

Robust Feature Matching with Alternate Hough and Inverted Hough Transforms

Hsin-Yi Chen^{1,2}

¹Academia Sinica, Taiwan

Yen-Yu Lin¹

Bing-Yu Chen²

²National Taiwan University

Abstract

We present an algorithm that carries out alternate Hough transform and inverted Hough transform to establish feature correspondences, and enhances the quality of matching in both precision and recall. Inspired by the fact that nearby features on the same object share coherent homographies in matching, we cast the task of feature matching as a density estimation problem in the Hough space spanned by the hypotheses of homographies. Specifically, we project all the correspondences into the Hough space, and determine the correctness of the correspondences by their respective densities. In this way, mutual verification of relevant correspondences is activated, and the precision of matching is boosted. On the other hand, we infer the concerted homographies propagated from the locally grouped features, and enrich the correspondence candidates for each feature. The recall is hence increased. The two processes are tightly coupled. Through iterative optimization, plausible enrichments are gradually revealed while more correct correspondences are detected. Promising experimental results on three benchmark datasets manifest the effectiveness of the proposed approach.

1. Introduction

Establishing correspondences among two or more images has attracted great attention in the field of computer vision. Being a key component for image analysis and understanding, it is essential for a wide range of applications, such as object recognition [5], image retrieval [30], 3D reconstruction [25], image enhancement [16] and editing [2, 22]. Despite the great applicability, at least two difficulties hinder the advance in establishing correspondences of high quality. First, the predominant paradigm has been starting from local features to yield the candidates of correspondences. Although the design of local descriptors has gained significant progress, methods of this category often suffer from corrupted matches caused by large deformations, illumination changes, or clutter backgrounds. It leads to *low precision* in feature matching. Second, many refined methods, such as [1, 19, 27, 34], tackle this prob-

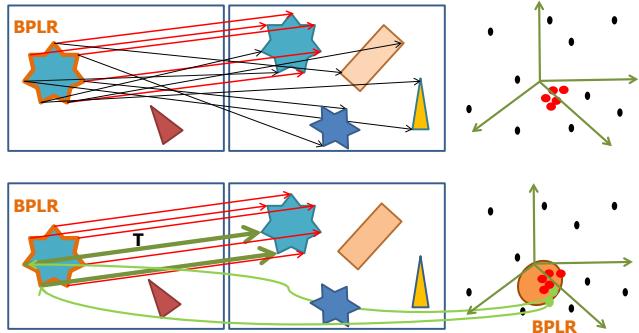


Figure 1. (Top) We project correspondences into the transformation space, and distinguish correct (red) and wrong (black) correspondences by their densities. (Bottom) Potential (green) correspondences are incrementally inferred by exploring density distributions of transformations grouped by BPLRs [17].

lem by ensuring the geometric consistency. These methods typically do not scale very well due to high-order geometric checking. They often work on a pre-selected, small subset of correspondence candidates, and result in *low recall*.

We aim to address the two aforementioned problems simultaneously. Our approach is developed upon the insight that nearby features on the same object typically share similar homographies if they are matched correctly. It follows that their homographies tend to gather together in the *transformation space*. Besides, each wrong matching is usually wrong in its own way. It implies that the *density* of each correspondence in the transformation space can verify its correctness. We leverage this property and cast the task of feature matching into a density estimation problem. Specifically, we identify correct correspondences by comparing the densities among mutually exclusive correspondences, *i.e.*, those violating one-to-one constraints. On the other hand, it is allowed to dynamically *recommend* potential correspondences by exploring the density distributions of locally grouped features. See Figure 1 for an illustration.

The proposed approach carries out *Hough transform* and *inverted Hough transform* alternately to establish robust feature correspondences. It can distinguish itself with the following three main contributions. First, every correspondence candidate is projected into the Hough space

spanned by the transformations. In addition, we grow BPLRs (*Boundary Preserving Dense Local Regions*) [17] for each feature. Only the correspondences associated to features within the same BPLR are considered in *Hough voting*. In this way, mutual verification with relevant correspondences boosts the precision of matching. Furthermore, it makes the complexity of geometric checking independent to the number of correspondences, and leads to one order speed-up in matching. Second, an inverted variant of Hough transform is developed. The *inverted Hough transform* recommends each feature additional transformations by investigating density distributions of nearby features covered by the same BPLR. These transformations enable the dynamical construction of potential correspondences. It allows relevant features to propagate their transformations to each other until consistency is reached. It considerably increases the recall of feature matching. Third, our approach is comprehensively evaluated and compared to the state-of-the-art systems on several benchmark datasets. The superior performance demonstrates its effectiveness.

2. Related Work

The literature of feature correspondence is quite extensive. Our review focuses on those that are relevant to the development of the proposed approach.

Matching via feature descriptor. Point-to-point matching with local feature descriptors is a principal way for correspondence problems. Some of notable researches, *e.g.*, [3, 4, 6, 21], have brought about significant progress in this area. Although these local descriptors are distinctive and powerful, the general conclusion is still that no descriptor is sufficient for handling variation caused by complex combinations of nonrigid deformations, illumination and pose changes, in nowadays vision applications.

Matching via graph partition. One way to address matching ambiguity with additional geometric checking is to cast feature correspondence as a graph matching problem. By defining an objective function based on both photometric similarity and pairwise geometric compatibility between correspondences, promising results via graph matching have been demonstrated [11, 12, 18, 33]. However, these methods typically work well on one common object with simple backgrounds, and do not deal with the cases that multiple sets of common features appear. As mentioned in [20], graph matching is sensitive to corrupt correspondences and outliers. In addition, high computational cost may restrict its applicability, especially when solving a generalized eigenvalue problem is required.

Matching via clustering. Research efforts on clustering-based mechanisms have been made to handle unconstrained matching cases. Bottom-up clustering can integrate locally adaptive constraints to aggregates coherent bundles

of matches. Cho *et al.* [8] carry out object-based image matching via hierarchical agglomerative clustering. Yacov *et al.* [16] adopt a coarse-to-fine scheme and the coherence property of images to achieve dense matching. On the other side, Hai *et al.* [20] propose a top-down clustering approach to detect dense neighborhoods on an affinity graph, and find common visual patterns among images. Despite the effectiveness, one major weakness of these methods lies in their time-complexity. Moreover, the optimal cluster numbers, criteria of cluster merging, and similarity thresholds typically vary from image to image.

Matching via voting. RANSAC [14], a geometric verification model, can be incorporated with local descriptors to enhance the performance. Yuan *et al.* [32] treat each correspondence as a voter, and maintain an affinity matrix to encode how these correspondences vote each other according to their compatibilities. Like RANSAC, their method only supports single object matching. Tolias and Avrithis [26] offer a variant of Hough transform for multi-object matching. They rank the correspondences by adopting the mechanism of *pyramid match* [15]. Their method evenly quantizes the transformation space for fast matching. However, the transformations of correct correspondences often distribute irregularly. It may result in accuracy degradation. Our approach is a voting-based system, and can be distinguished by the advantage that the complexity of Hough voting for each feature is *independent* to the number of correspondences. Furthermore, it dynamically enriches correspondences, and overcomes the *low recall* problem caused by working on a pre-selected, small subset of initial correspondences.

Correspondence enrichment. Most feature correspondence methods work with a small subset of pre-selected correspondences. Correspondence enrichment hence becomes an important task. Match-growing methods, *e.g.*, [11, 13], propagate individual matches to nearby regions based on local appearance, but their performances heavily depend on the quality of initial matching. On the other hand, Cech *et al.* [7] develop a region-growing algorithm to distinguish correct and incorrect correspondences. Cho *et al.* [10] instead describe a progressive graph matching framework to enrich initial matching. However, the yielded correspondences by their approach are biased to the density of features, and may be noisy due to diverse feature distributions in the two matched images. In contrast, our method works on feature bundles guided by BPLRs, so the concerted transformations with high probability are transferred through mutually relevant features. It turns out that the information can be propagated more efficiently and the resulting candidates of correspondences are much more targeted.

3. Problem Definition

Given two images I^P and I^Q , two sets of feature points, $V^P = \{v_i^P\}_{i=1}^{N^P}$ and $V^Q = \{v_i^Q\}_{i=1}^{N^Q}$, are respectively extracted. The region and the center of feature $v_i \in V^P \cup V^Q$ are denoted by S_i and x_i , respectively. The appearance of v_i is described by feature vector \mathbf{u}_i , and its orientation θ_i is estimated by a dominant orientation in the gradient histogram [23]. The product space $\mathcal{C} = V^P \times V^Q$ represents all the possible correspondences. Our goal is to find as many as possible correct correspondences in \mathcal{C} .

3.1. Transformation space

The local shape and the position of feature v_i can be described by a 3×3 matrix $T(v_i)$, which specifies an affine transform of v_i with regards to the *normalized patch* [23]:

$$T(v_i) = \begin{bmatrix} A(v_i) & \mathbf{x}_i \\ \mathbf{0}^\top & 1 \end{bmatrix}, \quad (1)$$

where $A(v_i)$ is a 2×2 non-singular matrix.

Given a feature pair $v_i^P \in V^P$ and $v_{i'}^Q \in V^Q$, the *relative transformation* $H_{ii'}$ from v_i^P to $v_{i'}^Q$ can be derived as

$$H_{ii'} = T(v_{i'}^Q) * T(v_i^P)^{-1}. \quad (2)$$

In this work, we represent a feature *correspondence* as a triplet $m_{ii'} = (v_i^P, v_{i'}^Q, H_{ii'})$, i.e., two features in the opposite images and their relative transformation. As $H_{ii'}$ is a 6-dof affine homography, $m_{ii'}$ can be considered as a point in the 6-dimensional transformation space.

3.2. Distance metric in the transformation space

Given two correspondences $m_{ii'} = (v_i^P, v_{i'}^Q, H_{ii'})$ and $m_{jj'} = (v_j^P, v_{j'}^Q, H_{jj'})$, the projection error of $m_{jj'}$ with respect to $m_{ii'}$ can be defined as

$$d_{jj'|ii'} = \|\mathbf{x}_{j'}^Q - \rho(H_{ii'} \begin{bmatrix} \mathbf{x}_j^P \\ 1 \end{bmatrix})\|, \quad (3)$$

$$\text{where } \rho(\begin{bmatrix} a & b & c \end{bmatrix}^\top) = \begin{bmatrix} a/c & b/c \end{bmatrix}^\top. \quad (4)$$

It checks if $H_{ii'}$ projects \mathbf{x}_j^P around $\mathbf{x}_{j'}^Q$.

For a pair of correspondences $m_{ii'}$ and $m_{jj'}$, they are considered compatible if the corresponding homographies are similar. We hence adopt the *re-projection error* for dissimilarity measure, i.e.,

$$d(m_{ii'}, m_{jj'}) = \frac{1}{4}(d_{jj'|ii'} + d_{ii'|jj'} + d_{j'j|i'i} + d_{i'i|j'j}). \quad (5)$$

Note that it is symmetric and is used to compute the distances among correspondences in the transformation space.



Figure 2. The procedure of the proposed approach.

4. The Proposed Approach

Features with compatible geometric configurations are mutually dependent in matching. We investigate feature dependence via BPLR detector [17], and cast feature matching as a density estimation problem. The proposed approach carries out this idea by alternate Hough and inverted Hough voting. While the former discovers the consistent homographies by projecting correspondences into the transformation space, the latter incrementally recommends potential correspondences driven by the concerted homographies. The procedure of our approach is illustrated in Figure 2. In the following, we first describe the construction of initial correspondence. Then the Hough and inverted Hough transforms for feature matching are introduced, respectively.

4.1. Initial correspondence candidates

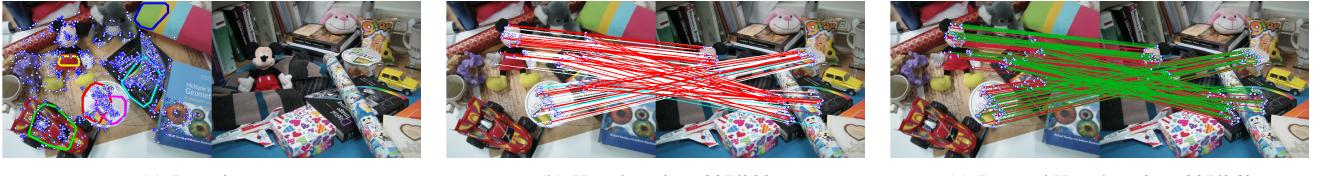
Our approach starts from the construction of initial correspondence candidates. For each feature $v_i^P \in I^P$, we find its r potential matchings $\{v_{i_k}^Q\}_{k=1}^r$ in I^Q according to their appearance similarity and with the constraint that none of the r matchings highly overlap. This can be accomplished by sorting the similarity scores and sequentially excluding those that have strong overlapping with the prior ones. Specifically, we evaluate the similarity between two features by the SIFT descriptor, i.e., $\|\mathbf{u}_i^P - \mathbf{u}_{i_k}^Q\|$, while overlapping is measured by the area of intersection divided by the area of union, i.e., $S_i^P \cap S_{i_k}^Q / S_i^P \cup S_{i_k}^Q$. With $\{v_{i_k}^Q\}_{k=1}^r$, the set of initial correspondences associated with v_i^P is

$$\mathcal{M}_i = \{m_{ii_k} = (v_i^P, v_{i_k}^Q, H_{ii_k})\}_{k=1}^r, \quad (6)$$

where H_{ii_k} is the relative transformation from v_i^P to $v_{i_k}^Q$. This process is repeated for each feature in I^P . Then the set of initial correspondences is constructed by

$$\mathcal{M} = \bigcup_{i=1}^{N^P} \mathcal{M}_i. \quad (7)$$

The initial set \mathcal{M} is of size $|\mathcal{M}| = r \times N^P$. It contains many corrupted matchings since there exists at most one correct correspondence in each \mathcal{M}_i . In complex matching tasks, it is usually the case that only a small subset of correct correspondences in \mathcal{C} is included in \mathcal{M} . Empirically, we set $r = 5$ in all the experiments, because the precision of correspondences decreases rapidly when r is larger than 5.



(a) Input images

(b) Hough voting: 207/222

(c) Inverted Hough voting: 337/369

Figure 3. Feature matching by our approach. (a) Input images I^P and I^Q , together with all the feature points and some of the BPLRs (contours) in I^P . (b) Hough voting and its comparison with SIFT. 207 out of 222 correct correspondences in \mathcal{M} are identified via Hough voting. White lines denote the correct correspondences detected by both approaches. Red and cyan lines are the correct correspondences by only Hough voting and the nearest SIFT searching, respectively. (c) Inverted Hough voting. It recommends 147 ($= 369 - 222$) correct candidates and leads to additional 130 ($= 337 - 207$) correct correspondences (green lines) detected by the successive Hough voting.

4.2. Hough transform for homography verification

The goal at this stage is to detect the correct correspondences in \mathcal{M} , which is either the initial correspondence set or the enriched set by the following stage. We investigate the property that the transformations of correct correspondences are concerted while those of incorrect correspondences are different in their own ways. Hough voting for homography verification is employed since it can handle a high percentage of incorrect correspondences and detect correct correspondences via density estimation. Specifically, the relative transformation of each correspondence is treated as a point in *Hough space*, and it is considered as a hypothesis about the underlying homography of interest.

Despite its robustness, Hough transform is developed upon the assumption that the hypotheses are a sum of independent votes, and thereby neglects the spatial dependence among features. As pointed out in [31], choosing proper *voters* is critical in Hough transform, especially when voters are dependent. We are inspired by the fact that nearby features on the same object are mutually dependent, and group relevant correspondences via BPLR detector [17], which respects object boundary and captures the local shape of an object. It turns out that the performance of Hough voting is remarkably boosted. Furthermore, only relevant, small-size correspondences are involved in density estimation, instead of the whole \mathcal{M} . It significantly speeds up the process.

To formalize, let $\mathcal{B} = \{b_e\}$ be the set of the detected BPLRs in image I^P . For each feature $v_i^P \in I^P$, we use $\pi(v_i^P) \subseteq \mathcal{B}$ to denote the set of BPLRs that cover the center of v_i^P . For v_i^P that is not covered by any BPLR, *i.e.*, $\pi(v_i^P) = \emptyset$, we just simply assign it to the nearest BPLR, though this case rarely occurs because BPLRs are densely sampled. We then cluster features relevant to v_i^P by checking if they reside in at least one common BPLR, *i.e.*,

$$G(v_i^P) = \{v_j^P | \pi(v_i^P) \cap \pi(v_j^P) \neq \emptyset\}. \quad (8)$$

We assume that the grouped features with high probability undergo similar transformations in matching. It follows that the correspondences relevant to v_i^P in Hough voting can

be collected by

$$R(v_i^P) = \bigcup_{v_j^P \in G(v_i^P)} \mathcal{M}_j. \quad (9)$$

According to Eq.(6), there exists at most one correct correspondence in \mathcal{M}_i . Hough voting as well as voters $R(v_i^P)$ are adopted to pick the most plausible correspondence associated with feature v_i^P . Specifically, it is accomplished by *normalized kernel density estimation* (KDE):

$$m_{ii'}^* = \underset{m_{ii'} \in \mathcal{M}_i}{\operatorname{argmax}} \frac{1}{|R(v_i^P)|} \sum_{m \in R(v_i^P)} \exp\left(-\frac{d(m_{ii'}, m)}{\sigma}\right), \quad (10)$$

where σ is a positive constant whose value is set as the average distance among the existing correspondences. Note that the normalization term $1/|R(v_i^P)|$ does not affect the result in Eq.(10), but it is required in comparing densities across feature points.

The procedure of correspondence selection is repeated for each feature in image I^P . It results in N^P selected correspondences $\mathcal{M}^* = \{m_{ii'}^*\}_{i=1}^{N^P}$. We then sort them according to their associated densities in Eq.(10), and return the top correspondences by a proper threshold. In the experiments, our approach is evaluated by precision-recall curves, plotted with various thresholds. An example of the verification results by Hough voting is shown in Figure 3b.

4.3. Inverted Hough transform for correspondence recommendation

While Hough transform identifies correct correspondences $\mathcal{M}^* \subseteq \mathcal{M}$ and boosts the precision in matching, the goal of inverted Hough transform is to enrich \mathcal{M} so that the recall can be increased. The locally clustered features by BPLRs have consensus transformations and can assist each other in finding plausible correspondences. We investigate this property and develop the inverted Hough transform, which allows grouped features to propagate their homographies to each other and recommends each feature concerted correspondences by exploring the propagated homographies.

Algorithm 1 The procedure of the proposed framework

```

1: Input: Feature sets  $V^P$  and  $V^Q$ ; Max iteration  $T$ 
2: Output: Matched correspondences  $\mathcal{M}^*$ 
3: Initialize correspondence sets  $\{\mathcal{M}_i\}_{i=1}^{N^P}$  via (6)
4: while  $t < T$  do
5:    $M^* \leftarrow \emptyset$ 
6:   for all  $v_i^P \in V^P$  do
7:     Detect correspondence  $m_{ii'}^* \in \mathcal{M}_i$  via (10);
8:      $\mathcal{M}^* \leftarrow \mathcal{M}^* \cup m_{ii'}^*$ ;
9:   for all  $v_i^P \in V^P$  do
10:    Identify recommended feature  $v_k^Q$  via (13);
11:    Construct  $m_{ik} = (v_i^P, v_k^Q, H_{ik})$ ;
12:     $\mathcal{M}_i \leftarrow \mathcal{M}_i \cup m_{ik}$ ;
13: Sort elements in  $\mathcal{M}^*$  with thresholding;

```

For each $v_i^P \in I^P$, we search the relevant features, $G(v_i^P)$ in Eq.(8). Each of these features delivers a hypothesis about the homography of v_i^P . These hypotheses are collected in

$$\tilde{M}_i = \{m_{jj'}^* | v_j^P \in G(v_i^P)\}, \quad (11)$$

where $m_{jj'}^*$ is the selected correspondence of v_j^P through Hough voting. The set \tilde{M}_i may contain outliers caused by corrupted matchings. Hence, we pick the homography of the most plausible correspondence $\tilde{m}_{jj'} \in \tilde{M}_i$ for recommendation, where

$$\tilde{m}_{jj'} = \operatorname{argmax}_{m_{jj'}^* \in \tilde{M}_i} \sum_{m \in \tilde{M}_i} \exp\left(-\frac{d(m_{jj'}, m)}{\sigma}\right). \quad (12)$$

Suppose that the relative transformation of $\tilde{m}_{jj'}$ is $H_{jj'}$. The projected region of v_j^P from I^P to I^Q through $H_{jj'}$ is denoted by S . The matching feature in I^Q is determined by

$$v_k^Q = \operatorname{argmax}_{v_k^Q \in V^Q} \frac{S \cap S_k^Q}{S \cup S_k^Q}. \quad (13)$$

It follows that correspondence $m_{ik} = (v_i^P, v_k^Q, H_{ik})$ is recommended, i.e., $\mathcal{M}_i \leftarrow \mathcal{M}_i \cup m_{ik}$. This process is done for each feature in I^P . The resulting correspondence set \mathcal{M} in Eq.(7) is incrementally enriched. An example of inverted Hough transform is given in Figure 3c.

Hough voting and its inverted variant are tightly coupled. While the former densely detects correct correspondences from the enriched candidates, the latter provides harmonic enrichment owing to better detection results. The convergency of the alternate voting procedure is guaranteed. The number of all the correspondences, i.e., \mathcal{C} , is finite and fixed. At each iteration, the number of correspondence candidates, i.e., \mathcal{M} , is monotonically strictly increasing. Since \mathcal{M} is a

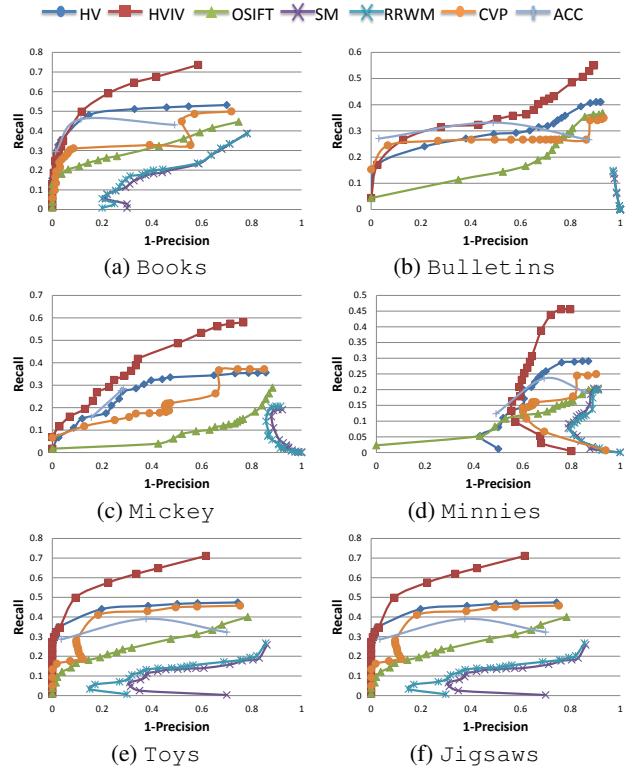


Figure 4. Comparison among various approaches on each test image of the SNU dataset. Precision is the fraction of correct correspondences among the detected matches. Recall is the fraction of correct detections among the correct correspondences in \mathcal{C} .

subset of \mathcal{C} , the iterative procedure must converge. Empirically, it rapidly converges after a few iterations, typically $2 \sim 4$, in all our experiments. We conclude this section by summarizing our approach in Algorithm 1.

5. Experimental Results

We conduct three sets of experiments for performance evaluation. First, our approach is compared with other state-of-the-art techniques in identifying multiple common objects. Second, we show the advantages of our approach in progressive enrichment of correspondence candidates. Third, we show that our approach can collaborate with different detectors and descriptors, and match features across images with large illumination changes.

5.1. Matching with multiple common objects

The SNU dataset [11] is used in the experiments. It contains six image pairs. The appearance of multiple common objects, partial occlusions, and clutter backgrounds make matching over this dataset quite challenging. However, it provides a good test bed to manifest the importance of geometric verification and correspondence recommendation, since the initial correspondences are not reliable enough.



Figure 5. The matching results of various approaches on (a) ~ (c) image Books, (d) ~ (f) image Bulletins, and (g) ~ (i) image Toys. In each figure, the approach as well as its performance (correct detections / correct candidates in \mathcal{M}) are shown.

We follow [24] and consider a correspondence to be correct if the area of the intersection of the predicted and true regions divided by the area of the union of the two regions is larger than 40%. We measure the performance of feature matching by both precision and recall. While precision is the fraction of detected correspondences that are correct, recall is the fraction of correct correspondences that are detected. For each adopted approach, all the detected correspondences are ranked by its own criterion, such as the element values of the eigenvector in spectral matching [18] or the estimated density Eq.(10) in our approach. With proper thresholds, the performance of each approach is presented by a precision-recall curve.

Our approach, *Hough voting and inverted Hough voting* (HVIV), is compared with other state-of-the-art systems, including descriptor-based approach, *opponent SIFT* (OSIFT) [28], clustering-based approaches, *common visual pattern discovery* (CVP) [20] and *agglomerative correspondence clustering* (ACC) [8], and graph-based approaches, *spectral matching* (SM) [18] and *reweighted random walks* (RRWM) [9]. Besides, we implement one additional baseline *Hough voting* (HV), which carries out only the Hough transform part of our approach. For the sake of fair comparison, all the approaches work on the same feature points, each of which is detected by the Hessian affine detector [23] and depicted by the opponent SIFT descriptor [28]. The initial correspondence set of all the approaches is selected according to the nearest search of the opponent SIFT.

The quantitative results are summarized in Figure 4. The performances of the approaches based on graph matching,

i.e., SM and RRWM, are not stable due to their sensitivity to outliers. In this dataset, wrong correspondences are often far more than correct ones in initialization. The baseline OSIFT does not work well, since the unary local features are not sufficient to handle clutter backgrounds and complex deformations and transformations of objects. Instead, CVP, ACC, HV and HVIV, improve the performances by investigating geometric consistence. Among them, our HVIV considerably outperforms CVP, ACC and HV in most cases. This is because our approach clusters mutually relevant features for correspondence verification and recommendation. It avoids the adverse effect of distracting outliers, resolves the limitation of initial candidate construction, and thus achieves both high precision and recall. Some matching results of various approaches are shown in Figure 5. It can be observed that our approach gives to more accurate and dense correspondences.

5.2. Incremental correspondence enrichment

In this experiment, we show the advantages of our approach in incremental correspondence enrichment, and compare it with baseline *Hough voting* (HV) and the *progressive graph matching framework* (PROG) [10], which is one of the best approaches in correspondence enrichment.

The benchmark dataset collected by Cho *et al.* [10] is adopted for performance evaluation. It is composed of 30 image pairs. Each of them has one common object for matching. The experiments are conducted with the same settings as those described above, including the use of the same local feature detector, descriptor and evaluation crite-

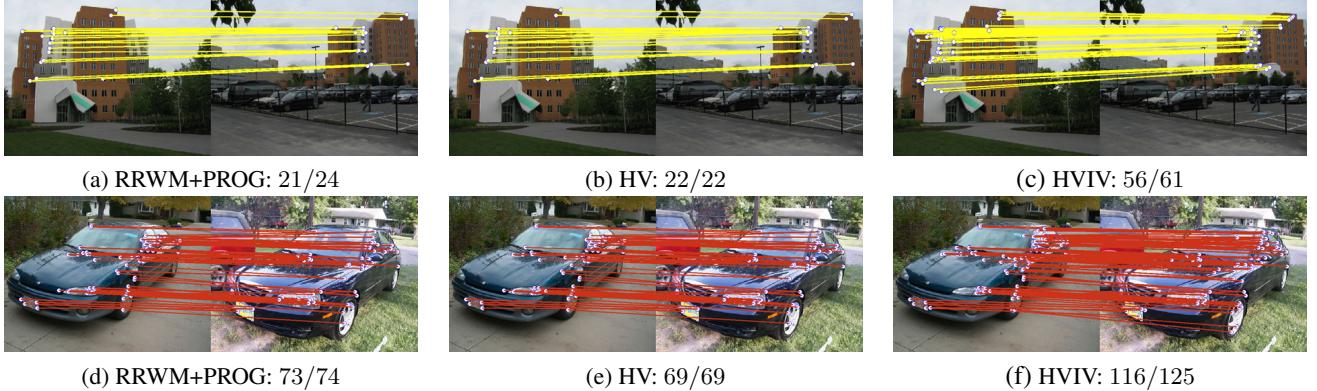


Figure 6. The matching results by RRWM+PROG, HV and our HVIV. While RRWM+PROG and HV detect almost all the correct correspondences, our approach further improves the results by progressively enriching the correspondence set.

ria. As PROG provides a general mechanism for progressive graph matching, it can work with any graph-based approaches. RRWM [9] is adopted here as the graph-matching module of PROG. The parameters k_1 and k_2 pertain to PROG are set as 10 and 5, respectively.

With the same detector, descriptor, matching criterion, and initial correspondences, Figure 6 displays the correct correspondences detected by RRWM+PROG, HV and our HVIV, respectively. Note that the results obtained by RRWM+PROG are different from the ones reported in original paper, this may be because the matching results are sensitive to the adopted detector, descriptor and matching criterion.

It can be observed that HVIV remarkably increases the number of true candidate matches, and results in better matching outcomes. This is because the objective functions of correspondence verification and recommendation steps in our approach are both derived upon densities and hence coherent. The two steps complement each other to jointly lead to better results. Besides, the grouped features by BPLRs faithfully identify relevant voters in both Hough and inverted Hough transforms. It excludes distracting voters and provides recommendations of high quality. With the same precision, our approach averagely achieves 54.0% growth rate in true correspondences when comparing with HV, and 44.4% growth rate when comparing with PROG.

5.3. Plug-in with other feature descriptors

Our approach can be treated as a geometric filter. It drops the corrupted correspondences by geometric checking and enhances the matching by propagating concerted transformations among dependent features. It can be applied to various types of feature descriptors, and improve the performance. In the experiment, our approach collaborates with the *LIO*P descriptor [29], and establishes correspondences across images with drastic illumination changes.

We perform the quantitative analysis on complex illumination dataset used in [29]. It contains two image

pairs Desktop and Corridor. Besides, the image pair Leuven with exposure change from Oxford dataset¹ is also adopted. For each image pair, the Hessian-affine detector is used to localize feature positions. The initial matching candidates are constructed by LIO, which is designed to be invariant to dramatic illumination changes.

HV and our HVIV are applied to the correspondences discovered by LIO. The results in form of precision-recall curves are plotted in Figure 7. Despite the robustness to illumination changes, the performance of LIO can still be enhanced by ensuring homography consistency and enriching correspondence candidates.

6. Conclusion and Future Work

We have presented a simple but effective approach that carries out alternate Hough voting and its inverted variant to establish correspondences in complex matching tasks, and boosts the performances in both precision and recall. It formulates feature matching as a density estimation problem. Through iterative optimization, more correct correspondences are detected from the enriched candidates, while plausible enrichments are gradually revealed by the propagated beliefs in the concerted homographies. Besides, we group mutually dependent features via BPLRs. It not only increases the accuracy but also speeds up the process. The proposed approach is comprehensively evaluated on three datasets coupled with different descriptors. The promising results consolidate the usefulness of our approach. For future work, we will apply our approach to handling co-segmentation, image synthesis and reconstruction. This is because high-quality, dense matches generally facilitate the accomplishments of these applications.

Acknowledgments. We thank the anonymous reviewers for their comments. This work is supported in part by grants NSC 101-2221-E-001-018 and 101-2622-E-002-006-CC2.

¹<http://www.robots.ox.ac.uk/vgg/research/affine/>

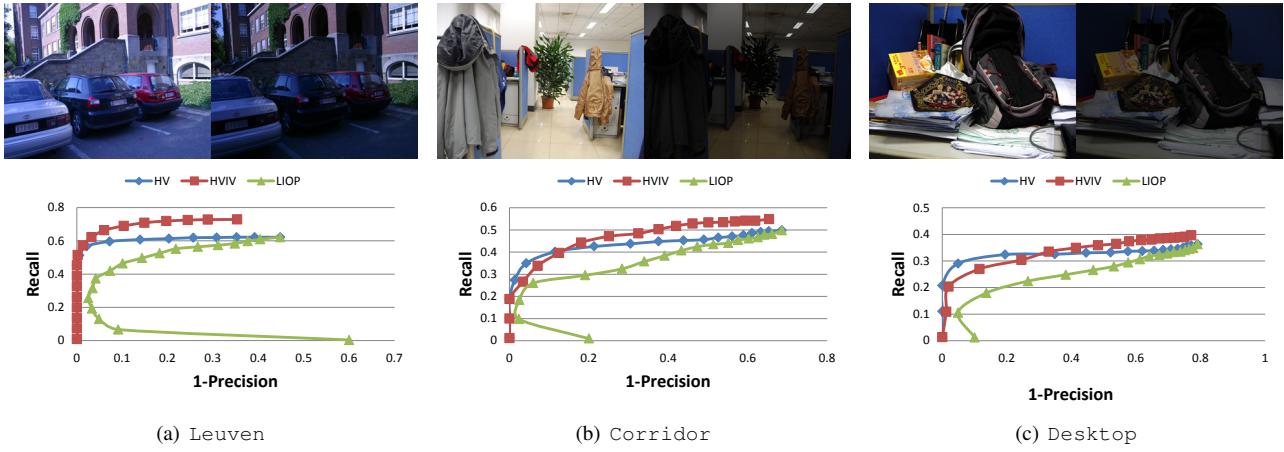


Figure 7. Plug-in comparison with the LIOP descriptor on three image pairs.

References

- [1] A. Albarelli, E. Rodolà, and A. Torsello. Imposing semi-local geometric constraints for accurate correspondences selection in structure from motion: A game-theoretic perspective. *IJCV*, 97(1):36–53, 2012.
- [2] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman. Patch-Match: A randomized correspondence algorithm for structural image editing. *ACM TOG*, 28(3):24:1–24:11, 2009.
- [3] H. Bay, T. Tuytelaars, and L. V. Gool. SURF: Speeded up robust features. In *ECCV*, volume 1, pages 404–417, 2006.
- [4] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE TPAMI*, 24(4):509–522, 2002.
- [5] A. C. Berg, T. L. Berg, and J. Malik. Shape matching and object recognition using low distortion correspondence. In *IEEE CVPR*, volume 1, pages 26–33, 2005.
- [6] A. C. Berg and J. Malik. Geometric blur for template matching. In *IEEE CVPR*, volume 1, pages 607–614, 2001.
- [7] J. Cech, J. Matas, and M. Perdoch. Efficient sequential correspondence selection by cosegmentation. *IEEE PAMI*, 32(9):1568–1581, 2010.
- [8] M. Cho, J. Lee, and K. M. Lee. Feature correspondence and deformable object matching via agglomerative correspondence clustering. In *IEEE ICCV*, pages 1280–1287, 2009.
- [9] M. Cho, J. Lee, and K. M. Lee. Reweighted random walks for graph matching. In *ECCV*, volume 5, pages 492–505, 2010.
- [10] M. Cho and K. M. Lee. Progressive graph matching: Making a move of graphs via probabilistic voting. In *IEEE CVPR*, pages 398–405, 2012.
- [11] M. Cho, Y. M. Shin, and K. M. Lee. Co-recognition of image pairs by data-driven monte carlo image exploration. In *ECCV*, volume 4, pages 144–157, 2008.
- [12] T. Cour, P. Srinivasan, and J. Shi. Balanced graph matching. In *NIPS*, pages 313–320, 2006.
- [13] V. Ferrari, T. Tuytelaars, and L. van Gool. Simultaneous object recognition and segmentation by image exploration. In *ECCV*, volume 1, pages 40–54, 2004.
- [14] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography. *CACM*, 24(6):381–395, 1981.
- [15] K. Grauman and T. Darrell. The pyramid match kernel: Discriminative classification with sets of image features. In *IEEE ICCV*, volume 2, pages 1458–1465, 2005.
- [16] Y. HaCohen, E. Shechtman, D. B. Goldman, and D. Lischinski. Non-rigid dense correspondence with application for image enhancement. *ACM TOG*, 30(4):70:1–70:9, 2011.
- [17] J. Kim and K. Grauman. Boundary preserving dense local regions. In *IEEE CVPR*, pages 1153–1160, 2011.
- [18] M. Leordeanu and M. Hebert. A spectral technique for correspondence problems using pairwise constraints. In *IEEE ICCV*, volume 2, pages 1482–1489, 2005.
- [19] M. Leordeanu, M. Hebert, and R. Sukthankar. An integer projected fixed point method for graph matching and map inference. In *NIPS*, pages 1114–1122, 2009.
- [20] H. Liu and S. Yan. Common visual pattern discovery via spatially coherent correspondences. In *IEEE CVPR*, pages 1609–1616, 2010.
- [21] D. Lowe. Object recognition from local scale-invariant features. In *IEEE ICCV*, volume 2, pages 1150–1157, 1999.
- [22] D. Mahajan, F.-C. Huang, W. Matusik, R. Ramamoorthi, and P. Belhumeur. Moving gradients: A path-based method for plausible image interpolation. *ACM TOG*, 28(3):42:1–42:12, 2009.
- [23] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. *IJCV*, 60(1):63–86, 2004.
- [24] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE TPAMI*, 27(10):1615–1630, 2005.
- [25] N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: Exploring photo collections in 3D. *ACM TOG*, 25(3):835–846, 2006.
- [26] G. Tolias and Y. Avrithis. Speeded-up, relaxed spatial matching. In *IEEE ICCV*, pages 1653–1660, 2011.
- [27] L. Torresani, V. Kolmogorov, and C. Rother. Feature correspondence via graph matching: Models and global optimization. In *ECCV*, volume 2, pages 596–609, 2008.
- [28] K. van de Sande, T. Gevers, and C. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE TPAMI*, 32(9):1582–1596, 2010.
- [29] Z. Wang, B. Fan, and F. Wu. Local intensity order pattern for feature description. In *IEEE ICCV*, pages 603–610, 2011.
- [30] Z. Wu, Q. Ke, M. Isard, and J. Sun. Bundling features for large scale partial-duplicate web image search. In *IEEE CVPR*, pages 25–32, 2009.
- [31] P. Yarlagadda, A. Monroy, and B. Ommer. Voting by grouping dependent parts. In *ECCV*, volume 5, pages 197–210, 2010.
- [32] Y. Yuan, Y. Pang, K. Wang, and M. Shang. Efficient image matching using weighted voting. *PRL*, 33(4):471–475, 2012.
- [33] R. Zass and A. Shashua. Probabilistic graph and hypergraph matching. In *IEEE CVPR*, pages 1–8, 2008.
- [34] F. Zhou and F. D. Torre. Factorized graph matching. In *IEEE CVPR*, pages 127–134, 2012.