

Virtual Line Descriptor and Semi-Local Matching Method for Reliable Feature Correspondence

(Extended abstract — see details in BMVC 2012 full paper)

Zhe Liu

zhe.liu@enpc.fr

Renaud Marlet

renaud.marlet@enpc.fr

University Paris-Est, LIGM (UMR CNRS),

Center for Visual Computing,

Ecole des Ponts ParisTech,

6-8 av. Blaise Pascal, 77455 Marne-la-Vallée, France

Matching detected features in two images based on the similarity of their descriptor often provides good correspondences (inliers). But it often also includes false matches (outliers). Eliminating those false matches while preserving true correspondences remains challenging for images with numerous ambiguities or strong transformations, e.g., due to strong occlusions. In these cases, individual feature matching is not enough; global methods are required, such as RANSAC or graph matching.

However, RANSAC-like methods hardly treat low inlier rates (less than 10%) and, when estimating a fundamental matrix, they cannot eliminate outliers corresponding to points that have matches near their epipolar line. As for graph matchers, they can cope with higher-order constraints (involving more than one match) and optimize a global consistency score. However, most of them are based on geometric constraints rather than photometry. Besides, they are not well suited for a high outlier rate, and their time and space complexity grows exponentially with the order, which limits in practice applications to a few hundred points.

In this paper, we define a 2nd-order photometric descriptor for virtual lines joining two neighbouring feature points. We show it can be used in existing graph matchers to significantly improve their accuracy. We also define a scalable, semi-local matching method based on this descriptor. We show that it is robust to strong transformations and more accurate than existing graph matchers for scenes with significant occlusions, including for very low inlier rates. If used as a preprocessor, it also significantly improves the robustness of RANSAC and reduces camera calibration errors.

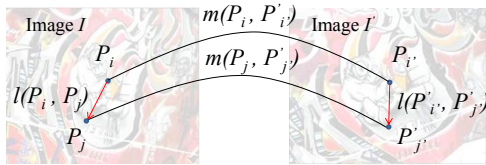


Figure 1: Lines (P_i, P_j) and (P'_i, P'_j) are unlikely to be similar unless matches (P_i, P'_i) and (P_j, P'_j) are correct.

Virtual line descriptor (VLD). Our approach is based on the fact that, for any two points P_i, P_j in image I , and any two points P'_i, P'_j in image I' , it is unlikely to find similar photometric information around lines (P_i, P_j) and (P'_i, P'_j) unless both (P_i, P'_i) and (P_j, P'_j) are correct matches (see Fig. 1). To measure this property, we define a virtual line descriptor (VLD) that captures photometric information between any two points.

We consider a regular disk covering, with overlap, of an image strip between P_i and P_j , and use a SIFT-like descriptor to represent each disk. The covering consists of U disks D_u of radius $r = \frac{\text{dist}(P_i, P_j)}{U+1}$ (see Fig. 2). Each disk is described at image scale $\max(r/r_{\min}, 1)$ where r_{\min} is a minimum description radius. Scales are actually discretized and precomputed. The disk descriptor includes a small-size histogram of gradients and an orientation (see Fig. 2). The global, virtual line descriptor is the concatenation of all disk descriptors. Although it requires less dimensions, VLD inherits SIFT descriptor's robustness to noise and changes of scale, orientation and illumination. Last we define a distance between VLDs that can be used in the pairwise score of a graph matcher.

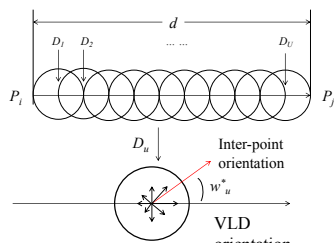


Figure 2: Disk covering of line (P_i, P_j) and histogram of gradient orientation.

K-VLD matching method. We introduce K-VLD, a novel semi-local, 2nd-order matching method. It relies both on geometric and photometric consistency (based on VLD). Two matches (P_i, P'_i) and (P_j, P'_j) are considered *VLD-consistent* iff virtual lines (P_i, P_j) in I and (P'_i, P'_j) in I'

have a VLD-distance less than a given threshold. In the paper, we also define a geometric consistency measure for two matches based on the scale and main orientation of feature points, assuming the local transformation is close to a similitude; if this measure is under a given threshold, then the matches are considered *geometry-consistent*. Finally, two matches are *gVLD-consistent* iff they are both geometry- and VLD-consistent.

The basic idea of K-VLD relies on the fact that, given a potential match (P_i, P'_i) , if there are in the neighborhood of P_i and P'_i at least K other matches $(P_{j_k}, P'_{j_k})_{k \in \{1, \dots, K\}}$ that are gVLD-consistent with (P_i, P'_i) , then (P_i, P'_i) is likely to be a correct match. Given a match m among a set of matches M , the paper defines a notion of neighborhood $\mathcal{N}_M(m)$ whose size adapts to the density of feature points. Experimentally, requiring that good matches have at least K gVLD-consistent neighbors eliminates many outliers, but some may still remain, especially in ambiguous scenes. We found that adding an extra constraint on the proportion of geometry-consistent neighbors and on their average measure of geometric consistency helped in removing many of these remaining outliers.

The K-VLD algorithm starts with all the potential matches and iteratively removes matches that have less than K gVLD-consistent neighbors and matches that do not satisfy the extra geometric constraint, until no match is removed. Ambiguous matches are solved too, based on a preference for matches with many gVLD-consistent neighbors or, when the number of such neighbors is equal, a preference for matches with lowest average VLD-distance. Additional optimizations and heuristics ensure a performance that is in practice quasi linear in the number of matches.

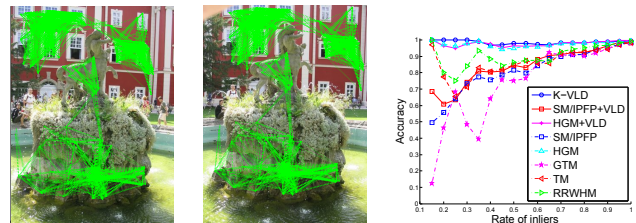


Figure 3: D tenice fountain: K-VLD clusters & average accuracy.

Evaluation. We experimented with various matching methods: probabilistic hypergraph matching (HGM), tensor matching, hypergraph matching via reweighted random walks, spectral matching (SM) / integer projected fixed point, and game-theoretic matching. We also augmented methods SM and HGM with our VLD. And we compared with K-VLD.

We evaluated matching accuracy w.r.t. changing imaging conditions with Mikolajczyk's dataset. K-VLD often outperforms other methods and VLD significantly improves existing methods, especially for scenes with viewpoint or scale changes. We also evaluated the case of strong occlusion with D tenice fountain's dataset (see Fig. 3). K-VLD creates clusters of consistent matches despite occlusions, mostly outperforming other methods. Last we tested K-VLD as a pre-filter to RANSAC-based calibration (ORSA) using Strecha's castle dataset (see Fig.). It considerably improves the quality of match selection. As it can eliminate false matches near epipolar lines, it greatly improves precision, as well as stability. It also substantially reduces the number of iterations, which improves speed.

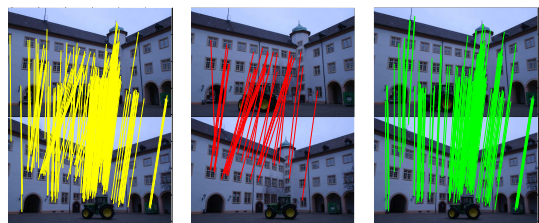


Figure 4: Left: inliers by ORSA. Middle: false matches near epipolar lines by ORSA, rejected by K-VLD. Right: inliers by K-VLD + ORSA.