An Iterative 5-pt Algorithm for Fast and Robust Essential Matrix Estimation

Vincent Lui

Tom Drummond

Department of Electrical & Computer Systems Engineering, Monash University

The essential matrix, first introduced by Longuet-Higgins [5], is a 3×3 matrix encoding the relative pose information between two views. Conventional approaches for relative pose estimation is to solve a system of linear equations. The 5-pt algorithm [6] is the current state-of-the-art algorithm in relative pose estimation. It is a minimal-set direct solver which solves the essential matrix as a system of polynomial equations.

We show in this paper an iterative method which provides robust and real-time essential matrix estimation, capable of 30Hz, frame-rate performance. The benefit of an iterative approach lies in its simplicity and speed. The use of high degree polynomials may lead to ill-conditioning [1] and are often difficult to solve, leading to alternative methods which sacrifice speed for simplicity [1, 3, 4]. Although convergence is not guaranteed, when used within RANSAC, more hypotheses can be evaluated in the same block of time, yielding improved performance. While iterative solvers which minimizes the algebraic epipolar reprojection error have previously been proposed [2, 7], our parametrization is based on a novel geometric error which incorporates the half-plane constraint [8] and thus enforces orientation consistency between points.

Figure 1 illustrates the concept of our iterative 5-pt algorithm. A coordinate frame is chosen such that the z-axis e_z joins the two camera centres. In this frame, vectors $\hat{v}_i$ and $\hat{v}'_i$ and e_z are coplanar. The goal is to find a rotation for each of the two cameras that maps from their internal coordinate frame to that of Figure 1. We parametrize the image for each camera as a unit 2-sphere, mapping image points in normalized camera coordinates $[x, y, 1]^T$ to unit vectors by dividing by $\sqrt{x^2+y^2+1}$.

At each iteration, the normalized point correspondences $\hat{u}_i \leftrightarrow \hat{u}'_i$ are left multiplied with the rotations R and R', giving the rotated point correspondences $\hat{v} = R\hat{u}$, $\hat{v}' = R'\hat{u}'$. This rotates the two unit spheres, changing the direction of the epipoles. By rotating the unit spheres such that the z-axis e_z is aligned with the epipoles, *i.e.* $e_z = Re = R'e'$, $\hat{v}_i \leftrightarrow \hat{v}'_i$ become coplanar with the epipoles e, e'. The epipoles can then be computed as

$$e = R^T e_z, \quad e' = R'^T e_z. \tag{1}$$

The error function to be minimized is based on the idea that if R and R' are incorrect, there will be a non-zero angle between the plane defined by $\hat{v}_i$ and e_z and the plane defined by $\hat{v}'_i$ and e_z. We compute this angle by projecting $\hat{v}_i$ and $\hat{v}'_i$ onto the x-y plane, using the atan2() function:

$$A(\hat{v}_i) = \mathrm{atan2}(y_i, x_i), \tag{2}$$

where $[x_i \; y_i \; z_i]^T$ is the 3-vector of a rotated point $\hat{v}_i$. The residual error associated with each rotated point correspondence is then given as

$$r_i = A(\hat{v}_i) - A(\hat{v}'_i), \tag{3}$$

We constrain R' to have no component of rotation about its z-axis as two rotations provide 6 degrees of freedom (DOF) but the essential matrix has only 5DOF—there is a rotational gauge freedom about the joint z-axis e_z. This gives 5 parameters α_j, where α_{1-3} represents R and $\alpha_{4,5}$ represents R'. At each iteration, the rotational parameters are updated as

$$R^{k+1} = e^{\sum_{i=1}^{3} \alpha_i G_i} R^k, \quad R'^{k+1} = e^{\alpha_4 G_1 + \alpha_5 G_2} R'^k. \tag{4}$$

where G_i represents the SO(3) Lie group generators. To update the rotation parameters, the Jacobian of the error relative to α_{1-5} has to be computed. Representing the projected point p_i as the 2-vector $[x_i, y_i]$, the Jacobian is given as

$$J = \frac{(x_i)((G_j\hat{v}_i)[1]) - (y_i)((G_j\hat{v}_i)[0])}{x_i^2 + y_i^2}. \tag{5}$$

Using the LM algorithm, the change to be applied to the rotational parameters α_j is $\delta = (J^T J + \lambda I)^{-1} J^T r$. This process is repeated iteratively until the algorithm converges. Using our iterative 5-pt algorithm as a hypothesis generator within RANSAC, the essential matrix can then be recovered simply from the best hypothesis, with the translation $\hat{t}$ given by the second epipole e', and the rotation matrix given by $R'^T R$.

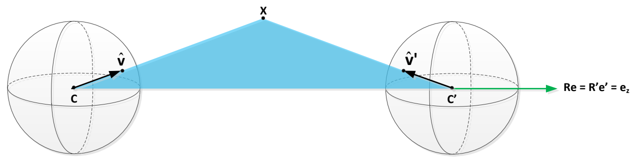

Figure 1: Visual concept of the iterative 5-pt algorithm. Image planes are represented as unit spheres, constraining the epipole to lie on the sphere's surface. Here, the rotated point pairs are co-planar with the epipoles.

In our approach, detected inliers can be easily triangulated using cylindrical coordinates. The vectors $\hat{v}_i$ and $\hat{v}'_i$ can be transformed into cylindrical coordinates by dividing by $\sqrt{x_i^2+y_i^2}$ and $\sqrt{x_i'^2+y_i'^2}$ respectively. In this coordinate frame, $||x_i, y_i|| = 1$ and $||x'_i, y'_i|| = 1$, so the inliers should have $x_i = x'_i$ and $y_i = y'_i$ since inliers have $[x_i, y_i]$ pointing in the same direction as $[x'_i, y'_i]$. The disparity $|z_i - z'_i|$ now encodes the cylindrical radial distance to the point, so setting

$$P_i = \frac{\hat{v}_i}{|z_i - z'_i|} \quad \text{and} \quad P'_i = \frac{\hat{v}'_i}{|z_i - z'_i|} \tag{6}$$

gives the coordinates of the triangulated point in the coordinate frame of Figure 1. This calculation guarantees that $p_i = p'_i \pm e_z$ because the vector e_z represents the motion between the two cameras in the frame of $\hat{v}_i$ and $\hat{v}'_i$. The coordinates of the triangulated point in the original camera frames can then be recovered by simply multiplying P_i and P'_i by R^T and R'^T, *i.e.* $Q_i = R^T P_i$ and $Q'_i = R'^T P'_i$.

Full details of our iterative 5-pt algorithm, including the derivation of the Jacobian, implementation details, and comprehensive experiments. In conclusion, the simplicity of the method, coupled with the ability to easily triangulate inlier points, makes it a useful method for relative pose estimation.

[1] Dhruv Batra, Bart Nabbe, and Martial Hebert. An Alternative Formulation for Five Point Relative Pose Problem. In *Proc. WMVC*, pages 21–27, 2007.

[2] Uwe Helmke, Knut Hüper, Pei Yean Lee, and John Moore. Essential Matrix Estimation Using Gauss-Newton Iterations on a Manifold. *IJCV*, 74(2):117–136, January 2007.

[3] Li Hongdong and Richard Hartley. Five-Point Motion Estimation Made Easy. In *Proc. ICPR*, pages 630–633, 2006.

[4] Zuzana Kukelova, Martin Bujnak, and Tomas Pajdla. Polynomial Eigenvalue Solutions to the 5-pt and 6-pt Relative Pose Problems. In *Proc. BMVC*, 2008.

[5] H.C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–134, 1981.

[6] David Nistér. An efficient solution to the five-point relative pose problem. *TPAMI*, 26(6):756–77, June 2004.

[7] Edward Rosten, Gerhard Reitmayr, and Tom Drummond. Improved RANSAC performance using simple, iterative minimal-set solvers. 2010.

[8] Tomas Werner and Oxford Ox. Constraint on Five Points in Two Images. In *Proc. CVPR*, pages 2–7, 2003.