动分量$(\boldsymbol{t}, \boldsymbol{\omega}_z)$的求解提供了两个独立的约束.

总结:

- 如果只使用object points，则四个运动分量$(\boldsymbol{t}, \boldsymbol{\omega}_z)$的求解需要至少两个点以及至少一个已知的$\boldsymbol{\varphi}$估计.
- 如果使用object-shadow pair，则四个运动分量$(\boldsymbol{t}, \boldsymbol{\omega}_z)$的求解需要至少两个点对(不需要$\boldsymbol{\varphi}$).

Synthetic Data

- Pure translation: t_x, t_y两个分量的精度更高，因为它们的估计不需要用到φ.

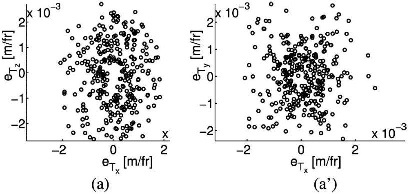


(a) (a’)

(b) (b’)

Synthetic Data

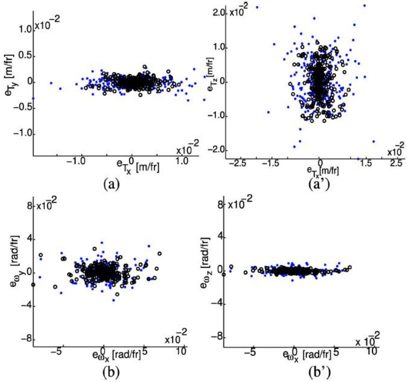
- Pure translation (Object-shadow pair): t_z分量的精度提高.



(a) (a')

Synthetic Data

- General motion:
  - *Method L1:* 先计算$\boldsymbol{t}$，再用得到的$\boldsymbol{t}$计算$\boldsymbol{\omega}$ (black circles).
  - *Method L2:* 同时计算$\boldsymbol{t}$和$\boldsymbol{\omega}$ (blue dots).

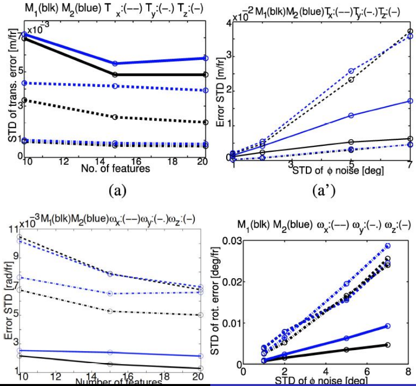


(a) (a’)

(b) (b’)

Synthetic Data

- General motion:
    - *Method L1:* 先计算$\boldsymbol{t}$，再用得到的$\boldsymbol{t}$计算$\boldsymbol{\omega}$ (black circles).
    - *Method L2:* 同时计算$\boldsymbol{t}$和$\boldsymbol{\omega}$ (blue dots).



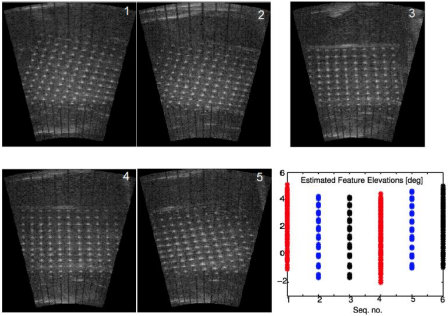
(a)

(a')

Real Data

Real Data Experiment I:

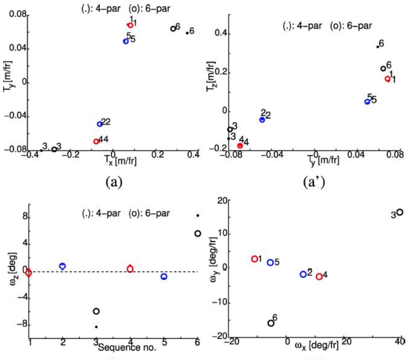
- used motion pairs: 1-2 (seq. 1), 3-4 (seq. 2), 3-5 (seq. 3).
- (reversing the order) 2-1 (seq. 4), 4-3 (seq. 5), 5-3 (seq. 6).
- seq. 3 and seq. 6 are with a larger rotational component.



Real Data

Real Data Experiment I:

- The estimated motion components show consistency in translational $\boldsymbol{t}$ and ω_z rotational components for sequences 1, 2, 4, and 5 but not for 3 and 6.



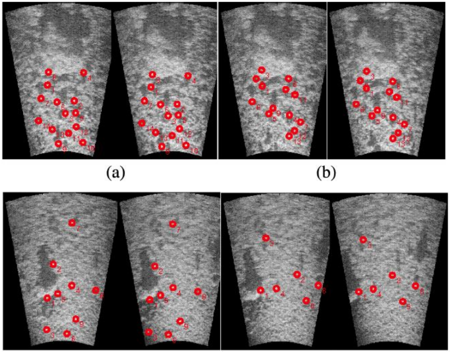
(a) (a’)

(b) (b’)

Real Data

Real Data Experiment II:

- Applying the L1 method to 17 frames from a DIDSON video, recorded at 1.8 MHz frequency setting, in Lake Osceola, on the campus of the University of Miami.



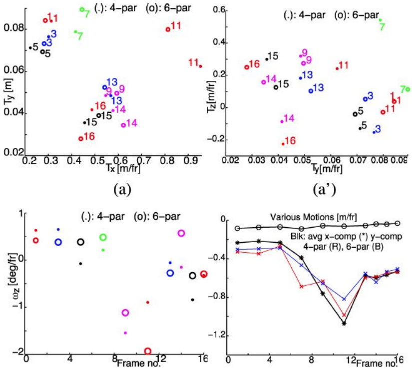
(a) (b)

(c) (d)

Real Data

Real Data Experiment II:

- These behaviors agree with the synthetic data simulations.



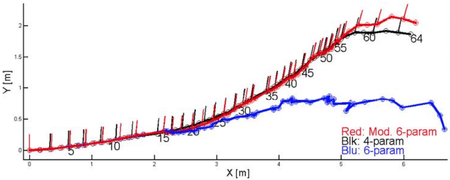
(a) (a')

(b) (c)

Real Data

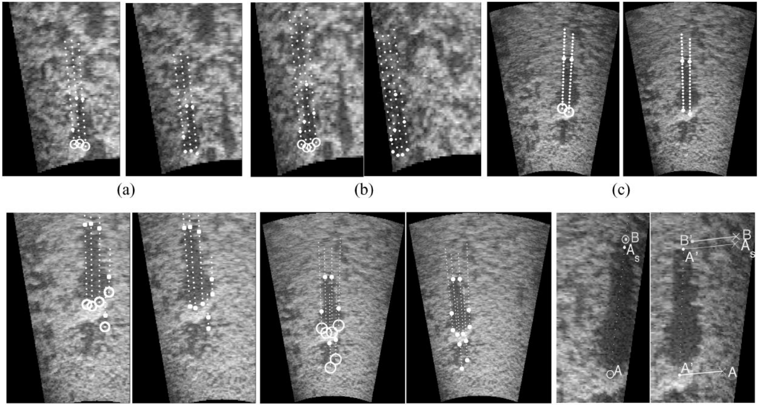
Real Data Experiment III:

- A final experiment utilizes a longer lake sequence, 64 frames at 2 fr/sec data rate, by taking every fifth frame of a 316-frame sequence recorded at 10 [Hz].
- The "drift" of the six-parameter trajectory is primarily due to inaccurate (although small) pitch and roll estimates. To verify, zeroing out these two components from the motions that are estimated by the six-parameter method yields a third trajectory, denoted modified six-parameter solution (red circles).



Shadow Motion Analysis

- The potential application for scene classification.



(a) (b) (c)

(d) (e) (f)