Improving Keypoint Orientation Assignment

Steffen Gauglitz
sgauglitz@cs.ucsb.edu
Matthew Turk
mturk@cs.ucsb.edu
Tobias Höllerer
holl@cs.ucsb.edu

Four Eyes Lab
Department of Computer Science
University of California, Santa Barbara

Detection and description of local image features has proven to be a powerful paradigm for a variety of applications in computer vision. Often, this process includes an orientation assignment step to render the overall process invariant to in-plane rotation. In this paper, we propose two novel, very efficient algorithms for orientation assignment and present a detailed quantitative evaluation and analysis of our two algorithms as well as four competing algorithms [1, 2, 4, 5] under a variety of conditions and in combination with six keypoint detectors. In this process, we generalize one existing method [5] and, by doing so, significantly improve its robustness.

"Center-of-Mass" Orientation Assignment. To derive a single orientation, we propose the following simple and very efficient algorithm:

Consider a circular neighborhood R around the keypoint (x_p, y_p) and compute the following sums:

$$c_x = \frac{1}{S} \sum_{(x,y)} w(r) \cdot (x - x_p) \cdot I(x,y) \quad ; \quad c_y = \frac{1}{S} \sum_{(x,y)} w(r) \cdot (y - y_p) \cdot I(x,y)$$

where $r = \|(x,y) - (x_p, y_p)\|$, $w(r)$ is a radial weighting function, and $S = \sum w(r) \cdot I(x,y)$. (c_x, c_y) is the (weighted) "center of mass" (CoM) of the neighborhood if each pixel's intensity is interpreted as its "mass." Its direction relative to the keypoint location is a function of the local intensities that is robust to small shifts in the keypoint location and image noise, and is thus well-suited as the keypoint's orientation.

We evaluated different radial weighting schemes $w(r)$ and describe a method to achieve subpixel-accuracy without using costly interpolation.

"Histogram-of-Intensities" Orientation Assignment. One significant advantage of SIFT's histogram-based approach is that multiple dominant orientations can be extracted. Lowe [4] reports that this improves matching precision significantly, and we confirm this finding in our evaluations.

"Histogram of Intensities" (HoI) was created by combining the intensity-based computation of CoM with SIFT's histogram concept: Each pixel (x,y) in a region R around the keypoint (x_p, y_p) is entered into a histogram based on the angle $\theta(x,y) = \arctan2(y - y_p, x - x_p)$ between the two points, weighted by its intensity $I(x,y)$ and a radial weighting function $w(\|(x,y) - (x_p, y_p)\|)$. From this histogram, dominant orientations can be extracted in the same fashion as done in SIFT.

Despite some structural similarity to SIFT's orientation assignment, a fundamental difference lies in the value that the histogram is based upon (relative orientation to keypoint vs. direction of local gradient). Our algorithm can be implemented much more efficiently since expensive components can be pre-computed.

Evaluation. To evaluate the aforementioned algorithms, we used them in a keypoint matching framework and measured the time to compute the orientation as well as the average precision achieved by the different algorithms under a variety of conditions. We used our dataset published in [3], which contains videos of six different planar textures, each under several camera paths, as well as the ground truth position of the target in each frame.

Keypoints are detected with one of six popular keypoint detectors, then each keypoint gets assigned one or (for SIFT and HoI) multiple orientations using the orientation assignment algorithm under evaluation. An accordingly rotated image patch is used as image descriptor. For comparison, we also include an "oracle" which derives orientations directly from the ground truth, thus indicating what precision could be achieved with perfect orientation assignment given a particular condition. Results are analyzed by detector, image texture, rotation between the images, and size of the neighborhood from which the orientation is estimated.

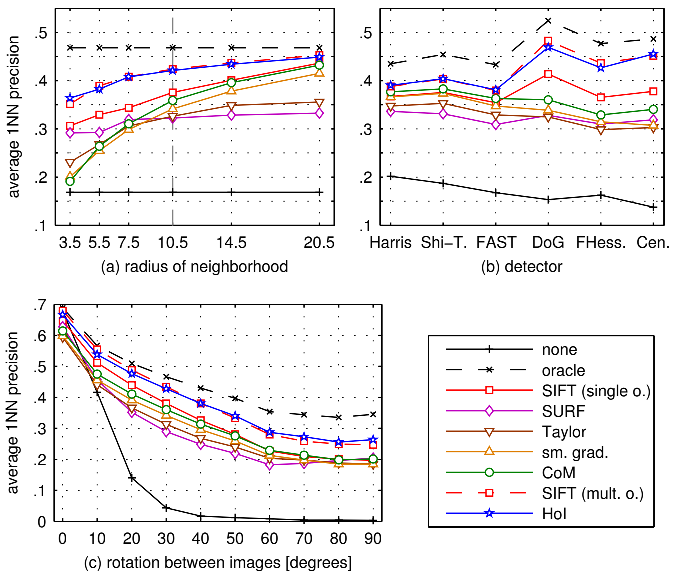

Figure 1: Selected results: Average precision of keypoint matching using an image patch rotated according to the assigned orientation as descriptor (a) as function of the size of the neighborhood R (average for all detectors); (b) for different detectors (radius 10.5); (c) broken down by actual rotation between the images (radius 10.5). Every data point shown in (a) aggregates 3,416,945 keypoint matches.

Results. Our first algorithm, CoM, is very easy to implement, fast and effective. In particular, it is faster to compute than all other methods except Taylor and Drummond's [5] and performs best among all single-orientation methods if used after a corner detector.

In general, the ability to create multiple orientations significantly increases robustness, especially if used with blob detectors. Our second algorithm, HoI, performs very similar to SIFT's orientation assignment (after the latter has been optimized for the metric used) while being significantly cheaper to compute. However, SIFT achieves this performance with fewer orientations per keypoint on average.

Furthermore, we discuss implications about the problem in general as well as for existing algorithms: We analyze the performance/number of orientations trade-off for SIFT, improve Taylor and Drummond's very fast method [5] by using larger rings, and we suggest that the orientation assignment of SURF [1] could be improved by using more samples in the case of large neighborhoods. We also show cases in which orientation assignments are counterproductive.

[1] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Speeded-up robust features (SURF). *CVIU* 2008.

[2] Matthew Brown, Richard Szeliski, and Simon Winder. Multi-image matching using multi-scale oriented patches. *CVPR* 2005.

[3] Steffen Gauglitz, Tobias Höllerer, and Matthew Turk. Evaluation of interest point detectors and feature descriptors for visual tracking. *IJCV* 2011.

[4] David G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV* 2004.

[5] Simon Taylor and Tom Drummond. Binary histogrammed intensity patches for efficient and robust matching. *IJCV* 2011.