Estimation of Automotive Pitch, Yaw, and Roll using Enhanced Phase Correlation on Multiple Far-field Windows

M. Barnada, C. Conrad, H. Bradler, M. Ochs and R. Mester

Visual Sensorics and Information Processing Lab
Goethe-University, Frankfurt
and
Computer Vision Laboratory
Linköping University, Sweden
Motion estimation

From far-field sub-images at two instants $t - 1$ and $t$, the module determines:

- Global 2D translational motion
- Rate changes in the camera angle (roll/pitch/yaw)
- Rotation matrix between $t - 1$ and $t$

The motion is analysed for each pair of corresponding windows at instants $t - 1$ and $t$
System Overview

For each subwindow

- check structure
- phase correlation
- threshold back projected spectrum
- find multiple translations

If distance too large:

- estimate roll
- check mapping distance
- estimate pitch and yaw
- Intrinsic

Pitch, yaw, roll angles
The motion is only measured on those cells with enough structure information. To find those, we compute the windowed average

\[ \hat{m}_k = \frac{1}{\sum_{x_i} w_T(x_i)} \cdot \sum_{i=1}^M w_T(x_i) \cdot s(x_i) \]  

(1)

and the variance:

\[ \hat{\sigma}_k^2 = \frac{1}{\sum_{x_i} w_T(x_i)} \cdot \sum_{i=1}^M w_T(x_i) \cdot (s(x_i) - \hat{m}_k)^2 \]  

(2)

Further computations are performed only on cells with sufficient structure.

\[ \hat{\sigma}_k^2 \quad \leq \quad T_{var} \quad \text{too weak structure} \]

\[ \text{sufficient structure} \]  

(3)
Determining the translation motion

We use a regularized version of Phase-only-Correlation.

\[
P(f) = \frac{S(f) \cdot G^*(f)}{|S(f) \cdot G^*(f)| + \lambda}
\] (4)

Where \(\lambda\) performs a soft thresholding of the FFT, leaving almost intact the votes from amplitudes larger than \(\lambda\) and giving small weight to the rest.

The displacement \(d\) between cells is determined as the location of the maximum of the inverse Fourier transform of \(P(f)\):

\[
p(x) = F^{-1}\{P(f)\} \approx \delta(x - d)
\] (5)
Due to non-pure translation motion, the back-projected spectrum might not concentrate in one peak. We compute the variance of the noise

$$\sigma_n^2 = \frac{1}{M} \|p(x)\|_F^2$$

(6)

and then we select a subset $\mathcal{P}$ of pixel coordinates belonging to the pixels of the peaks that passed the test

$$p(x) \overset{!}{>} 5 \cdot \sigma_n \land p(x) \overset{!}{>} 0.6 \cdot \max(p(x))$$

(7)

Multiple translations are obtained, one for each cluster of pixels.
Estimating the roll angle

We define the matrix $A$ with the centres of the windows and matrix $B$ with position after applying a candidate translation $d$.

Both sets of points are aligned to the origin of coordinates by

$$A' = A - \frac{1}{k} \sum_{i=1}^{k} c_i \quad \text{and} \quad B' = B - \frac{1}{k} \sum_{i=1}^{k} c'_i$$

(8)
We find the rotation matrix $\mathbf{R}$ that minimizes the element-wise difference of the points in $\mathbf{A}'$ and the rotated points in $\mathbf{B}'$:

$$\min \| \mathbf{A}' \mathbf{R} - \mathbf{B}' \|_F$$  \hspace{1cm} (9)$$

From the sought rotation matrix $\mathbf{R}$:

$$\mathbf{R} = \begin{pmatrix} \cos(\Phi) & -\sin(\Phi) \\ \sin(\Phi) & \cos(\Phi) \end{pmatrix}$$  \hspace{1cm} (10)$$

We obtain the roll angle $\Phi$ according to:

$$\Phi = \arctan (R_{2,1}, R_{1,1})$$  \hspace{1cm} (11)$$
Additionally we determine the average translation $\bar{d}$ between the point sets $A$ and $B$, for latter estimating the pitch and yaw angles:

$$\bar{d} = \frac{1}{k} \sum_{i=1}^{k} (c'_i - c_i) = \frac{1}{k} \sum_{i=1}^{k} d_i$$

(12)

The transformation $T = (R, \bar{d})$ maps the $B$ points into $A$. 

M. Barnada, C. Conrad, H. Bradler, M. Ochs and R. Mester
Testing the validity of the mapping

We validate if the mapping of the point sets is correct:

\[
\frac{1}{k} \sum_{i=1}^{k} \| T \cdot c_i - c'_i \|_2^2 \quad \begin{array}{c}
\text{valid} \\
\text{not valid}
\end{array} \quad T_{\text{dist}} \leq T_{\text{dist}}
\]  

(13)

If the distance is too large, we iteratively repeat the previous process using a leave-one-out strategy and we take the motion vector with the smallest distance.
The pitch $\theta$ and yaw $\Psi$ are computed as:

$$\theta = \arctan\left(\frac{c_x - p_{2x}'}{f_x}\right) - \arctan\left(\frac{c_x - p_{1x}'}{f_x}\right) \quad (14)$$

$$\Psi = \arctan\left(\frac{c_y - p_{2y}'}{f_y}\right) - \arctan\left(\frac{c_y - p_{1y}'}{f_y}\right) \quad (15)$$
Running on KITTI odometry sequences
Off-road sequence
Motion results

Estimation of camera angles for Kitti sequence 3

![Graphs showing roll rate, pitch rate, and yaw rate](image-url)
Motion results

Estimation of camera angles for Kitti sequence 7
# Motion results

Table with results for the KITTI odometry sequences

<table>
<thead>
<tr>
<th>Seq</th>
<th>Valid</th>
<th>$\Delta \Phi$[°]</th>
<th>Corr</th>
<th>$\Delta \Theta$[°]</th>
<th>Corr</th>
<th>$\Delta \Psi$[°]</th>
<th>Corr</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>98.83</td>
<td>0.113</td>
<td>0.657</td>
<td>0.044</td>
<td>0.964</td>
<td>0.129</td>
<td>0.992</td>
</tr>
<tr>
<td>1</td>
<td>99.73</td>
<td>0.072</td>
<td>0.584</td>
<td>0.024</td>
<td>0.95</td>
<td>0.064</td>
<td>0.992</td>
</tr>
<tr>
<td>2</td>
<td>97.02</td>
<td>0.144</td>
<td>0.539</td>
<td>0.052</td>
<td>0.954</td>
<td>0.12</td>
<td>0.987</td>
</tr>
<tr>
<td>3</td>
<td>100.0</td>
<td>0.091</td>
<td>0.73</td>
<td>0.038</td>
<td>0.969</td>
<td>0.09</td>
<td>0.99</td>
</tr>
<tr>
<td>4</td>
<td>100.0</td>
<td>0.136</td>
<td>0.433</td>
<td>0.041</td>
<td>0.875</td>
<td>0.106</td>
<td>0.285</td>
</tr>
<tr>
<td>5</td>
<td>99.53</td>
<td>0.096</td>
<td>0.621</td>
<td>0.032</td>
<td>0.967</td>
<td>0.106</td>
<td>0.991</td>
</tr>
<tr>
<td>6</td>
<td>100.0</td>
<td>0.087</td>
<td>0.518</td>
<td>0.026</td>
<td>0.969</td>
<td>0.106</td>
<td>0.995</td>
</tr>
<tr>
<td>7</td>
<td>98.54</td>
<td>0.091</td>
<td>0.63</td>
<td>0.03</td>
<td>0.968</td>
<td>0.13</td>
<td>0.962</td>
</tr>
<tr>
<td>8</td>
<td>99.48</td>
<td>0.096</td>
<td>0.681</td>
<td>0.034</td>
<td>0.967</td>
<td>0.117</td>
<td>0.992</td>
</tr>
<tr>
<td>9</td>
<td>97.11</td>
<td>0.124</td>
<td>0.502</td>
<td>0.04</td>
<td>0.95</td>
<td>0.121</td>
<td>0.986</td>
</tr>
<tr>
<td>10</td>
<td>96.5</td>
<td>0.124</td>
<td>0.543</td>
<td>0.041</td>
<td>0.972</td>
<td>0.125</td>
<td>0.984</td>
</tr>
</tbody>
</table>

**Table**: Angular errors are measured in degrees
Conclusions

What we achieved:

- A reliable method for determining 2D global motion vectors and rotation angles for a pair of consecutive frames of an automotive sequence.
- An algorithm which can run smoothly on real-time at 140fps (10-7ms/frame) and assist on the prediction of tracked keypoints.
Future work

- Boost the accuracy of the 2D motion estimates by using sub-pixel measurements.
- Measure the motion in more windows, and use depth estimates from SfM of previous frame.
- Use more complex consistency checks to reduce outliers still further.
Questions?
Thank you

M. Barnada, C. Conrad, H. Bradler, M. Ochs and R. Mester

barnada,conrad,bradler,ochs,mester@vsi.cs.uni-frankfurt.de