

Globally Convergent Range Image Registration by Graph Kernel Algorithm

Radim Šára

Czech Technical University, Czech Republic

sara@cmp.felk.cvut.cz

Ikuko Shimizu Okatani

Tokyo University of Agriculture and Technology, Japan

ikuko@cc.tuat.ac.jp

Akihiro Sugimoto

National Institute of Informatics, Japan

sugimoto@nii.ac.jp

Abstract

Automatic range image registration without any knowledge of the viewpoint requires identification of common regions across different range images and then establishing point correspondences in these regions. We formulate this as a graph-based optimization problem. More specifically, we define a graph in which each vertex represents a putative match of two points, each edge represents binary consistency decision between two matches, and each edge orientation represents match quality from worse to better putative match. Then strict sub-kernel defined in the graph is maximized. The maximum strict sub-kernel algorithm enables us to uniquely determine the largest consistent matching of points. To evaluate the quality of a single match, we employ the histogram of triple products that are generated by all surface normals in a point neighborhood. Our experimental results show the effectiveness of our method for rough range image registration.

1. Introduction

Automatic modeling of 3D objects in the real world is required in many applications such as CG or CAD/CAM because the traditional manual method of constructing models is too labor-intensive to be feasible. A range sensor, which is a sensing device directly measuring 3D information of an object surface, is a useful tool in modeling of 3D objects.

An image of an object captured by a range sensor is called a range image and it provides a partial shape of the object in terms of the 3D coordinates of surface points in which the coordinate system is defined by the position and pose of the range sensor. To obtain the full shape of an object, therefore, we have to align range images captured from different viewpoints.

Finding the rigid transformation that best aligns given range images is called range image registration. One of the widely used methods for range image registration is the it-

erative closest point (ICP) method which was proposed by Besl *et al.* [2]. Basically, it consists of two steps: closest point search and matching similarity maximization, which are iteratively applied to two range images.

The problems that make matching difficult are occlusion and digitization depending on viewpoint. In other words, some points of an object surface observed from one viewpoint are not observed from another viewpoint, and the same points on an object surface cannot be measured in different range images even if they are commonly observed. To overcome these problems, the ICP method has been extended [4, 6, 8, 9, 13, 14, 17, 18, 20, 22]. As a result, the robustness and stability of matching are enhanced.

These methods, however, require a sufficiently good initial estimate to avoid convergence to a local similarity maximum. This is because the error function employed in registration has a number of local maxima, in particular around the true transformation. Methods to obtain such initial estimates are called *coarse registration*.

Over the decades, a number of methods have been proposed to obtain good coarse registrations in which invariant features are detected and matched with rigid transformations [3]. Johnson and Hebert [11] proposed to use a *spin image* as an invariant feature, which is a 2D histogram of surface locations around a point, and to match the spin images based on their cross-correlation. Stein and Medioni [19] showed the usefulness of a *splash* which represents changes of the encoded directions of normal vectors around points of interest, while Chua and Jarvis [5] proposed a *point signature* which encodes the minimum distances of points on a 3D contour from a reference plane. Higuchi *et al.* [10], on the other hand, proposed a *spherical attribute image* which represents on a sphere the normal vectors scaled by curvatures measured at points on an object surface. These features were used as indices representing surface structures in a hash table for matching. A hypothesize and test approach for matching was also proposed in which curvatures and normal vectors [7], curves containing points with zero mean curvatures [12], and the bi-tangent

curves of surfaces [21] are used for matching.

The idea underlying the above conventional approaches is that introducing sufficiently discriminative point features and defining a single metric using the features to find the closest point enable us to find point matches independently, by a maximum-similarity search. This is feasible if one can compute such features in a reliable way. Combining different kinds of features and defining one common metric, however, are still difficult problems.

In this paper, we take a different approach to establishing point matches for coarse registration. That is, we first represent all putative point matches in an graph and then select as many consistent matches among them as possible. This is solved as one global discrete optimization problem, not by point-by-point decisions. Each vertex of the graph represents a putative match of two points, each edge represents binary consistency decision between two matches, and each edge orientation represents match quality from worse to better possible match. Then the maximum strict sub-kernel (SSK) of the graph is found. The maximum strict sub-kernel algorithm [16, 15] enables us to uniquely determine the largest consistent matching of points. This means that our method does not require any initial good estimation and, at the same time, guarantees that global solution is achieved.

2. Graph-Based Matching by SSK Algorithm

In this section, we formalize the matching problem as a discrete optimization problem in an oriented graph and then give an algorithm.

2.1. 3D Point Matching Problem

Consider first a set of surface normals \mathbf{n}_i , $i = 1, \dots, n$ collected over a small surface point neighborhood. We can construct a 3×3 symmetric structure matrix $\mathbf{S} = \sum_{i=1}^n \mathbf{n}_i \mathbf{n}_i^\top$. This matrix is non-singular if $n > 2$ and the vectors \mathbf{n}_i are in general position. If the surface undergoes a rigid motion with rotation matrix \mathbf{R} and translation vector \mathbf{t} , the structure matrix maps to $\mathbf{S}' = \mathbf{R} \mathbf{S} \mathbf{R}^\top$. Using SVD, we can write the relation as

$$\mathbf{S}' = \mathbf{U}' \mathbf{D}' (\mathbf{U}')^\top = \mathbf{R} \mathbf{U} \mathbf{D} \mathbf{U}^\top \mathbf{R}^\top, \quad (1)$$

where $\mathbf{U} \mathbf{D} \mathbf{U}^\top$ is the SVD of matrix \mathbf{S} . Suppose the matrices \mathbf{S} and \mathbf{S}' are constructed from the neighborhoods of two points \mathbf{x} and \mathbf{y} on two different surfaces \mathcal{S}_1 and \mathcal{S}_2 that differ by rigid motion (\mathbf{R}, \mathbf{t}) . If \mathbf{x} and \mathbf{y} are corresponding, two conditions must hold:

$$\mathbf{D}' = \mathbf{D}, \quad (2)$$

$$\mathbf{U}' \mathbf{P} = \mathbf{R} \mathbf{U}, \quad (3)$$

where the 3×3 diagonal matrix $\mathbf{P} = \text{diag}(s_1, s_2, s_1 \cdot s_2)$, $|s_1| = |s_2| = 1$ represents the four-fold ambiguity in signs.¹

The relation (2) requires the *invariant features* to be equivalent and the relation (3) requires the *covariant features* to be geometrically consistent. From the former we can construct a similarity criterion for matching and from the latter we can construct a geometric matching constraint.

Hence, geometric feature correspondence search must achieve two goals: similarity of invariant features and consistency of covariant features. Extending this and saying it in more precise words, if P is the set of all putative point correspondences (\mathbf{x}, \mathbf{y}) from surface \mathcal{S}_1 to surface \mathcal{S}_2 , the correspondence problem requires finding a subset $M \subseteq P$ satisfying four requirements:

- R1 *Similarity*: every correspondence $p = (\mathbf{x}, \mathbf{y}) \in M$ matches points $\mathbf{x} \in \mathcal{S}_1$, $\mathbf{y} \in \mathcal{S}_2$ of similar invariant features.
- R2 *Consistency with search range*: every $p \in M$ is consistent with some admissible rigid motion $(\mathbf{R}_p, \mathbf{t}_p)$.
- R3 *Geometric concordance constraint*: $(\mathbf{R}_p, \mathbf{t}_p) = (\mathbf{R}_q, \mathbf{t}_q)$ for all pairs $p, q \in M$.
- R4 *Uniqueness constraint*: every surface point is matched (occurs in some $p \in M$) at most once.

From this a formal task is constructed as follows. First, unoriented graph $\mathcal{D} = (P, E)$ is created. Its vertex set P represents all putative correspondences $p = (\mathbf{x}_i, \mathbf{y}_j)$, $\mathbf{x}_i \in \mathcal{S}_1$, $\mathbf{y}_j \in \mathcal{S}_2$.² There are two kinds of edges in the set E .

First, let $M \subseteq P$ be some matching. Following R4, an edge joins two vertices p, q if either p can occur in M or if q can occur in M , but they can never occur in M simultaneously. For instance, if point $\mathbf{x}_1 \in \mathcal{S}_1$ is matched to $\mathbf{y}_1 \in \mathcal{S}_2$ then \mathbf{x}_1 cannot be matched to any other point in \mathcal{S}_2 . Hence, if $p = (\mathbf{x}_1, \mathbf{y}_1) \in M$ then no pair of the form $q_i = (\mathbf{x}_i, \mathbf{y}_1)$, $i \neq 1$ or $s_j = (\mathbf{x}_1, \mathbf{y}_j)$, $j \neq 1$ can be in M together with p . Therefore, the p is connected by edges to every such q_i and s_j in \mathcal{D} . These are, for example, the horizontal and the vertical edges in graph \mathcal{D} in Fig. 1(a).

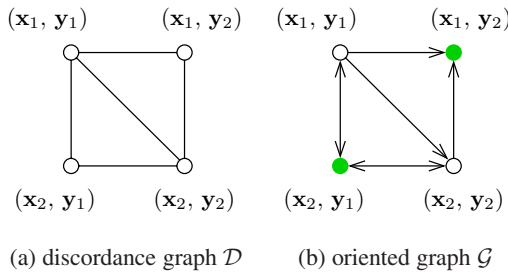
Geometric constraints induce additional edges in \mathcal{D} . Two vertices, $p = (\mathbf{x}_i, \mathbf{y}_j)$ and $q = (\mathbf{x}_k, \mathbf{y}_l)$, are connected by an edge in \mathcal{D} if there is *no* rigid transform (\mathbf{R}, \mathbf{t}) that maps the surface neighborhood of \mathbf{x}_i to \mathbf{y}_j and, simultaneously, the neighborhood of \mathbf{x}_k to \mathbf{y}_l . In other words, if there is no special orthogonal matrix \mathbf{R} such that

$$[\mathbf{y}_l - \mathbf{y}_j, \mathbf{m}_l, \mathbf{m}_j] = \mathbf{R} [\mathbf{x}_k - \mathbf{x}_i, \mathbf{n}_k, \mathbf{n}_i], \quad (4)$$

$$\begin{aligned} \mathbf{S}_l &= \mathbf{R} \mathbf{S}_k \mathbf{R}^\top, \\ \mathbf{S}_j &= \mathbf{R} \mathbf{S}_i \mathbf{R}^\top, \end{aligned} \quad (5)$$

¹It can be reduced to at most two-fold ambiguity if an additional oriented normal vector is given besides \mathbf{S} . The analysis here assumes singular values of multiplicity one in \mathbf{D} .

²Optionally, we can require the corresponding structure matrices $\mathbf{S}(\mathbf{x})$ and $\mathbf{S}(\mathbf{y})$ to be consistent for all $(\mathbf{x}, \mathbf{y}) \in P$. This is enforced in match pre-selection, using R2 and based e.g. on maximal allowed rotation.



| | y_1 | y_2 |
|-------|--------------|--------------|
| x_1 | [0.60, 0.70] | [0.85, 0.95] |
| x_2 | [0.65, 0.75] | [0.72, 0.82] |

(c) similarity intervals $c(p)$

Figure 1. A 2×2 matching problem with the SSK solution $M = \{(x_1, y_2), (x_2, y_1)\}$.

where \mathbf{n}_i is the surface normal vector at $\mathbf{x}_i \in \mathcal{S}_1$ and \mathbf{m}_j is the surface normal vector at $\mathbf{y}_j \in \mathcal{S}_2$, \mathbf{S}_l and \mathbf{S}_k are the respective structure matrices, etc. The $[\cdot, \cdot, \cdot]$ is a matrix composed of three column vectors. This binary relation over pairs from P will be called *geometric discordance* and corresponds to (the complement of) R3. The diagonal edge in \mathcal{D} in Fig. 1(a) represents such geometric discordance edge.

2.2. Matchings and Strict Sub-Kernels

From the above, it should be clear that if M is a solution to the matching problem, no pair of elements in M must be connected by edge in \mathcal{D} . In other words, every admissible matching M must be an independent vertex subset of \mathcal{D} [1]. There will, of course, be many independent vertex sets in \mathcal{D} . The goal of matching is to select the one that is in best agreement with data. We use a criterion based on invariant feature similarity. Suppose we have a statistic that assigns a real interval $c(p) = [\underline{c}(p), \bar{c}(p)]$ to every putative correspondence $p = (\mathbf{x}, \mathbf{y}) \in P$. The statistic compares the neighborhood of surface point $\mathbf{x} \in \mathcal{S}_1$ with the neighborhood of $\mathbf{y} \in \mathcal{S}_2$ based on invariant features. We may think of $\bar{c}(p)$ as the usual similarity and the difference $\bar{c}(p) - \underline{c}(p) \geq 0$ as the sensitivity of the similarity to small data perturbation. Using $c(\cdot)$, we orient the edges of the graph \mathcal{D} to obtain oriented graph $\mathcal{G} = (P, A \cup A^*)$, in which A is the set of reversible (bidirectional) arcs,³ A^* is the set of irreversible (unidirectional) arcs, and $A \cap A^* = \emptyset$. An arc $(p, q) \in A \cup A^*$ is said to be irreversible if $(q, p) \notin A \cup A^*$, otherwise it is said to be reversible. The edge $\{p, q\} \in E$ will become an irreversible arc $(p, q) \in A^*$ (i.e. an oriented edge from p to q) if the intervals $c(p)$ and $c(q)$ do not intersect and $c(p) \prec c(q)$. All other edges in E (i.e. those when the intervals intersect) will become reversible arcs in

³Following standard terminology, arcs are oriented edges [1].

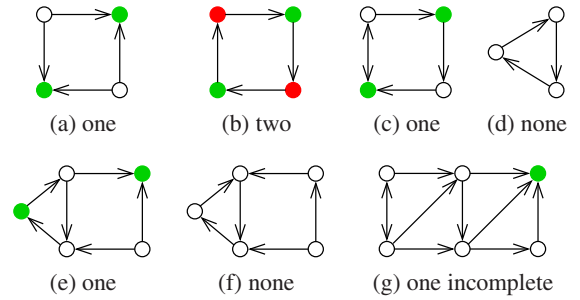


Figure 2. A few examples of strict sub-kernel problem in general oriented graphs with the solutions showed as filled vertices. Number of solutions is indicated. Only the case (a) is an interval orientation.

A. Such orientation of graph edges will be called *interval orientation*. Fig. 1(b) shows an example of interval orientation.

The interval $c(p)$ is related to the quality of a putative match $p \in P$. The quality of the whole matching M will be defined in terms of the strict sub-kernel [16]. A vertex subset $K \subseteq P$ is called a *strict sub-kernel* (SSK) of oriented graph $\mathcal{G} = (P, A \cup A^*)$ if

1. K is an independent vertex subset in \mathcal{G} ,
2. for every $p \in K$ it holds: if there is $(p, q) \in A \cup A^*$ there must be an arc $(q, r) \in A^*$ such that $r \in K$.

Fig. 2 shows a few examples of oriented graphs and their maximum strict sub-kernels.

The matching M we are looking for is identical with the maximum strict sub-kernel K of the graph \mathcal{G} : it is an independent vertex subset of the underlying graph \mathcal{D} and acquires global optimality in the sense of the SSK definition. Maximality means it is the largest-cardinality matching that still fulfills the SSK definition. If the intervals $c(p)$ are non-degenerate, i.e. when $\bar{c}(p) > \underline{c}(p)$ for all $p \in P$, the strict sub-kernel will exhibit robustness, i.e. invariance to small data perturbations. A full discussion of this property is beyond the scope of this paper.

As an example, consider the small matching problem in Fig. 1. We have two points per surface, $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}_1$ and $\mathbf{y}_1, \mathbf{y}_2 \in \mathcal{S}_2$. As we have discussed, the discordance over pairs of putative matches is captured by the unoriented graph \mathcal{D} . The diagonal edge between $(\mathbf{x}_1, \mathbf{y}_1)$ and $(\mathbf{x}_2, \mathbf{y}_2)$ represents a fact there is no rigid motion that can bring $\mathbf{x}_1, \mathbf{x}_2$ into simultaneous alignment with $\mathbf{y}_1, \mathbf{y}_2$ (as verified by (4) and (5)). The invariant feature similarity intervals $c(\cdot)$ are given in the table in Fig. 1(c), and the maximum strict sub-kernel of the graph \mathcal{G} oriented by the intervals is the set of green vertices in \mathcal{G} in Fig. 1(b). The matching $M = \{(x_1, y_2), (x_2, y_1)\}$ is then the solution of the registration problem.

Note two things. First, *the matching does not necessarily optimize any additive cost function*.⁴ Second, the matching *need not be complete*: it does not necessarily have the largest possible cardinality of all matchings.⁵ This is a favorable property of SSK: if evidence for correspondence is insufficient in data or if it is ambiguous, a part of the data is rejected rather than forcefully interpreted. As a result, the matching will not suffer from *mismatches*, but may suffer from *sparsity*. We prefer this behavior over getting wrong solutions.

In the case when a (simple) graph is oriented by intervals as discussed above, it can be proved that there is always at most one maximum strict sub-kernel [16]. Hence, *range image registration based on SSK will provably find a unique solution or it rejects data as ambiguous*. The rejection typically occurs if the object is perfectly symmetric, if it is featureless, or very noisy.

2.3. Strict Sub-Kernel Algorithm

It can be proved [16] that the following polynomial algorithm for finding SSK can be used for interval orientations.⁶ We say a vertex p of an oriented graph $\mathcal{G} = (P, A \cup A^*)$ is a sink if it has no successor, i.e. if there is no $q \in P$ such that $(p, q) \in A \cup A^*$. Note that an isolated vertex is a sink. The algorithm proceeds as follows.⁷

Input: Interval-oriented graph $\mathcal{G} = (P, A \cup A^*)$.

Output: Maximum strict sub-kernel K .

Procedure:

1. Initialize $K := \emptyset$.
2. If there is no sink in P , terminate.
3. Find a sink $s \in P$.
4. Add s to K .
5. Remove s and all its predecessors from P .
6. Go to Step 2.

If this algorithm is used for range image registration where we have n interest points (IPs) on surface \mathcal{S}_1 and n IPs on \mathcal{S}_2 , the computational complexity of matching is $O(n^4)$. The $O(n^4)$ time includes constructing \mathcal{G} . There is an $O(n^3)$ SSK algorithm that is more complex [16, 15]. It is omitted from description here for lack of space.

⁴It maximizes the sum of $\bar{c}(\cdot)$ in the above example just coincidentally, it is easy to construct an example when it does not.

⁵Although it does in the example, again by coincidence. But in the extreme case when all arcs are reversible, for instance, there would be no SSK. Check also Fig. 2(g).

⁶The class of orientations for which this is a valid algorithm is somewhat larger than just interval orientations. In general oriented graphs, the problem is NP-complete. Similarly, the uniqueness theorem holds for a larger class than interval orientations [16] but does not hold generally.

⁷The reader may try testing the algorithm on the example in Fig. 1 or Fig. 2(a) (but not on the other examples in Fig. 2, they are not interval orientations).

3. Range Image Registration

Using the formalisms presented in the previous section the range image registration procedure is described here. The task is to find a mapping from the set of interest points (IPs) on surface (range image) \mathcal{S}_1 to the set of IPs on surface \mathcal{S}_2 . We first choose the sets of IPs I_1, I_2 , on \mathcal{S}_1 and on \mathcal{S}_2 independently. Then similarity intervals $c(p)$ are calculated for all pairs in $P = I_1 \times I_2$. Finally, the IPs are matched using the SSK algorithm.

3.1. Surface Features

Surface features should be sufficiently local to enable registering complex range images with many discontinuities. On the other hand they should be as discriminative as possible. By discriminability we mean the ability to distinguish the correct match using local decision based on feature similarity alone.

The features we will use are computed from augmented triangular mesh which includes all possible triangles among triples of vertices in a small vertex neighborhood, see Fig. 3(a). We have chosen three local features: (1) oriented surface normal, (2) structure matrix, (3) triple features. Given the surface as an augmented triangular mesh, the triple feature at point \mathbf{x} is a collection of elementary features $F(\mathbf{x}) = \{f_i(\mathbf{x}), i = 1, \dots, t\}$ over all t oriented triangles in the neighborhood of point \mathbf{x} that share \mathbf{x} :⁸

$$f_i(\mathbf{x}) = \frac{\det[\mathbf{n}, \mathbf{n}_1, \mathbf{n}_2]}{\|(\mathbf{x}_1 - \mathbf{x}) \times (\mathbf{x}_2 - \mathbf{x})\|}, \quad (6)$$

where $\mathbf{n}, \mathbf{n}_1, \mathbf{n}_2$ are oriented surface normals at triangle vertices \mathbf{x}, \mathbf{x}_1 , and \mathbf{x}_2 , respectively, the numerator is the triple product of the three vectors including the sign and the denominator is the area of the triangle. The normalization by area reduces the influence of surface discretization. Triple product was chosen because it discriminates convex and concave surface point neighborhoods, unlike pairwise features defined over edges in the triangulation. For non-degenerate neighborhoods, the collection $F(\mathbf{x})$ is quite large, which allows working with the *distribution* of the elementary feature values at a surface point.

The surface normal is first computed as the mean normal over all elementary triangles of the augmented mesh in a small neighborhood. In our experiments we used 7×7 neighborhood, see Fig. 3(b). This gives satisfactory results as can be seen from re-illuminated maps in Fig. 5. These normal vectors are then used to compute the structure matrices $\mathbf{S}(\mathbf{x})$ and the elementary triple feature collections $F(\mathbf{x})$ at each surface vertex \mathbf{x} . The triple features and structure

⁸For the central node e in Fig. 3(a) we have $t = 28 - 4 = 24$ and the set of triangles is $\{aeb, aec, ade, aef, age, ahe, bec, bde, bef, bge, bei, cde, cef, ceh, cei, dge, dhe, die, egf, ehf, eif, egh, egi, ehi\}$.

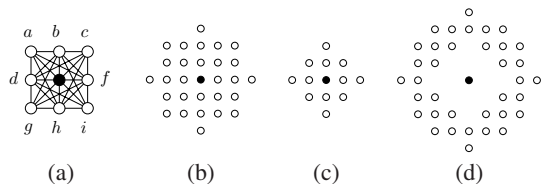


Figure 3. Augmented triangulation over 3×3 vertex e neighborhood gives 24 elementary triangles that share the central vertex (a), local surface vertex neighborhood used to estimate local normal vector \mathbf{n} from 332 triangles (b), the neighborhoods for computing the first two triple feature sets F^1 (c) and F^2 (d) give 52 and 604 oriented triangles, respectively.

matrices are computed over k different non-overlapping circular neighborhoods of increasing radius. In our experiments we used $k = 2$, see Fig. 3(c,d).

To summarize, every vertex \mathbf{x} of the discrete surface \mathcal{S} is augmented with (1) unit normal vector $\mathbf{n}(\mathbf{x})$, (2) k structure matrices $\mathbf{S}^j(\mathbf{x})$, $j = 1, \dots, k$, one per each circular neighborhood and (3) k feature collections $F^j(\mathbf{x})$, one per circular neighborhood. The similarity between two collections $F^j(\mathbf{x})$ and $F^j(\mathbf{y})$ is the complement of Kolmogorov-Smirnov distance $1 - \text{KS}(F^j(\mathbf{x}), F^j(\mathbf{y}))$. The aggregated similarity $\bar{c}(\mathbf{x}, \mathbf{y})$ of surface points \mathbf{x} and \mathbf{y} is then the product over all circular neighborhoods

$$\bar{c}(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^k \left(1 - \text{KS}(F^j(\mathbf{x}), F^j(\mathbf{y}))\right). \quad (7)$$

The sensitivity $\bar{c}(\mathbf{x}, \mathbf{y}) - \underline{c}(\mathbf{x}, \mathbf{y})$ is computed from the difference in respective K-S distances computed from the triple feature collections and the collections artificially perturbed by uniform-distribution noise.⁹

3.2. Interest Point Detection and Selection

Not all surface points are equally suitable for matching. For efficiency, the interest points should be well localized and mutually non-interchangeable. Two IPs are interchangeable if they are similar and lie both within the valid range of rigid transforms. We perform three steps.

The IPs are first detected as local maxima of localizability and then they are selected (pruned) based on mutual geometric concordance. In our current implementation, localizability $L(\mathbf{x})$ at point \mathbf{x} is measured as the standard deviation of the set of all triple features

$$L(\mathbf{x}) = \text{std} \bigcup_{i=1}^k F^i(\mathbf{x}). \quad (8)$$

This favors points whose neighborhood is non-uniform.

⁹A simple closed-form formula exists but is omitted for lack of space.

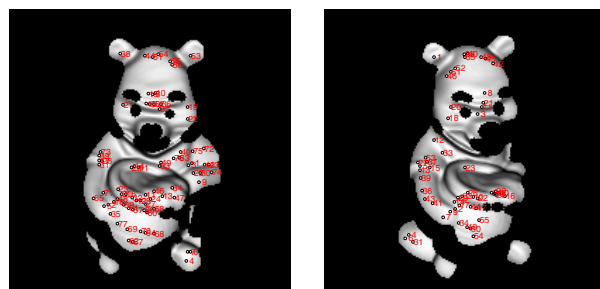


Figure 4. The first 77 interest points (dots) ranked by $L(\mathbf{x})$ on the 1st pair of the `PooH` dataset. See Fig. 5 for IPs that become selected for matching. The point numbers in this figure and in Fig. 5 are not consistent.

Working independently in each image, the IP pre-selection procedure gives a subset of detected IPs that are mutually geometrically *discordant* and are ranked high by (8). This guarantees that IPs that are interchangeable do not enter matching. Hence, repeated structures do not enter matching. This is important because the subsequent matching algorithm would otherwise tend to reject all data.

The selection problem is also formalized using strict sub-kernels. The vertices of the unoriented graph \mathcal{D} are all detected IPs, as those in one of the images in Fig. 4. Two IPs $\mathbf{x}, \mathbf{y} \in \mathcal{S}$ are joined by edge in \mathcal{D} if there is an orthogonal matrix \mathbf{R} in the allowed range of rotations that satisfies (3) for some \mathbf{P} preserving surface normals: $\mathbf{n}(\mathbf{y}) = \mathbf{R} \mathbf{n}(\mathbf{x})$. The orientation is constructed from the degenerate intervals $\underline{c}(\mathbf{x}) = \bar{c}(\mathbf{x}) = L(\mathbf{x})$.¹⁰ Note that in this problem we can neither use any pairwise similarities like those in (6) nor we can use the geometric constraint (4). On the other hand the computational complexity is kept at $O(n^2)$, since the graph \mathcal{D} has only n vertices (n is the number of interest points).

Finally, in the 3rd step, matching table P is constructed from all possible matches of pre-selected IPs in the two range images. Putative pairs in this table failing to pass the Requirement R2 test are deleted from P . Rotation is estimated by (4). The user-defined search range is a parameter of this procedure.

The effect of interest point selection is illustrated in Fig. 4 (detected IPs) and Fig. 5 (selected IPs).

3.3. Matching

Given the selected interest points I_1, I_2 in both images, the matching solves the maximum strict sub-kernel problem. The matching table P , which is a subset of the product $I_1 \times I_2$ defines the vertices of the unoriented graph \mathcal{D} . Two vertices in \mathcal{D} are joined by edge if they are geometrically discordant based on (4), (5) or if the uniqueness constraint is not satisfied, as already discussed in Sec. 2. The orien-

¹⁰This will have to be improved in the future.

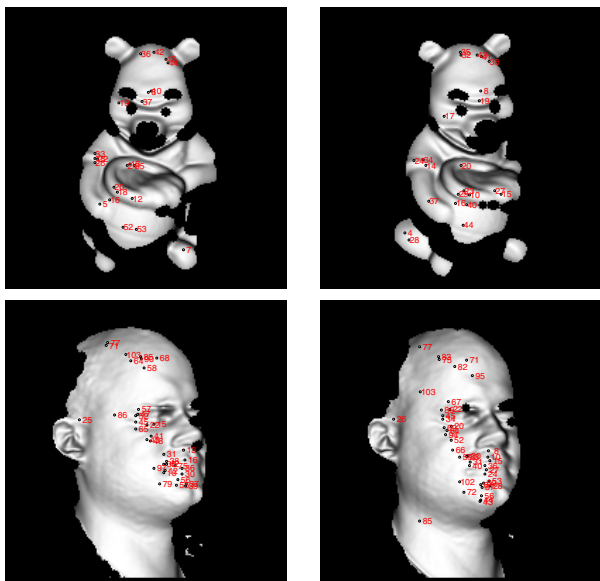


Figure 5. Selected IPs in the Pooh (top row) and the Rick1 datasets (bottom row).

tation of the graph is constructed from (7), as described in Sec. 2. The SSK algorithm is then used, as described above.

4. Experiments

This experiment shows the performance of the proposed registration procedure on two real datasets Pooh and Rick1 obtained from [23]. The Pooh object is an example of a plastic toy of relatively smooth, featureless surface scanned at 20 degree rotation steps. The Rick1 dataset represents a non-trivial surface with realistic noise due to subsurface light scattering. The rotation in this set is irregular and it is about three times larger in the last pair of the sequence as in the other pairs. All range images are 200×200 pixels. Color images that are part of the datasets were not used. All parameters are set so that they are not critically important for the success of matching.

The detected interest points for two images of the Pooh set are shown in Fig. 4. The top-rank points are at high-curvature locations. The selected IPs for a pair in each set are shown in Fig. 5. Note that they focus on prominent local features quite well. The selection assumed the rotation is within $20^\circ \pm 15^\circ$ in the Pooh dataset and $-10^\circ \pm 15^\circ$ in the Rick1 dataset except for the last pair where it is assumed $-20^\circ \pm 15^\circ$. The location of the IPs was determined to sub-pixel accuracy using local quadratic fit to the response of the localizability function L . All local features were interpolated at this position.

The final matching for one pair of range images in each dataset is shown in Fig. 6. The lines show the motion vectors of each matched IP. Using the matched points the two

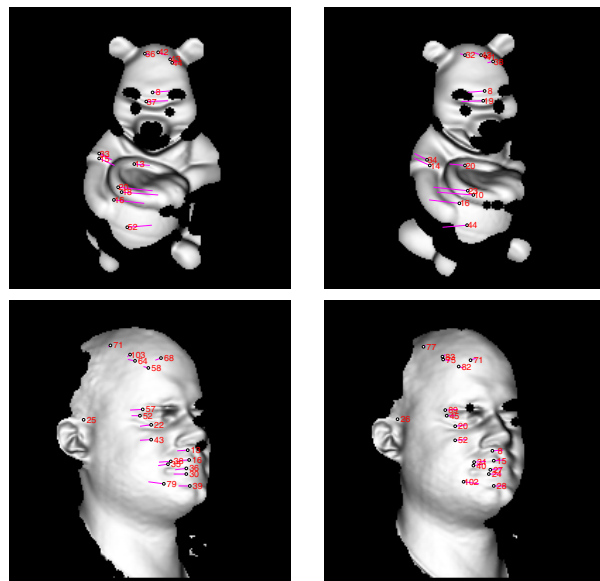


Figure 6. Matched points in the Pooh (top row) and the Rick1 datasets (bottom row).

range images can be registered.

To test the performance we run the entire matching pipeline on the whole datasets. Only the neighboring-image pairs were matched in the sequence. Table 1 overviews the results and Figs. 7–8 show them in detail, including 3D registrations. In one case (pair 8–9 of the Pooh set) no matching was found due to lack of discriminable IPs on a relatively featureless surface. The relative lack of local surface features makes the Pooh dataset rather difficult for the proposed algorithm. A few small matching errors occurred (the largest one in the last pair of the Rick1 set), due to a relatively wide tolerance when testing geometric discordance.

5. Discussion and Conclusions

In this paper we described range image registration method that is globally convergent. Namely, we first represent all putative point matches in an oriented graph, whose edges represent geometric matching constraints with orientation representing feature similarity, and then select as many consistent matches among them as possible. The registration problem is formalized as a discrete optimization task of finding a maximum strict sub-kernel in the graph.

Our future work will focus on more discriminable surface features. Note we used very small neighborhood features in this paper. Discriminability of such features is not large. This is the reason why the matching is so sparse. But thanks to the richness of covariant features the estimated rigid motion is quite close to the correct one as can be seen in Figs. 7–8.

To improve discriminability, color can be used. It is pos-

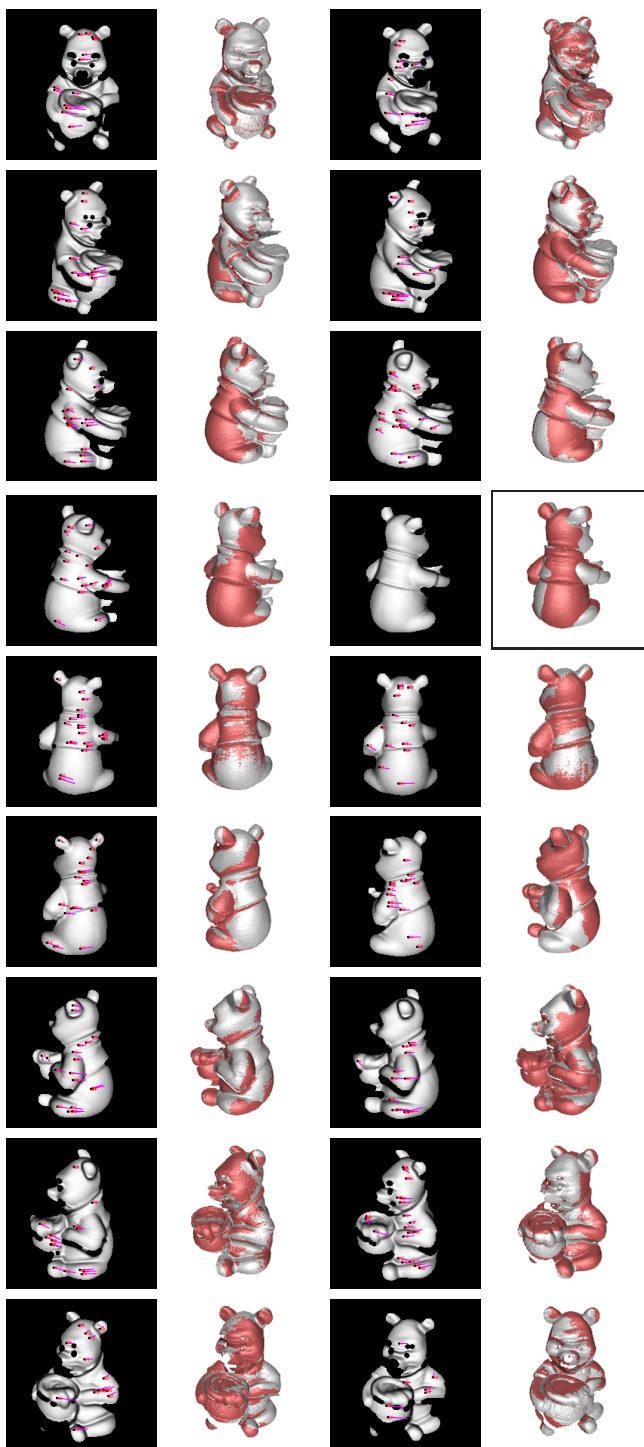


Figure 7. All pairwise coarse registrations in the Pooh sequence. The first and third columns show the matching and the second and fourth columns show the 1st image registered onto 2nd in a respective pair. The boxed result shows unregistered maps because no matching was found for that pair. See also Tab. 1.



Figure 8. All pairwise coarse registrations in the Rick1 sequence. Wrong correspondence 26 on the lip in the last pair caused a slight registration bias that shows around the nose.

sible to generalize the strict sub-kernel problem for graphs whose orientation is given by the combination of two orientations. The problem is still polynomial but the algorithm is different and of higher algorithmic complexity, see [16]. This extension is left for future work.

The matching essentially relies on finding a correspondence that “is clearly correct” (as measured by the similarity statistic). Such correspondence does not necessarily have the largest similarity value. The first match then constrains the matching by limiting the acceptable rigid motions to only those that are consistent with the first match. As the matching procedure progresses, it is more and more guided by the matches found so far, since the consistency is ever more hard to fulfill. This counter-balances the relative decrease in similarity. As a result, the matching is sparse but, on the other hand, some of the correspondences are at locations of not necessarily strong surface features. This, we believe, is a favorable property.

Acknowledgements This work is in part supported by Grant-in-Aid for Scientific Research of the Ministry of Education, Culture, Sports, Science and Technology of Japan under the contracts of 13224051, 14380161 and 16650040, in part by the Czech Academy of Sciences under the contract of 1ET101210406 and by the Czech Ministry of Education under the contract ME 678. A part of this work was done under the framework of Memorandum of Understanding between the Czech Technical University and National Institute of Informatics.

Table 1. Whole sequence matching and registration results. Per rows: image sequence number i ; number of selected IPs in i -th and $i + 1$ -th map, respectively; matching cardinality $|M|$ for pair $i, i + 1$; rotation angle α in degrees estimated from coarse registration; the angle in degrees between rotation axis estimated from the given pair $(i, i + 1)$ and the rotation axis estimated from all pairs: ϕ_{CR} for the coarse registration and ϕ_{ICP} after ICP initialized from coarse registration, respectively; mean closest-point distances before registration ε_0 , after coarse registration ε_{CR} , and after ICP initialized from coarse registration ε_{ICP} .

| Pooh | | | | | | | | | |
|---------------------|------|------|------|------|------|------|------|------|------|
| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| $ I_1 $ | 23 | 37 | 38 | 44 | 61 | 54 | 70 | 57 | 63 |
| $ I_2 $ | 23 | 41 | 43 | 52 | 54 | 64 | 55 | 74 | 63 |
| $ M $ | 13 | 11 | 17 | 14 | 17 | 22 | 18 | 0 | 22 |
| α | 19.6 | 20.5 | 18.9 | 19.6 | 18.5 | 18.8 | 19.0 | - | 19.5 |
| ϕ_{CR} | 1.4 | 1.9 | 0.8 | 2.2 | 0.7 | 3.8 | 3.4 | - | 0.6 |
| ϕ_{ICP} | 2.1 | 1.3 | 0.4 | 1.5 | 0.6 | 1.0 | 3.4 | - | 0.6 |
| ε_0 | 1.05 | 1.19 | 1.42 | 1.54 | 1.53 | 1.49 | 1.26 | 0.99 | 0.82 |
| ε_{CR} | 0.21 | 0.09 | 0.28 | 0.22 | 0.21 | 0.23 | 0.35 | 0.99 | 0.08 |
| ε_{ICP} | 0.18 | 0.08 | 0.24 | 0.21 | 0.17 | 0.13 | 0.26 | 0.82 | 0.05 |

| i | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---------------------|------|------|------|------|------|------|------|------|------|
| $ I_1 $ | 54 | 67 | 53 | 64 | 64 | 70 | 62 | 49 | 37 |
| $ I_2 $ | 60 | 68 | 59 | 61 | 70 | 69 | 63 | 38 | 25 |
| $ M $ | 14 | 18 | 14 | 18 | 13 | 15 | 20 | 14 | 7 |
| α | 18.5 | 18.7 | 18.8 | 19.3 | 20.2 | 19.7 | 21.2 | 19.6 | 20.8 |
| ϕ_{CR} | 2.0 | 2.1 | 6.5 | 0.5 | 1.3 | 1.4 | 1.5 | 1.4 | 1.1 |
| ϕ_{ICP} | 1.8 | 1.5 | 5.5 | 0.4 | 1.3 | 0.8 | 2.5 | 0.5 | 1.3 |
| ε_0 | 0.72 | 0.77 | 0.88 | 1.24 | 1.54 | 1.86 | 1.72 | 1.79 | 1.20 |
| ε_{CR} | 0.14 | 0.25 | 0.33 | 0.15 | 0.18 | 0.11 | 0.28 | 0.28 | 0.27 |
| ε_{ICP} | 0.11 | 0.21 | 0.20 | 0.15 | 0.10 | 0.10 | 0.21 | 0.15 | 0.26 |

| Rickl | | | | | | | | |
|---------------------|------|------|------|------|------|------|------|------|
| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| $ I_1 $ | 37 | 35 | 28 | 34 | 32 | 35 | 31 | 52 |
| $ I_2 $ | 41 | 37 | 28 | 34 | 33 | 32 | 33 | 58 |
| $ M $ | 18 | 17 | 10 | 9 | 14 | 18 | 10 | 11 |
| $-\alpha$ | 10.2 | 11.6 | 9.0 | 13.4 | 14.3 | 9.4 | 13.0 | 29.1 |
| ϕ_{CR} | 5.0 | 6.2 | 15.0 | 3.1 | 5.0 | 7.4 | 5.8 | 5.1 |
| ϕ_{ICP} | 5.0 | 4.9 | 4.9 | 3.2 | 4.2 | 4.2 | 3.1 | 1.6 |
| ε_0 | 1.79 | 5.71 | 4.72 | 3.25 | 9.38 | 2.51 | 6.07 | 8.06 |
| ε_{CR} | 0.35 | 0.43 | 0.96 | 0.58 | 0.43 | 0.45 | 0.60 | 1.14 |
| ε_{ICP} | 0.24 | 0.26 | 0.46 | 0.39 | 0.26 | 0.31 | 0.37 | 0.72 |

References

- [1] C. Berge. *Graphs and Hypergraphs*. North-Holland, 1973.
- [2] P. J. Besl and N. D. McKay. A Method for registration 3-D shapes. *IEEE Trans. on PAMI*, 14(2):239–56, 1992.
- [3] R. J. Campbell and P. J. Flynn. A Survey of free-form object representation and recognition techniques. *CVIU*, 81:166–210, 2001.
- [4] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. *IVC*, 10(3):145–55, 1992.
- [5] C. S. Chua and R. Jarvis. 3D free-form surface registration and object recognition. *IJCV*, 17(1):77–99, 1996.
- [6] C. Dorai, G. Wang, A. K. Jain, and C. Mercer. Registration and integration of multiple object views for 3D model construction. *IEEE Trans. on PAMI*, 20(1):83–9, 1998.
- [7] J. Feldmar, N. Ayache, and F. Berrig. Rigid, affine and locally affine registration of free-form surfaces. *IJCV*, 18(2):99–119, 1996.
- [8] G. Godin, D. Laurendeau, and R. Bergevin. A method for the registration of attributed range images. In *Proc. of 3DIM*, pages 179–86, 2001.
- [9] E. Guest, E. Berry, R. A. Baldock, M. Fidrich, and M. A. Smith. Robust point correspondence applied to two and three-dimensional image registration. *IEEE Trans. on PAMI*, 23(2):165–79, 2001.
- [10] K. Higuchi, M. Hebert, and K. Ikeuchi. Building 3-D models from unregistered range images. *GMIP*, 57(4):315–33, 1995.
- [11] A. E. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Trans. on PAMI*, 21(5):433–49, 1999.
- [12] P. Krsek, T. Pajdla, and V. Hlavac. Differential invariants as the base of triangulated surface registration. *CVIU*, 87:27–38, 2002.
- [13] T. Masuda. Registration and integration of multiple range images by matching signed distance fields for object shape modeling. *CVIU*, 87:51–65, 2002.
- [14] S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Proc. of 3DIM*, pages 145–52, 2001.
- [15] R. Šára. Finding the largest unambiguous component of stereo matching. In *Proc ECCV*, pp. 900–14, 2002.
- [16] R. Šára. Strict sub-kernels. In preparation.
- [17] G. C. Sharp, S. W. Lee, and D. K. Wehe. ICP registration using invariant features. *IEEE Trans. on PAMI*, 24(1):90–102, 2002.
- [18] L. Silva, O. R. P. Bellon, and K. L. Boyer. Enhanced, robust genetic algorithms for multiview range image registration. In *Proc. of 3DIM*, pages 268–75, 2003.
- [19] F. Stein and G. Medioni. Structural indexing: Efficient 3-D object recognition. *IEEE Trans. on PAMI*, 14(2):125–45, 1992.
- [20] G. Turk and M. Levoy. Zipped polygon meshes from range images. In *ACM SIGGRAPH Computer Graphics*, pages 311–8, 1994.
- [21] J. V. Wyngaerd and L. V. Gool. Automatic crude patch registration: Toward automatic 3D model building. *CVIU*, 87:8–26, 2002.
- [22] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *IJCV*, 13(2):119–52, 1994.
- [23] The Ohio State University Range Image Repository. <http://sampl.eng.ohio-state.edu/~sampl/data/3DDB/RID/minolta/>.