Handout 03 Local Feature Descriptors and Matching

1. Basic Concepts

(1) SIFT (scale invariant feature transform) is an effective and efficient way to build scale-invariant feature descriptors. It identifies interest points by looking for extrema in the DoG scale space; it figures out the local dominant orientation by using local histogram of orientation; the final descriptor actually is concatenated from 16 histograms, each of which is an 8-bin orientation histogram.

(2) For a normal point, its homogeneous coordinate can be converted to inhomogeneous coordinate by $\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} \Rightarrow \begin{pmatrix} \frac{x'}{z'} \\ \frac{y'}{z'} \end{pmatrix}$ (Note that this is for the 2D case).

(3) The coordinate of the infinity point has the form $(kx_0, ky_0, 0)^T$; all infinity points on the projective plan form an infinity line; two "parallel" projective planes meet at the infinity line; homogeneous equation of the infinity line is z=0; two points determine a line; two lines determine a point; two parallel lines intersect at an infinity point, which means one infinity point corresponds to a specific orientation.

(4) On the projective plane, the line passing two points **x** and **x**' is $\mathbf{l} = \mathbf{x} \times \mathbf{x}'$; the intersection of two lines **l** and **l**' is the point $\mathbf{x} = \mathbf{l} \times \mathbf{l}'$.

(5) Duality principle: to any theorem of projective geometry, there corresponds a dual theorem, which may be derived by interchanging the roles of points and lines in the original theorem.

(6) Hierarchy of the geometric transformations, isometry transformation, similarity transformation, affine transformation, and projective transformation. (refer to our DIP course for more details)

(7) Least square is a method for solving an overdetermined linear system

$$A\mathbf{x} = b, A \in \mathbb{R}^{m \times n}, m > n, rank(A) = n$$

The closed-form solution is $\mathbf{x} = \left(A^T A\right)^{-1} A^T b$

(8) To estimate the homography transform between two images, we need at least four non-degenerate correspondence pairs. In practice, for the consideration of robustness, more than 4 correspondence pairs are used; in this case, we use the RANSAC algorithm to estimate the homography matrix; RANSAC is robust to outliers.

2. Matlab Programming

(1) Study the demo program "HomographyEstimation" provided on our course website. Get two images with sharing content and test the demo.

(2) Study the code provided along with the VLFeat lib. It demonstrates how to align two images by SIFT and RANSAC. http://www.vlfeat.org/applications/sift-mosaic-code.html. Download VLFeat library and make this demo run.