Curvature Based Robust Descriptors

Farlin Mohideen
farlin.mohideen@anu.edu.au
Ranga Rodrigo
ranga@ent.mrt.ac.lk

Australian National University

Department of Electronic and Telecommunication,
University of Moratuwa

Feature descriptors have enabled feature matching under varying imaging conditions, while mostly being backed by experimental evidence. In addition to imposing some restrictions in imaging conditions needed to ensure matching, extending the existing descriptors is not straightforward due to the lack of sound mathematical bases. In this work, by using a surface bending versus shape histogram based on the principal curvatures, we are able to produce a descriptor which is not sensitive to the errors in dominant orientation assignment. Experimental evaluations show that our descriptor outperforms existing descriptors in the areas of viewpoint, rotation, scale, zoom, lighting and compression changes, with the exception of resilience to blur. Further, we apply this descriptor for accuracy demanding applications such as homography estimation and pose estimation. The experimental results show significant improvements in estimated homography and pose in terms of residual error and Sampson distance respectively.

Eigenvalues of Hessian matrix H are the principal curvatures $\lambda_{\max}, \lambda_{\min}$ of a surface $I(x,y)$ at any given point. Let $\vec{p}$ be any point in patch S. We introduce a metric the amount of bending $m(\vec{p})$ based on rotationally invariant principle curvatures as below

$$H = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{bmatrix}, \quad H\vec{v} = \lambda\vec{v}, \quad m(\vec{p}) = \sqrt{\lambda_{\max}^2 + \lambda_{\min}^2}. \tag{1}$$

We propose to represent the dominant orientation by finding where the maximum bending of the surface occurs in a sliding arc-window of $30°$ (Figure 1a). We compute this statistic in a circular patch and with radius proportional to scale (Figure 1a).
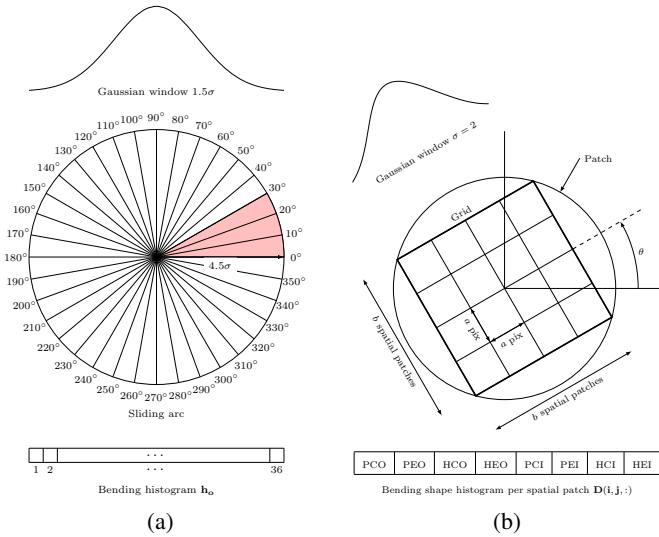

(a)　　　　　　　　(b)

Figure 1: An image patch divided in to 36 arc-windows. A grid superimposed on an image patch, and a grid divided into spatial patches. A Gaussian damping window is overlayed on each patch. (a) Bending histogram for dominant orientation. (b) Descriptor histogram. $a = 3\sigma$ is the width of a spatial patch. Descriptor grid width $b = 4$. Number of classification bins $c = 8$. The gaussian damping window g for descriptor is of variance $\sigma_0 = 2$. To meet these requirements and considering extra patches needed in distributing bending among adjacent bins we need to consider a patch of radius $w = \sqrt{2}a(b+1)/2 + 0.5$.

We compute the histogram h_o of bending of the surface $m(\vec{p})$ (1), vs. polar angle of the pixel $\vec{p}$ w.r.t. keypoint, multiplied by a Gaussian damping window. Thus, the bending histogram $h_o(i)$ is computed where each bin corresponds to the $10°$-arc's total bending of the surface under the influence of the Gaussian. Binning is done by distributing values among adjacent bins by trilinear interpolation. Responses for the $30°$-arc with $10°$ sliding is found by creating the final histogram $\hat{h}_o(i)$ by summing three adjacent bins of $h_o(i)$ by
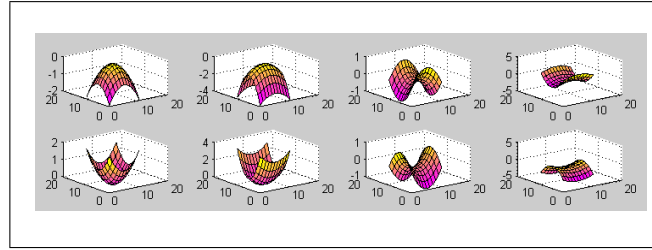

Figure 2: Shape classifications according to parabola (P), hyperbolae (H), corner (C), edge (E), outward (O) and inward (I). (a) PCO (b) PEO (c) HCO (d) HEO (e) PCI (f) PEI (g)HCI (h)HEI

$$\hat{h}_o(i) = \sum_{j=1,...,3} [h_o((i+j) \bmod N)], \quad i = 1,...,N. \tag{2}$$

In $\hat{h}_o(i)$ orientation values are found for the highest peak and those above 75% of the highest, followed by interpolation as in SIFT. Each dominant orientation is used to find descriptors; as in SIFT, one keypoint may have multiple descriptors. In summary, a good descriptor is preferably patch based, the grid width being proportional to the scale, resilient to misorientation due to relying on a metric like curvature which characterizes the shape of the surface at any point.

There are four steps in our feature description: (1) Computing the rotated, normalized spatial patch coordinate frame (2) Surface classification for each patch (3) Generating the descriptor vector and (4) Normalization. We use a 4×4 spatial patch grid for our descriptor, with width 12σ. For each spatial patch, we create a surface bending $m(\vec{p})$ vs shape histogram $D(i, j, :)$ (must see Figure 1b,) based on the eight classifications of the surface (must see Figure 2).

We classify the amount of surface bending according to shapes based on the ratio of principal curvatures (Figure 2). Each spatial patch produces an eight-element descriptor and all 16 spatial patches produce a 128D descriptor.

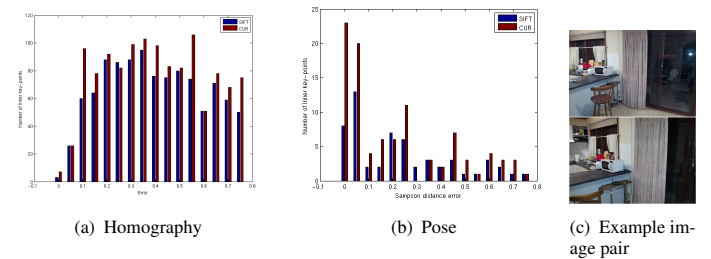(a) Homography　　(b) Pose　　(c) Example image pair

Figure 3: Number of inlier key-points vs residual error for homography estimation and number of inlier key-points vs Sampson distance for pose estimation.

In addition to evaluation of homography estimation we evaluated pose estimation in terms of first-order geometric error (Sampson distance) by $\frac{(x'^T F x)^2}{(Fx)_1^2 + (Fx)_2^2 + (F^T x')_1^2 + (F^T x')_2^2}$. We represent the accuracy by a histogram of number of inlier key-points vs their Sampson distance as shown by Figure 3.b.