1 Proof of Equation 8

We first show an elementary result in geometric form, namely the basic geometric equation of an epipolar line and then, intersect two epipolar lines to get the desired result.

**Proposition 1.** Given two images, \((I_1, I_2)\), then the equation of epipolar line corresponding to the point \(\gamma_1 = \begin{pmatrix} \xi_1 \\ \eta_1 \\ 1 \end{pmatrix}\) is

\[
(a_2 b_3 - a_3 b_2) \xi_2 + (a_3 b_1 - a_1 b_3) \eta_2 = (a_2 b_1 - a_1 b_2),
\]

(1)

where \(a_i = e_i^T R_{1,2} \gamma_1\), \(b_i = e_i^T T_{1,2}\), and where \((R_{1,2}, T_{1,2})\) is the relative pose of one camera with respect to another, i.e., a point \(\Gamma_1\) in camera \(I_1\) is expressed as \(\Gamma_2\) in camera \(I_2\) is

\[
\Gamma_2 = R_{1,2} \Gamma_1 + T_{1,2}.
\]

(2)

**Proof.** For \(\Gamma_1\) and \(\Gamma_2\) we have:

\[
\begin{cases}
\Gamma_1 = \rho_1 \gamma_1 \\
\Gamma_2 = \rho_2 \gamma_2
\end{cases}
\]

(3)

where \(\rho_1\) and \(\rho_2\) are the depths of the point in two different views, respectively. Using \(\Gamma_2 = R_{1,2} \Gamma_1 + T_{1,2}\), we have

\[
\rho_2 \gamma_2 = R_{1,2} \rho_1 \gamma_1 + T_{1,2},
\]

(4)

a vector equation with three rows. The third row of Equation 4 is then

\[
\rho_2 = \rho_1 e_3^T R_{1,2} \gamma_1 + e_3^T T_{1,2}
\]

(5)
Also, substituting this value of $p_2$ into the first and second row from Equation 4 given $p_2$, we have

\[
\begin{align*}
\xi_2 &= \frac{\rho_1 e^T R_1 \gamma_1 + e_i^T T_{1,2}}{\rho_1 e^T R_1 \gamma_1 + e_i^T T_{1,2}} \\
\eta_2 &= \frac{\rho_1 e^T R_1 \gamma_1 + e_i^T T_{1,2}}{\rho_1 e^T R_1 \gamma_1 + e_i^T T_{1,2}}
\end{align*}
\] (6)

or

\[
\begin{align*}
\xi_2 &= \frac{\rho_1 a_1 + b_1}{\rho_1 a_3 + b_3} \\
\eta_2 &= \frac{\rho_1 a_2 + b_2}{\rho_1 a_3 + b_3}
\end{align*}
\] (7)

where $a_i = e_i^T R_1 \gamma_1$ and $b_i = e_i^T T_{1,2}$. Isolating $\rho_1$ to eliminate it gives a relationship between $\xi_2$ and $\eta_2$,

\[
\begin{align*}
\rho_1 &= \frac{b_1 - b_3 \xi_2}{a_1 \xi_2 - a_1} \\
\rho_1 &= \frac{b_2 - b_3 \eta_2}{a_3 \eta_2 - a_2}
\end{align*}
\] (8)

Then the epipolar line on $I_2$ is then

\[
(a_2 b_3 - a_3 b_2) \xi_2 + (a_3 b_1 - a_1 b_3) \eta_2 = (a_2 b_1 - a_1 b_2)
\] (9)

where $a_i = e_i^T R_1 \gamma_1$, $b_i = e_i^T T_{1,2}$.

\[\Box\]

**Corollary 1.** Given three images $(I_1$, $I_2$, $I_3)$, and with relative pose $(R_{1,3}, T_{1,3})$ between cameras of $I_1$ and $I_3$ and $(R_{2,3}, T_{2,3})$ between cameras of $I_2$ and $I_3$ and given two corresponding point $\gamma_1$ in $I_1$ and $\gamma_2$ in $I_2$ project to $\gamma_3$ in $I_3$ given as

\[
\Gamma_3 = \begin{pmatrix} \xi_3 \\ \eta_3 \end{pmatrix} = \begin{pmatrix} (a_2 b_3 - a_3 b_2)(\bar{a}_3 \bar{b}_1 - \bar{a}_1 \bar{b}_3) - (a_3 b_1 - a_1 b_3)(\bar{a}_2 \bar{b}_1 - \bar{a}_1 \bar{b}_2) \\ (a_2 b_3 - a_3 b_2)(\bar{a}_3 \bar{b}_1 - \bar{a}_1 \bar{b}_3) - (a_3 b_1 - a_1 b_3)(\bar{a}_2 \bar{b}_1 - \bar{a}_1 \bar{b}_2) \\ \bar{a}_3 \bar{b}_3 - \bar{a}_3 b_2 \xi_2 + (\bar{a}_3 \bar{b}_1 - \bar{a}_1 \bar{b}_3) \eta_2 = (\bar{a}_2 \bar{b}_1 - \bar{a}_1 \bar{b}_2) \end{pmatrix} \] (10)

where $a_i = e_i^T R_1 \gamma_1$, $b_i = e_i^T T_{1,3}$, $\bar{a}_i = e_i^T R_{2,3} \gamma_2$ and $\bar{b}_i = e_i^T T_{2,3}$.

**Proof.** By using Proposition 1, we can build the epipolar constraint in the set $(I_1, I_3)$ and $(I_2, I_3)$, which gives us two epipolar lines:

\[
\begin{align*}
(a_2 b_3 - a_3 b_2) \xi_2 + (a_3 b_1 - a_1 b_3) \eta_2 &= (a_2 b_1 - a_1 b_2) \\
(\bar{a}_2 \bar{b}_3 - \bar{a}_3 b_2) \xi_2 + (\bar{a}_3 \bar{b}_1 - \bar{a}_1 \bar{b}_3) \eta_2 &= (\bar{a}_2 \bar{b}_1 - \bar{a}_1 \bar{b}_2)
\end{align*}
\] (11)

where $a_i = e_i^T R_1 \gamma_1$, $b_i = e_i^T T_{1,3}$, $\bar{a}_i = e_i^T R_{2,3} \gamma_2$ and $\bar{b}_i = e_i^T T_{2,3}$. Then solving this linear system gives the expression of $\gamma_3$:

\[
\Gamma_3 = \begin{pmatrix} (a_2 b_1 - a_1 b_2)(\bar{a}_3 \bar{b}_1 - \bar{a}_1 \bar{b}_3) - (a_3 b_1 - a_1 b_3)(\bar{a}_2 \bar{b}_1 - \bar{a}_1 \bar{b}_2) \\ (a_2 b_1 - a_1 b_2)(\bar{a}_3 \bar{b}_1 - \bar{a}_1 \bar{b}_3) - (a_3 b_1 - a_1 b_3)(\bar{a}_2 \bar{b}_1 - \bar{a}_1 \bar{b}_2) \\ a_3 b_1 - a_1 b_3 \end{pmatrix} \] (12)

\[\Box\]

## 2 Results of Other Tracks

The paper presented results for tracks 00 and 07. Here we augment these tracks with the result of tracks 02, 03, 04, 05, 06, 08 and 09. Figure 1 shows the output paths from different approaches, demonstrating that dissecting scale performs much better than the classical PnP method.
3 Quantification of Improvement in Pose Estimation

A pose estimation, as proposed in this paper, consists of two separate phases: i) the estimation of 5 degrees of freedom, the rotation $R$ and the unit translation vector $\hat{T}$, and ii) the scale of the translation vector, $\lambda$, where the translation vector $T = \lambda \hat{T}$. The justification of improvement in pose estimation is best performed in two phases, one for estimating $(R, \hat{T})$, and the other in estimating $\lambda$. 

Figure 1: A comparison of our approach with the PnP-based method for tracks (a) 02, (b) 03, (c) 05, (d) 04, (e) 06, (f) 08, (g) 09.
Quantification of Improved Up-to-scale Relative Pose: In the experiment we use ground truth scale but estimate \((R, \hat{T})\) using the methods, one as proposed in the paper and one from the classical PnP-based approach. The plot, Figure 2, shows the percentage translation error of two methods with ground truth scale on different tracks.

Figure 2: A comparison of percentage translation errors for our method vs the PnP-based method with ground truth scale.

Quantification of Improved Scale of Translation: In this experiment the ground truth \((R, \hat{T})\) are used to compare the scale \(\lambda\) estimated from our method to that estimated from the PnP-based methods, Figure 3. The percentage error is normalized by ground truth scale. Note that the distribution of our method shows higher frequency for smaller errors and lower frequency for larger errors.

Figure 3: (a): the distribution of percentage error from PnP-based method is significantly higher than (b): the distribution of percentage error from our method.

Figure 4: (a) Stereo visual odometry provides absolute scale and prevents propagation of scale errors. (b),(c) different ways of establishing absolute scale.
4 Average of Forward and Backward Stereo

We proposed in the paper that the stereo odometry scheme can be done in "forward" and "backward" manner as shown in Figure 4. Figure 5(a) shows that the estimates are not correlated closely, thus, they are as independent observes whose average can be used as a better estimation of scale, as shown in Figure 5(b). This shows a slightly better estimation.

![Figure 5](image)

Figure 5: (a) The relative error of forward and backward stereo estimation. They are not correlated so that we can average them to get a better estimation. (b) The comparison of ground truth scale and the scale estimated using averaging.

Figure 6 illustrates the translation error of "forward" manner and the average of "forward" and "backward" manner. A minor improvement, as the figure tells, is achieved by including backward stereo information into the scheme.

![Figure 6](image)

Figure 6: (a) The translation error of two schemes of KITTI00. (b) The translation error of two schemes of KITTI03. (c) The translation error of two schemes of KITTI07.
Figure 7 and 8 show the track of averaging forward and backward scale of KITTI00 and KITTI07. Two results are nearly the same. However, from the magnified tracks, it is obvious that the average scale is slightly closer to the ground truth path.

Figure 7: Left: The output paths from two schemes on KITTI00. Right: Magnified view of the output paths.

Figure 8: Left: The output paths from two schemes on KITTI07. Right: Magnified view of the output path.