

A Robust Graph-Based Method for The General Correspondence Problem Demonstrated on Image Stitching

Martin Bujňák and Radim Šára
Center for Machine Perception
Czech Technical University, Prague
{bujnam1, sara}@cmp.felk.cvut.cz

Abstract We pose robust matching with parametric and non-parametric constraints as the problem of finding a stable independent set (SIS) in an oriented graph whose vertices are all possible correspondences, whose edges capture the structure of the constraints and whose edge orientation represents pairwise comparison ‘is better’ based on correspondence quality, including the uncertainty of this comparison. We show SIS possess properties of both robustness and weak optimality. The main contribution of this paper is algorithmic speedup that results from exploiting the dependence between the standard uniqueness constraint and the parametric constraint. The general theory is demonstrated on the example of image stitching using homography model. The algorithm needs at most $k N^2$ calls of a procedure testing if two ellipse correspondences are consistent with a general homography. The previous known SIS algorithm needed $O(N^4)$ tests. Experiments show the method gives good results and is fast in practice with $k \approx 0.3$.

1. Introduction

In this paper we are interested in image feature matching for problems like the one in Fig. 1. Such procedure must be robust: along with finding a matching, it must discover which subsets of image features are consistent with a global (geometric) model. The usual approach to robustness is based either on residual trimming, the most advanced version of which is the M-estimator [6], or on sampling like RANSAC [5] or LMS [12], or on explicit outlier rejection based on the EM-algorithm [17].

For solving many optimal problems in computer vision, energy optimization has become a widely popular approach. But robustness is difficult to reconcile with optimality without breaking the polynomial solvability barrier. In this paper we propose replacing optimality with stability to achieve robustness at low-order polynomial computational cost. To this end we borrow the game-theoretic formalization of stability [18]. Besides simplicity and computational efficiency, the advantage of the proposed method is its flex-

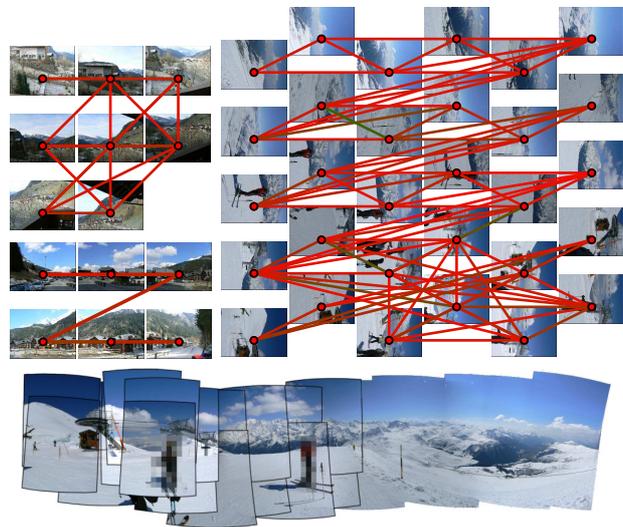


Fig. 1. Three panoramas identified among the set of 44 images (top) and the largest panorama stitched together, shown gradually from registered to blended images (bottom).

ibility that allows ‘robustification’ for a large class of problems, eg. other energy-based methods.

Let us begin by discussing stability of optimal solutions of a simple bipartite matching problem. Since we anticipate incomplete matchings (due to the required robustness), then, for technical reasons, we must consider maximum-cost problems instead of minimum-cost problems. In principle, minimum-energy problems can be transformed to max-cost problems by subtracting the energy values from a suitable fixed constant.

Consider a 4×4 maximum-cost bipartite matching problem which is given by a matching table shown in Fig. 2(a). The value $c(i, j)$ is the cost of including the pair (i, j) in the matching, e.g. $c(1, 4) = 0.0$. The matching problem is

$$M^* = \arg \max_{M \in \mathcal{M}} \sum_{(i,j) \in M} c(i, j), \quad (1)$$

where \mathcal{M} is the set of all possible bipartite matchings.

0.1	0.7	0.8	0.0
1.8	1.7	1.1	0.7
0.9	1.7	1.85	0.5
1.0	0.8	1.8	0.95

(a) M_1^*

0.1	0.7	0.8	0.0
1.8	1.7	1.1	0.7
0.9	1.7	1.85	0.5
1.0	0.8	1.8	0.95

(b) M_2^*

0.1	0.7	0.8	0.0
1.8	1.7	1.1	0.7
0.9	1.7	1.85	0.5
1.0	0.8	1.8	0.95

(c) M_3^*

Fig. 2. Max-cost solution M_1^* to a bipartite matching problem is not absorbant (a) because of the red unmatched element (3, 3). An absorbant solution M_2^* to the same problem with the same total cost of 5.3 and absorbance margin of 0.05 (b). Absorbant solution M_3^* of margin 0.1 is no longer complete (c).

The undershaded entries in Fig. 2(a) and 2(b) show two possible solutions to the problem, M_1^* and M_2^* . The cost of both matchings is 5.3. There is an interesting difference between them: the M_1^* is unstable: a small perturbation of the costs (like changing the cost $c(3, 3)$ by arbitrarily small $\varepsilon > 0$) causes the optimal solution M_1^* change to M_2^* .

The basis for stability is *absorbance* [2]. Let $U(a)$, $a = (i, j)$ be the set of matching table entries that contains the i -th row and the j -th column of the table, without the element (i, j) itself. Let $U(M)$ be the union $\bigcup_{a \in M} U(a)$. We say a matching M is absorbant if every element r in $U(M)$ has at least one element $a \in M \cap U(r)$ such that a is of strictly greater cost than r . We see the M_1^* is not absorbant because of $c(3, 3) = 1.85$. On the other hand, the M_2^* is absorbant, hence stable: Even if any combination of the unmatched elements increased their costs by 0.05, the matching M_2^* would still be absorbant. We call this maximum allowed increase an *absorbance margin*.

Stable solutions need not always be complete (of cardinality N , where N is the number of matching participants, $N = 4$ in Fig. 2): If the absorbance margin is required to be 0.1 then the stable matching is just $M_3^* = \{(2, 1)\}$ of cost 1.8, see Fig. 2(c). Note M_3^* does not contain the maximum-cost element (3, 3).

Interestingly, there is a simple polynomial algorithm that can find a stable solution given an absorbance margin as a parameter. This algorithm has been described in [13, 14, 16] and its new fast variant will be described in Sec. 4. It can be shown there is at most one maximal stable solution, see Sec. 3. This means that if there is an optimal max-cost solution that is stable, then the algorithm finds it, otherwise it finds a stable solution that approximates the optimal solution and the approximation factor can be characterized, see Sec. 3. But this is expected, since robustness must obviously come at the price of loosing max-cost optimality.

The stable set problem is constructed by transforming the problem at hand to an interval-oriented digraph in which putative solutions are independent vertex sets. The structure of the graph encodes the structure of the constraints (eg. matching uniqueness) and its orientation is constructed from confidence intervals encoding local evidence (eg. cor-

respondence quality) that are associated with graph vertices.

The point is that unstable solutions exist in other problems, not just in bipartite matching. They pertain to all min-cost and max-cost problems. We argue that stability is a prerequisite to robustness: to some extent data perturbation must have little influence on the solution. The proposed solution to robust matching is based on a generalization of the notion of a stable independent vertex set (SIS) in a problem graph, which has been studied since von Neumann [18]. An extensive review of the state of the art from the algorithmic point of view can be found in [15].

The stability principle has been applied to semi-dense stereo [7] and range image registration [16]. In this paper we discuss robust matching problems where the parametric model (homography) and the uniqueness constraint are not independent. This brings speedup from $O(N^4)$ worst-case complexity of algorithm described in [16] to $O(N^2)$, where N is the number of image features to be matched.

As a motivating application we have selected image stitching based on general homography estimation, since this problem is difficult enough to expose all properties of the proposed solution.

2. Parametric Matching Problem

In this paper we are interested in an exact polynomial algorithm for solving the following robust parametric matching problem.

Let X^1 be the set of geometric entities in the reference image and X^2 be the set of geometric entities in the target image. They can be image points, elliptic neighborhoods, LAFs [9], etc. or their (fixed) pairs, triples, etc. Let P^1 be a set indexing into X^1 and P^2 be a set indexing into X^2 . For the sake of clarity, we call the set $P^1 \cup P^2$ *participants* in the matching problem. Let further $x_i^1 \in X^1$ and $x_j^2 \in X^2$ for all $i \in P^1$, $j \in P^2$. Then the matching problem is:

1. Finding the largest subsets $S^1 \subseteq P^1$, $S^2 \subseteq P^2$, and
2. finding a bijective mapping $M: S^1 \rightarrow S^2$, and
3. a parameter set Θ , such that
4. every matched pair $p = (i, M(i))$, $i \in S^1$ satisfies a parametric constraint

$$C(x_i^1, x_{M(i)}^2 \mid \Theta) = 0, \quad (2)$$

5. and M is stable and as large as possible.

We shall see this is sufficient for a unique solution for M . It is assumed that there is at most one set of (true) parameters Θ that can explain the relation between the images. For the sake of proofs we assume the constraint (2) can be tested without any uncertainty (there is no measurement error). This will be relaxed in Sec. 4. We also assume that given Θ and a correspondence $(i, M(i))$ satisfying (2), no other participant $j \in P^2$, $j \neq M(i)$ satisfies $C(x_i^1, x_j^2 \mid \Theta) = 0$. In

other words, given x_i^1 and Θ , the set $Y(x_i^1, \Theta)$ of geometric entities y satisfying $C(x_i^1, y | \Theta) = 0$ must be a singleton.

For simplicity we assume two correspondences overconstrain the unknown parameters Θ , ie. they suffice for deriving a necessary condition from the model C that constrains all other correspondences.

As an example, if $\mathbf{x}^1, \mathbf{x}^2$ are corresponding points in image plane and $\mathbf{H}(\mathbf{t}, s)$ is a homography restricted to translation \mathbf{t} and scale change s , the constraint C can be

$$C(\mathbf{x}^1, \mathbf{x}^2 | \mathbf{t}, s) = \mathbf{x}^2 - s(\mathbf{x}^1 - \mathbf{t}) = \mathbf{0}. \quad (3)$$

This model has three parameters and gives two equations per correspondence. Two point correspondences suffice to test (3) even if we do not know \mathbf{t} and s . If there are $O(N)$ pre-selected tentative correspondences, a RANSAC solution to this matching problem requires a repeated draw of pairs from the $O(N)$ correspondences and looking for the largest consensus set by testing $O(N)$ remaining correspondences. Our problem solution does not require any tentative correspondences, if there are N participants, a graph is constructed whose vertices are all possible of the $O(N^2)$ single correspondences. Two correspondences are connected by an edge if they violate uniqueness or the parametric constraint (3). Putative matchings are independent vertex sets in this graph. The $O(N^4)$ edges of this graph will never be represented explicitly. We shall see that only $O(N^2)$ edges due to parametric incompatibility need to be traversed.

The question of this paper is how to find the stable independent vertex set in this graph efficiently. Note we require matching M to be a bijection, i.e. we have a uniqueness constraint which requires every participant be matched at most once. This makes the uniqueness and the parametric constraints dependent. We will make use of this dependency to design an efficient algorithm.

3. Stable Independent Sets and Their Robustness and Optimality

In this section we define stability precisely and then review major results on robustness and optimality of stable solutions. More details and proofs are found in [15].

Recall we have a matching table whose entries are image similarities $c(p)$ for each possible correspondence $p = (i, j)$, where larger $c(p)$ is better. Instead of single-valued similarities we will use confidence intervals $e(p) = [\underline{e}(p), \bar{e}(p)]$. We will assume the top value $\bar{e}(p)$ is a measure of match quality. The width of the interval $\bar{e}(p) - \underline{e}(p)$ corresponds to our uncertainty about the reliability of the similarity estimate. In this paper we assume that the ‘true’ similarity value (without noise or image discretization artifacts) falls within this interval. The case when it does not is briefly discussed in Sec. 7.

We say two correspondences $p = (i, j)$, $q = (k, l)$ are *mutually exclusive* if the uniqueness constraint is violated,

ie. when $q \in U(p)$, (see Sec. 1 for the definition of U) or if they violate the parametric constraint C defined by (2). We construct an unoriented graph $\mathcal{G} = (V, E)$ whose vertices V are all possible correspondences and whose edges E are all mutually exclusive correspondence pairs. We call this the *underlying graph*. Hence, the structure of this graph is given by the constraints of the matching problem. An *independent set* $K \subseteq V$ of \mathcal{G} is one in which no two elements $p, q \in K$ are joined by an edge.

Edges of the underlying graph are now oriented as follows: the edge $\{p, q\} \in E$ is oriented from p to q if $\bar{e}(p) < \underline{e}(q)$, ie. when match quality strictly favors q to p . We denote the set of unidirectionally oriented edges as \vec{A} . They are ordered pairs: $(p, q) \in \vec{A} \Rightarrow (q, p) \notin \vec{A}$. Remaining edges are oriented bidirectionally, this corresponds to the case when confidence intervals overlap, $e(p) \cap e(q) \neq \emptyset$, and we cannot compare p, q at the given confidence level. We denote the set of bidirectional edges as \overleftrightarrow{A} and we have $(p, q) \in \overleftrightarrow{A} \Leftrightarrow (q, p) \in \overleftrightarrow{A}$. We write $\vec{A} \cup \overleftrightarrow{A} = A$, where $\vec{A} \cap \overleftrightarrow{A} = \emptyset$. We will call the oriented edges unidirectional/bidirectional arcs, respectively, in an oriented graph $\mathcal{D} = (V, \vec{A} \cup \overleftrightarrow{A})$ (or $\mathcal{D} = (V, A)$ when we do not distinguish arc types). We will call this the *problem graph*.

Orientations induced by real intervals have special properties that make the problem of finding a stable set of polynomial complexity [15]. Moreover, they belong to the class of orientations that guarantee a unique maximal stable independent set (SIS) [15] according to the following definition:

Def. 1. *Let M be an independent vertex set in $\mathcal{D} = (V, A)$. We say M is stable if for every $p \in M$ there is a $r \in M$ such that $(q, r) \in \vec{A}$ for every $(p, q) \in A$.*

This is indeed equivalent to the informal stability definition discussed above. Consider the example in Fig. 2(b). We have 16 vertices in V corresponding to the 4×4 table. Every vertex $p = (i, j)$ is associated with interval $[c(p) - 0.05, c(p)]$. Since the only constraint is uniqueness, edges E of the underlying graph occur only between vertices sharing a row or a column of the table (each row and each column corresponds to a clique in the graph, the graph is too dense to be drawn here). It is easy to verify the set M_2^* is stable: we take every $p \in M_2^*$ in turn and test if every element $q \in U(p)$ such that $\bar{e}(q) \geq \underline{e}(p)$ has an element $r \in U(q) \cap M_2^*$ such that $\underline{e}(r) > \bar{e}(q)$. Indeed, it has.

If the SIS is to be robust, it has to exhibit reasonable behavior with respect to uncertainty encoded by the bidirectional arcs \overleftrightarrow{A} . Two propositions summarize this behavior (in interval orientations of graphs):

Proposition 1 (Stability wrt uncertainty increase [15]). *Let $\mathcal{D} = (V, \vec{A} \cup \overleftrightarrow{A})$ have no SIS. We construct $\mathcal{D}' = (V, \vec{B} \cup \overleftrightarrow{B})$ from \mathcal{D} by making a subset $A_0 \subseteq \overleftrightarrow{A}$ bidirectional: $\vec{B} = \vec{A} \setminus A_0$, $\overleftrightarrow{B} = \overleftrightarrow{A} \cup A_0$. The \mathcal{D}' has a SIS for no choice of A_0 .*

Proposition 2 (Stability wrt uncertainty reduction [15]). *Let M be a SIS in $\mathcal{D} = (V, \vec{A} \cup \vec{A})$, let $\mathcal{D}' = (V, \vec{B} \cup \vec{B})$ be constructed from \mathcal{D} by making an arbitrary subset $A_0 \subseteq \vec{A}$ unidirectional, $\vec{B} = \vec{A} \cup A_0$, $\vec{B} = \vec{A} \setminus A_0$ and let M' be the maximal SIS in \mathcal{D}' . Then M is a SIS in \mathcal{D}' and $M \subseteq M'$.*

This implies that a SIS for a given set of intervals $e(p)$, $p \in V$ is the intersection of all solutions for any other choice of intervals $e'(p)$ as long as $e'(p) \subseteq e(p)$ for all $p \in V$ [14].

Before we turn our attention to optimality, we need to discuss equivalent reductions of the problem graph. We say two oriented graphs \mathcal{D} and \mathcal{D}' are S -equivalent if they have the same set of stable independent sets. A reduction rule is a procedure that simplifies \mathcal{D} to a smaller S -equivalent graph \mathcal{D}' . The simplest reduction rule is based on the notion of a *sink*. We say a vertex $p \in V$ is a sink in $\mathcal{D} = (V, A)$ if there is no outgoing arc from p in \mathcal{D} , ie. if there is no element $q \in V$ such that $(p, q) \in A$. An isolated vertex is a sink.

Reduction Rule 1 ([18]). *Let $s \in V$ be a sink in $\mathcal{D} = (V, A)$, $N(s)$ be the set of neighbors of s in \mathcal{D} . Construct \mathcal{D}' from \mathcal{D} by removing $N(s)$. Then \mathcal{D}' and \mathcal{D} are S -equivalent.*

Let us now return to optimality. Robustness should not destroy optimality properties completely. The following two propositions characterize the remaining weak optimality of SIS. We say $q \in V$ is a predecessor of $p \in V$ in $\mathcal{D} = (V, A)$ if $(q, p) \in A$. The set of all predecessors of p will be denoted as $P(p)$. Let $Q(p)$ be an independent¹ subset of $P(p)$ and $\mathcal{Q}(p)$ be the set of all such subsets $Q(p)$.

Proposition 3 (Weak Optimality I [14]). *Let \mathcal{D} be an interval-oriented graph and let M be a SIS which is also a maximal independent vertex set of \mathcal{D} . If there is a sequence of reductions by Rule 1 such that each sink s at which reduction occurs satisfies*

$$\underline{e}(p) \geq \max_{Q \in \mathcal{Q}(p)} \sum_{q \in Q} \bar{e}(q) \stackrel{\text{def}}{=} \phi(p) \quad (4)$$

then M also maximizes the cost sum over all possible independent vertex subsets \mathcal{M} of \mathcal{D} , ie.

$$M = \arg \max_{K \in \mathcal{M}} \sum_{p \in K} \bar{e}(p). \quad (5)$$

This is a sufficient condition. If a SIS is a maximal independent vertex set, and if we can find a reduction sequence satisfying (4) then we know the SIS is a max-sum independent vertex set, ie. an optimal solution. Clearly, greater discriminability of image features improves similarities e , which increases the chance of happening this. In the example in Fig. 2(b), the sets \mathcal{Q} are always of cardinality at most two

and the $\phi(p)$, $p = (i, j)$ is computed by taking the maximum of $\bar{e}(q_1) + \bar{e}(q_2)$ over all pairs of remaining elements $q_1, q_2 \in U(p)$ such that $q_1 \notin U(q_2)$ (or just $\bar{e}(q_1)$ if the other does not exist, or $\phi(p) = 0$ if neither exists). Eg. for the sinks $s_1 = (3, 3)$ and $s_2 = (2, 1)$ in Fig. 2(b) we have $\phi(s_1) = 1.7 + 1.8 = 3.5$ and $\phi(s_2) = 2.7$. We see the (4) is too strict: SIS are optimal more often than the condition tells us. In fact, given a SIS M , we can compute the upper bound on the optimal solution cost as

$$\sum_{s \in M} c(s) + \max(0, \phi(s) - c(s)),$$

giving 7.85 in the example in Fig. 2(b).

The converse proposition uses a much weaker condition: It requires *each vertex* of V to satisfy (4):

Proposition 4 (Weak Optimality II [15]). *Let \mathcal{D} be an interval-oriented graph, M be a solution to (5) and let each vertex $p \in V$ satisfy (4). Then M is a maximum SIS in \mathcal{D} .*

Often, optimal solution is a SIS even if it violates this condition. A tighter bound has not been found.

4. Algorithm

Reduction Rule 1 can be used as an algorithm because of the following proposition:

Proposition 5 ([14]). *Let \mathcal{D} be an interval-oriented graph with no sink. Then \mathcal{D} has no SIS.*

See [14] for a detailed description of an algorithm based on Rule 1. Note that vertices of the largest similarity $\bar{e}(p)$ tend to be sinks but only if their similarity differ enough from their neighbors. If not, they are not sinks and hence need not be members of the solution, as, for example in Fig. 2(c). More precisely, they could become part of the solution only after successive reductions of the graph removed all their successors but this cannot happen in interval orientations where unidirectional arcs \vec{A} represent a partial order.

The complexity of the algorithm based on Rule 1 is $O(N^4)$, which is slow. Testing if a vertex is a sink requires $O(|V|) = O(N^2)$ operations, since a typical problem graph is quite dense. A notable exception to this complexity is the case when edges in \mathcal{D} represent a bipartite matching problem (they are constructed from the uniqueness constraint only), as in the example in Fig. 2(a). In such case finding *all sinks* (which we will call U -sinks shortly) can be done in $O(N^2)$ steps as shown in Alg. 1.

It is easy to see all sinks of a graph \mathcal{D} constitute an independent vertex set. No independent set is larger than N in a graph of a bipartite matching problem. Hence, we need to make at most N passes to fully reduce the initial graph. Full reduction is achieved when there is no more change in the graph. By Proposition 5, isolated vertices in

¹An independent vertex subset in subgraph induced by $P(p)$ in \mathcal{D} .

Algorithm 1 Finding all U -sinks

- 1: Using the value of \bar{e} , find the first and the second largest elements $c_1(i)$, $c_2(i)$ for each matching table row i and $r_1(j)$, $r_2(j)$ for each column j . This gives vertices $p_1(i) = (i, c_1(i))$, $p_2(i) = (i, c_2(i))$ for all $i \in P^1$ and $q_1(j) = (r_1(j), j)$, $q_2(j) = (r_2(j), j)$ for all $j \in P^2$.
 - 2: U -sinks are all pairs $p_1(i)$ that satisfy $\underline{e}(p_1(i)) > \bar{e}(p_2(i))$ and $\underline{e}(q_1(j)) > \bar{e}(q_2(j))$, where $i = r_1(c_1(i))$.
-

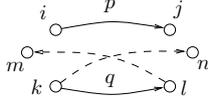


Fig. 3. Illustration to the proof of Lemma 1.

the fully reduced interval-oriented graph constitute the maximal (largest possible) SIS. This gives $O(N^3)$ worst-case complexity of the reduction algorithm in bipartite matching problems. Assuming absorbance margin of 0.05, the problem in Fig. 2(b) needs two passes: The first pass gives sinks (2, 1) and (3, 3), the second pass gives (1, 2) and (4, 4).

If we employ the parametric constraint (2), we no longer have a bipartite matching problem graph and finding sinks becomes more difficult. Given an oriented problem graph $\mathcal{D} = (V, \vec{A} \cup \vec{A})$ we consider its two subgraphs: The U -graph $\mathcal{U} = (V, \vec{A}_U \cup \vec{A}_U)$ contains only edges due to uniqueness constraint violations (as in bipartite matching) and the C -graph $\mathcal{C} = (V, \vec{A}_C \cup \vec{A}_C)$ contains only edges due to violations of pairwise parametric constraint (2), as explained above. Denoting $U(p)$ the set of neighbors of p in U -graph and $C(p)$ the set of its neighbors in C -graph, we see that $U(p) \cup C(p) = N(p)$.

We first observe that every sink in \mathcal{D} must be a sink in \mathcal{U} , which we call a U -sink. So Alg. 1 gives us $O(N)$ (ie. a small number) candidate sinks that only need to be tested over the edges of the graph \mathcal{C} . Since \mathcal{C} is not represented explicitly, this requires $O(N^2)$ calls of a procedure testing the constraint (2) to find all C -neighbors of a U -sink. In total, this requires $O(N^3)$ calls for the whole matching problem. We shall now see how to reduce this complexity to $O(N^2)$ calls. The constraint-testing procedure will be called a *geometric test*: $\text{gtest}(p, q)$ returns `true` if $q \in C(p)$.

Lemma 1. *Let p, q be not C -neighbors, $q \notin C(p)$. Then every $s \in U(q)$ is a C -neighbor of p and every $t \in U(p)$ is a C -neighbor of q .*

Proof. Consider Fig. 3 showing the geometric mapping among the involved participants. If there is no error and two correspondences p, q constrain the parametric model uniquely then both $s_1 = (k, n)$ and $s_2 = (m, l)$ must fail the geometric test. \square

Lemma 1 is useful for proving another reduction rule. If $T \subseteq V$ is a vertex subset, we denote $U(T) = \bigcup_{t \in T} U(t)$.

Reduction Rule 2. *Let $s \in V$ be a sink in $\mathcal{D} = (V, A)$. Let $T = V \setminus N(s)$ be the set of vertices of \mathcal{D} that are not neighbors of s . Construct \mathcal{D}' from \mathcal{D} by removing $N(s) \cup U(T)$. Then \mathcal{D} and \mathcal{D}' are S -equivalent.*

This rule is very potent: After reduction by Rule 2, the resulting graph becomes very small:

Lemma 2. *Let $\mathcal{D}' = (V', A')$ be obtained from $\mathcal{D} = (V, A)$ by Rule 2. Then $|V'| \leq \min(|P^1|, |P^2|)$, where $P^1 \cup P^2$ is the participant set in \mathcal{D} .*

Proof. Let $s \in V$ be a sink in \mathcal{D} . We first show by contradiction that no two vertices in \mathcal{D}' are U -neighbors, ie. for every $p, q \in V'$ it holds that $p \notin U(q)$. For suppose there are $p, q \in V'$ that are U -neighbors. Then $p \notin N(s)$ after the reduction. If $p \notin C(s)$ then $q \in C(s)$ by Lemma 1. Hence $q \notin V'$ which is a contradiction. Next, since no two vertices in \mathcal{D}' are U -neighbors, the set V' is a one-to-one matching in \mathcal{D}' , so $|V'| \leq \min(|P^1|, |P^2|)$. \square

Hence, the most expensive part of an SIS algorithm seems finding the first sink. We show this can be done fast. We say two vertices $p, q \in V$ are geometrically compatible if $p \notin C(q)$. Geometric compatibility is equivalence, hence it is symmetric and transitive. This gives two results:

Lemma 3. *Let u have the largest $\underline{e}(u)$ over all U -sinks. If there is a $p \in C(u)$ of $\bar{e}(p) \geq \underline{e}(u)$ then \mathcal{D} has no SIS.*

Proof. Clearly u is not a sink in \mathcal{D} . We show that no other U -sink $v \neq u$ is a sink: If $v \in C(u)$ then $\bar{e}(u) \geq \underline{e}(u) \geq \underline{e}(v)$. If $v \notin C(u)$ then $v \in C(p)$ by transitivity of geometric compatibility and we have $\bar{e}(p) \geq \underline{e}(u) \geq \underline{e}(v)$. \square

Let Q be the set of all vertices $\{p_1(i), q_1(j); i \in P_1, j \in P_2\}$ from Alg. 1. Let $s(q) = \max_{p \in U(q)} \bar{e}(p)$ be the cost of the second-largest element in $U(q)$ for every $q \in Q$, this is a by-product of Alg. 1 at no additional computational cost. The second result then states:

Lemma 4. *Let u be a U -sink of maximal $\underline{e}(u)$. If there is a $q \in Q \setminus C(u)$ such that $s(q) \geq \underline{e}(u)$ then \mathcal{D} has no SIS.*

Proof. By Lemma 1, $t \in C(q)$ for every $t \in U(q)$. Transitivity of compatibility and $q \notin C(u)$ require $t \in C(u)$. There is always a t such that $\bar{e}(t) = s(q)$. Hence $\bar{e}(t) \geq \underline{e}(u)$ and Lemma 3 applies. \square

Since every vertex $p \in V$ is either in Q or in $U(q)$ of some $q \in Q$ this means that if p passes the tests of Lemma 3 and 4 with Q as the test set, it must be a sink.

Alg. 2 gives the pseudo-code based on the results in this section. The reduction in the number of necessary geometric tests results in low algorithmic complexity, as opposed to the algorithm based on Rule 1:

Lemma 5. *The worst-case complexity of Alg. 2 is $O(N^2)$.*

Algorithm 2 Equivalent Reduction Algorithm

Require: Interval-oriented graph $\mathcal{H} = (V, A)$.

- 1: Initialize $M := \emptyset$.
 - 2: **loop**
 - 3: Run Alg. 1 to obtain U -sinks U , the set Q and the values $s(q)$ for every $q \in Q$.
 - 4: **if** $U = \emptyset$ **then return** M **end if**
 - 5: Find u of maximal $\underline{e}(u)$ in U .
 - 6: **if** there is $q \in Q$ such that $\text{gtest}(q, u)$ **and** $\bar{e}(q) \geq \underline{e}(u)$ **or** $s(q) \geq \underline{e}(u)$ **then break end if**
 - 7: **for** every $q \in V, q \neq u$ **do** {Rule 2 at U -sink u }
 - 8: **if** $\text{gtest}(q, u)$ **then** $D := \{q\}$ **else** $D := U(q)$ **end if**
 - 9: Delete D from \mathcal{H} .
 - 10: **end for**
 - 11: Update $M := M \cup \{u\}$, delete $U(u) \cup \{u\}$ from \mathcal{H} .
 - 12: **end loop**
-

Proof. If there is no SIS then there is either no U -sink which can be confirmed in $O(N^2)$ time in Step 3 or the U -sink of maximal $\underline{e}(u)$ is not a sink which is confirmed in $O(N)$ time since the size of Q is $O(N)$. After the first sink is found, we need $O(N^2)$ passes for the reduction in Steps 7–10 where the worst case happens when the geometric tests in every $q \in V$ succeeds only after all prior tests in $U(q)$ have failed. This is rare, as demonstrated by the experiments.

After the first sink is found, the problem graph is reduced to $O(N)$ vertices. Hence, there are $O(N)$ remaining passes through the main loop, each pass taking $O(N)$ steps. \square

The reader may have noticed that it should not be necessary to search for U -sinks in Step 3 after the first pass of the main loop (see Lemma 2). This is true as long as the geometric test can be done exactly. In practice it is often performed by testing if residual error exceeds a threshold. This may fail if there are two participants in the image that are closer to each other than the threshold. As a result, uniqueness and geometric compatibility are no longer dependent: It may happen that $q \notin C(p)$ even if $q \in U(p)$. A change is then needed in Step 8: Instead of deleting $U(q)$ we delete only $U(q) \cap C(q)$. It is easy to see that no change is necessary in Step 6 where gtest is also used. With this fix, the worst-case complexity of Alg. 2 should break the $O(N^2)$ barrier but we never observed this happen.

5. Application To Image Stitching

To demonstrate the theory introduced above and show the behavior of SIS in a practical setup, we describe the problem of finding a homography between two images given a sparse set of features. This allows us to describe the implementation technicalities in more detail.

We have selected ellipses oriented to a local affine frame (LAF) [9] as the image features. The ellipses were obtained by a fit to MSER regions [8]. Two ellipse correspondences

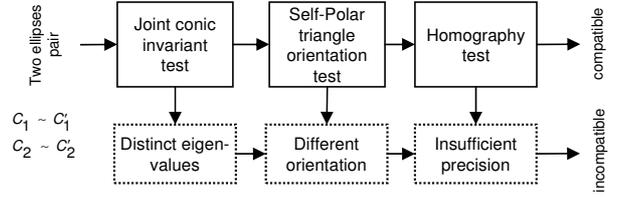


Fig. 4. The structure of the geometric test.

overconstrain a homography which gives a necessary condition for geometric compatibility. The LAF is used only to compute image similarity.

The geometric test is based on several necessary conditions that are progressively more expensive, see Fig. 4. The first subtest is based on joint conic invariants [10]: Given a pair of conics $\mathbf{C}_1 \leftrightarrow \mathbf{C}'_1, \mathbf{C}_2 \leftrightarrow \mathbf{C}'_2$, corresponding under homography, the matrix $\mathbf{C}'_2^{-1}\mathbf{C}_1$ has the same eigenvalues as $\mathbf{C}'_2^{-1}\mathbf{C}'_1$. Since the matrix $\mathbf{C}'_2^{-1}\mathbf{C}_1$ is homogeneous, we sort its eigenvalues $\lambda_1 \leq \lambda_2 \leq \lambda_3$ and the corresponding eigenvalues $\lambda'_1 \leq \lambda'_2 \leq \lambda'_3$ of $\mathbf{C}'_2^{-1}\mathbf{C}'_1$ and test if the ratio

$$r = 1 - \frac{2\lambda_1\lambda'_1}{\lambda_1\lambda'_3 + \lambda'_1\lambda_3} \cdot \frac{2\lambda_2\lambda'_2}{\lambda_2\lambda'_3 + \lambda'_2\lambda_3} \quad (6)$$

exceeds a given threshold τ_r . This test is necessary but insufficient. The second subtest is based on the eigenvectors of $\mathbf{C}'_2^{-1}\mathbf{C}_1$ corresponding to three vertices of the self-polar triangle [10]: If a correspondence is correct the triangle must preserve its orientation with respect to the vertex order given by the eigenvalues. The last subtest is based on homography \mathbf{H} estimated from the four bitangents of the ellipses $\mathbf{C}_1, \mathbf{C}_2$ and the corresponding bitangents between $\mathbf{C}'_1, \mathbf{C}'_2$ [11]. The bitangents are ordered by the point of tangency wrt the first ellipse. We test if the symmetric transfer error

$$e = \sum_{i=1}^3 \|\mathbf{t}'_i - \mathbf{H}\mathbf{t}_i\|^2 + \|\mathbf{t}_i - \mathbf{H}^{-1}\mathbf{t}'_i\|^2 \quad (7)$$

in pixels between the vertices of the transferred self-polar triangle $\mathbf{H}\mathbf{t}_i$ and their position in the second image \mathbf{t}'_i exceeds a threshold τ_e .

Efficiency of the compound test was examined as follows: Given 3000 ellipses in the reference image and 3500 in the target image we fixed one true match and performed the geometric test with respect to all other correspondences. We used $\tau_e = 2, \tau_r = 0.2$. The joint conic invariant test rejected 99.5% of all correspondences, the orientation test received the remaining 0.5% and rejected 43% of them and the homography test received 0.3% of all correspondences and rejected 88% of them.

With the help of LAF associated with the ellipse, image similarity is measured by normalized cross-correlation ρ so that $\bar{e}(p) = \rho$ and $\underline{e}(p) = \rho - \max(\alpha|\rho|, \beta)$, where we used



Fig. 5. Examples from the test suite of 900 images.

$\alpha = 0.01$, $\beta = 0.01$. The α makes the interval width inversely proportional to local image contrast and β is a fixed margin that helps identify repetitions in the image.

When there is more than a single homography that can explain the relation between the images, the SIS solution is empty. This happens because the highly correlated correspondences are mutually geometrically incompatible. As a result, the oriented problem graph has no sink. This is a correct behavior if there are repeated structures in the image. On the other hand, one also gets an empty solution in the case when there is even a single wrong correspondence of high correlation and narrow confidence interval. To cope with the latter case, we search for multiple models (homographies) as follows:

Algorithm 3 Finding The Largest-Support Solution

- 1: $i := 1$.
 - 2: **repeat**
 - 3: Run Alg. 2 to find a SIS M_i , $i := i + 1$.
 - 4: Remove $M_i \cup U(M_i)$ from the matching problem.
 - 5: **until** $M_i = \emptyset$
 - 6: Let M_j be of maximal size among all M_i 's found.
 - 7: **if** $|M_j| \leq \tau_M$ **then** $M_j := \emptyset$ **end if**
 - 8: **return** M_j
-

Making this procedure clean, in the spirit of the above theory, is a topic for ongoing work. Most often, the first set M_1 is correct. This is discussed in the next section.

6. Experiments

The first experiment is a performance study on a synthetically generated ground-truth dataset. Besides accuracy and speed we study two additional questions: (1) How often is the first set M_1 obtained by Alg. 3 the correct one? (2) What is the number of all geometric tests in Alg. 2? The second experiment uses real images to compare Alg. 2 to locally optimized [4] 2-ellipse [11] RANSAC we call LO-2-RANSAC. The last experiment shows an example of complex panorama stitching. Detailed results are found in [3].

The test suite consists of 900 images generated from the same underlying image by exhausting all combinations of rotations per 45° , scales $1/4$, $1/2$, 1 , $3/2$, 2 , skews 0° and $\pm 45^\circ$ and 9 different elations. The image is of a complex scene including strong repetitive structures. A few examples from the test suite are shown in Fig. 5.

Histograms in Fig 6 left show percentages of cases in which the i -th set M_i was the correct one, evaluated on the

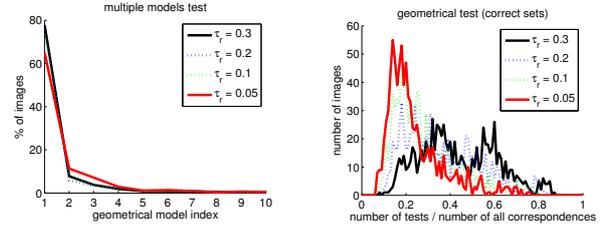


Fig. 6. Percentage of cases in which the i -th set M_i was the correct (left). Histograms of the number of geometric tests (right).

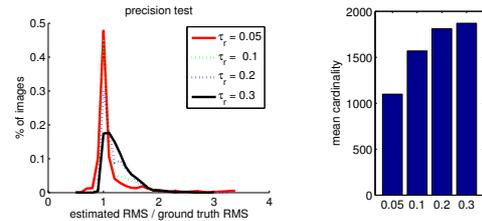


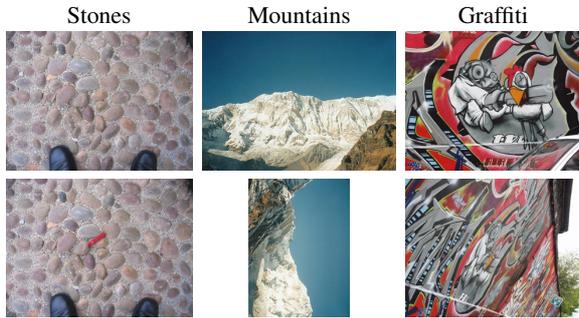
Fig. 7. RMS error relative to ground-truth error (left) and mean cardinality of the matching as a function of τ_r (right).

test suite. In over 60% of all cases the first matching is correct. Surprisingly, this success increases with relaxing the threshold τ_r . We observed in this test that every incorrect matching had fewer than 7 correspondences, which determined our threshold $\tau_M = 7$ in Alg. 3.

Histograms in Fig. 6 right show the distribution of the number of geometric tests per image pair relative to all N^2 correspondences measured in max-support matchings from Alg. 3. In none of the cases the N^2 bound was exceeded. On average $0.3N^2$ tests was needed. We can see the number of geometric tests is smaller with a tighter threshold τ_r .

Finally, Fig. 7 shows matching accuracy using the homography estimated by a non-linear LS fit from the matching. It is measured as the ratio of residual RMS error for the homography estimated by our algorithm and the RMS error for the ground truth homography. This is necessary since the MSER detector run in each image independently with finite resolution which made the ground-truth error non-zero. In some cases the estimated homographies have smaller error than the ground truth. It is caused by a small shift of MSER when the images differ in scale. The estimated homography fits this shift better. As expected, tighter τ_r gives more accurate results (at the cost of smaller matching).

The second experiment uses real image pairs shown in Fig. 8. The first pair of narrow-baseline contains many similar objects (small stones). Thus thresholding by correlations give many matches ($3 \cdot 10^5$ at 85% threshold) while selecting mutually best matches give at most 1570 matches. This makes selecting good tentative correspondences more difficult. The second is a standard panoramic pair of small overlap. The last is a wide-baseline stereo pair of a planar wall with extreme projective distortion. Our algorithm is compared to LO-2-RANSAC in Fig. 8. Tentative corre-



Scene	Alg. 2			LO-2-RANSAC		
	$ M $	RMS	time	$ M $	RMS	time
Stones	203	0.44	10.3	214	1.58	1.98
Mountains	14	0.65	4.2	14	1.33	13.8
Graffiti	36	1.81	18.1	34	1.82	10.6

Fig. 8. Real image pairs for the second experiment and the comparison results, $|M|$ is the matching size, time is given in seconds.

spondences for RANSAC are obtained by finding all mutually best matches using cross-correlation ρ , excluding them from the matching table and repeating this procedure $100\times$. We used $\tau_e = 2$, $\tau_r = 0.05$, and $\tau_M = 7$. ML fit was computed from the inliers in each method.

For a fair comparison, the implementations of LO-2-RANSAC and Alg. 2 were in Matlab 7 without using compiled C code. The test run on Mobile AMD Sempron 3000 CPU. Resulting inlier sets are of similar cardinality $|M|$. On average, Alg. 2 achieves about 40% better accuracy and the solution is found in about half the time. We observed RANSAC results are quite sensitive to the selection of tentative correspondences [3].

The last experiment uses our algorithm to identify and stitch panoramas among an unorganized set of 44 images. Each image pair was matched using Alg. 3. All three panoramas were successfully identified as shown in Fig. 1. Matches were used in *Hugin* [1] with *Panorama Tools* to blend images to panoramas. The largest panorama of 30 images is shown at the bottom in Fig. 1.

7. Conclusions

Our approach to robust matching is based on a radically different notion than standard optimization or robust statistics methods. Robustness is directly addressed via the property of stability, while the loss of optimality is characterized. The existence of a unique solution is guaranteed. The resulting algorithm is remarkably simple and is fast, in fact our matching has the same worst-case computational complexity as visiting all N^2 possible correspondences. No tentative correspondences are needed. The algorithm is deterministic which means we have a definite result in guaranteed time. The underlying theory is general and can be used for other problems in computer vision.

Experiments have shown the method behaves well in

practice except when the confidence intervals do not capture the true image similarity or if multiple homography models explain the images. As a result, the first matching need not be correct and a heuristic Alg. 3 for maximal-support search is needed. Finding a well-founded multi-model matching that will replace Alg. 3 is a topic for ongoing work.

Acknowledgments This work has been supported by the EC project MRTN-CT-2004-005439 VISIONTRAIN and by the Czech Academy of Sciences under project IET101210406.

References

- [1] *Hugin*. <http://hugin.sourceforge.net>.
- [2] E. Boros and V. Gurvich. Perfect graphs, kernels, and cores of cooperative games. RUTCOR RR 12-2003, Rutgers Center for Operations Research, 2003.
- [3] M. Bujňák and R. Šára. An application of strict sub-kernels in a sparse correspondence problem. RR CTU–CMP–2007–18, Center for Machine Perception, Czech Technical University, 2007.
- [4] O. Chum, J. Matas, and J. Kittler. Locally optimized RANSAC. In *Proc DAGM*, pp. 236–243, 2003.
- [5] M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm ACM*, 24(6):381–395, 1981.
- [6] P. Huber. *Robust Statistics*. Wiley, New York, 1981.
- [7] J. Kostková and R. Šára. Stratified dense matching for stereopsis in complex scenes. In *Proc BMVC*, pp. 339–348, 2003.
- [8] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *Proc BMVC*, pp. 384–393, 2002.
- [9] J. Matas, Š. Obdržálek, and O. Chum. Local affine frames for wide-baseline stereo. In *Proc ICPR*, pp. 363–366, 2002.
- [10] J. L. Mundy and A. Zisserman (eds). *Geometric Invariance in Computer Vision*, Chap. Geometric Interpretation of Joint Conic Invariants, pp. 77–86. The MIT Press, 1992.
- [11] C. A. Rothwell, A. Zisserman, and al. Relative motion and pose from arbitrary plane curves. *IVC*, 10(4):250–262, 1992.
- [12] P. J. Rousseeuw. Least median of squares regression. *J Am Stat Assoc*, 79(388):871–880, 1984.
- [13] R. Šára. Finding the largest unambiguous component of stereo matching. In *Proc ECCV*, pp. 900–914, 2002.
- [14] R. Šára. Robust correspondence recognition for computer vision. In *Proc COMPSTAT*, pp. 119–131, 2006.
- [15] R. Šára. The principle of stability applied to matching problems in computer vision. RR CTU–CMP–2007–17, Center for Machine Perception, Czech Technical University, 2007.
- [16] R. Šára, I. S. Okatani, and A. Sugimoto. Globally convergent range image registration by graph kernel algorithm. In *Proc 3DIM*, pp. 377–384, 2005.
- [17] P. H. S. Torr and A. Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. *CVIU*, 78(1):138–156, 2000.
- [18] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behaviour*. Princeton University Press, 1944.