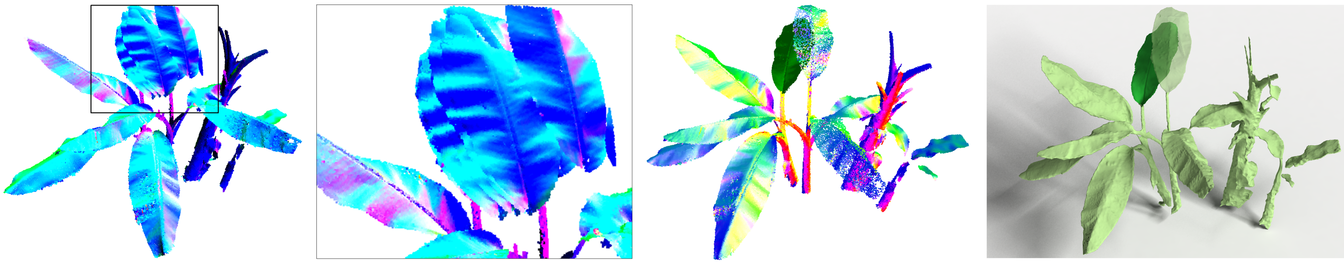
Proactive 3D Scanning of Inaccessible Parts

Feilong Yan* Andrei Sharf† Wenzhen Lin* Hui Huang* Baoquan Chen‡*
*Shenzhen VisuCA Key Lab / SIAT †Ben Gurion University ‡Shandong University



(a) (b) (c) (d)

Figure 1: *Proactive scanning of a tangled Banana tree. We freely move the scanner around to capture the scene while physically moving aside occluding leaves to scan the trunk and branches (a). In (b) we zoom on the interactive modification. Our algorithm reconstructs a 3D scene, including its interior parts (c) and generate a plausible mesh (d).*

Abstract

The evolution of 3D scanning technologies have revolutionized the way real-world object are digitally acquired. Nowadays, high-definition and high-speed scanners can capture even large scale scenes with very high accuracy. Nevertheless, the acquisition of complete 3D objects remains a bottleneck, requiring to carefully sample the whole object's surface, similar to a coverage process. Holes and undersampled regions are common in 3D scans of complex-shaped objects with self occlusions and hidden interiors. In this paper we introduce the novel paradigm of *proactive scanning*, in which the user actively modifies the scene while scanning it, in order to reveal and access occluded regions. We take a holistic approach and integrate the user interaction into the continuous scanning process. Our algorithm allows for dynamic modifications of the scene as part of a global 3D scanning process. We utilize a scan registration algorithm to compute motion trajectories and separate between user modifications and other motions such as (hand-held) camera movements and small deformations. Thus, we reconstruct together the static parts into a complete unified 3D model. We evaluate our technique by scanning and reconstructing 3D objects and scenes consisting of inaccessible regions such as interiors, entangled plants and clutter.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling

Keywords: 3D Scanning, piecewise-smooth segmentation, non-rigid registration, trajectory clustering

Links: DL PDF WEB VIDEO

1 Introduction

In recent years we observe significant advancements in 3D scanning technologies. With emergence of systems such as Kinect®, PrimeSense®and Asus Xtion®, scanners have become commercial off-the-shelf products introducing space-time 3D acquisition to end-users. As scanners continuously evolve, we experience a consistent increase in their resolution, frame-rate and accuracy. Nevertheless, the scanning process is still cumbersome and complex, requiring some level of expertise and tedious steps.

A major challenge in scanning 3D objects is to capture their complete geometry. Proper sampling of the object's surface requires the user to rotate around the object, capturing it from different views. This is a tedious process as the user is required to repeatedly check and rescan local parts that were not captured correctly in a *stop, evaluate and rescan* manner. This problem arises mainly, due to scan accessibility as parts in the 3D object are hidden by physical occlusions and cannot be scanned.

The accessibility problem is especially common in complex objects with large concavities, hidden interior parts and self occlusions which are inaccessible to the scanning device. Nowadays, this problem is handled through a tedious process where parts are scanned separately and registered in a post-process step [Brown and Rusinkiewicz 2007]. While feasible, this is a meticulous job, requiring careful planning and registration of the parts together. In our case, registration of hidden interior parts with their exterior is more challenging since typically there is no overlapping between them.

We present a novel proactive scanning process inspired by the emerging trend of robust scanning systems [Li et al. 2013]. We follow the observation that many times inaccessible regions can be made accessible through a simple and local modification of the scene. For example, given tangled plant, the user scans its dense foliage by moving scanner around and interacts with the scene by locally dragging away leaves to reveal the occluded branches and scan them (Figure 1). Thus, our method enhances the scanning process with user interaction (with the scene) allowing to modify and access inaccessible parts while continuously scanning it. User modifications are incorporated in a holistic manner, as an integral part of a 3D scanning process.

At the core of our method is a 3D registration and motion trajec-

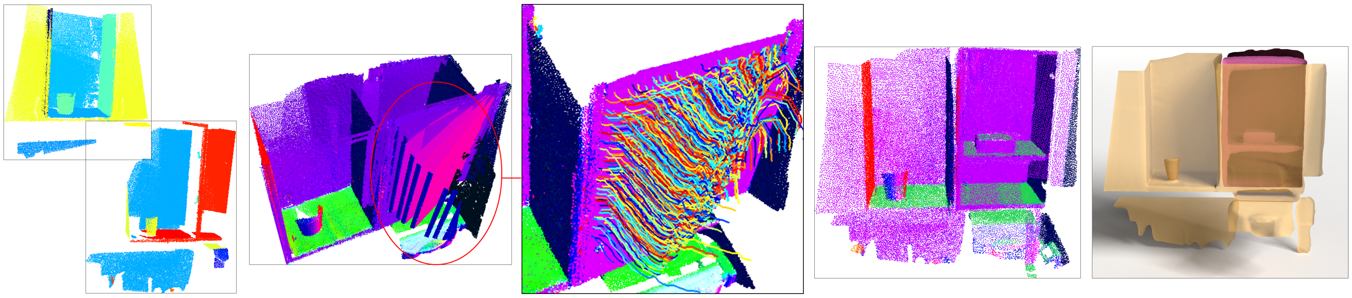


Figure 2: *Overview of our algorithm. We initially over-segment scanned frames into piecewise smooth patches(a). Next, we perform pairwise non-rigid registration of consecutive frames (b) and compute trajectory vectors. We cluster long trajectories (c), belonging to the user's interactive modifications, and accurately reconstruct the complete scene (d-e).*

tory analysis algorithm. Essentially, our dynamic scene consists of three distinct motion types: rigid part implied by the camera motion, small non-rigid local deformation occurring in the 3D scene and large non-rigid motions due to scene modifications by the user. The first two motion types are relatively easy to handle and can be recovered as showed in previous works [Li et al. 2009; Chang and Zwicker 2011]. However, motion due to user interaction is large, non-rigid and may locally affect the spatio-temporal neighborhood. Thus, a clear dichotomy is assumed to exist between the interaction and non-interaction deformations in the scene.

Our algorithm analyzes and registers the raw stream of scanned frames on-the-fly and classifies the different motions according to their specific patterns. This in turn allows to accurately reconstruct the full 3D model and making the following contributions:

- We define a novel *proactive scanning* method which allows the user to perform scene modifications while continuously scanning the 3D scene.
- We present a 3D reconstruction algorithm which registers together scanned frames on-the-fly by analyzing motion trajectories and segmenting out large motions.

2 Related work

Processing and modeling the dynamic scanned data captured by 3D scanners is a challenging computational problem. In recent years numerous methods have been introduced for dynamic 3D registration and reconstruction, presenting a wide range of solutions [Huang et al. 2008; Pekelny and Gotsman 2008; Sharf et al. 2008; Süssmuth et al. 2008; de Aguiar et al. 2008; Wand et al. 2009; Li et al. 2009; Chang and Zwicker 2011; Li et al. 2013]. Since at the core of our method is a dynamic 3D scan modeling algorithm, we focus our discussion on previous work in this field.

Rigid surface registration is a well known problem which has been explored for many years. An early solution for this problem is the classic iterative closest points (ICP) algorithm [Besl and McKay 1992; Chen and Medioni 1992] and its many extensions [Bouaziz et al. 2013]. With the rising popularity of 3D motion-acquisition scanners, there has been a renewed interest in the registration problem for non-rigid 3D surfaces as they move and deform in a scene. ICP variants for non-rigid registration of surfaces have been proposed in [Chui and Rangarajan 2003; Amberg et al. 2007; Brown and Rusinkiewicz 2007]. These methods replace the rigid component with a non-linear smooth deformation and iteratively align surfaces together in a global manner.

Mitra et al. [2007] use kinematic properties of the 4D space-time surface defined by the 3D scan motion in time, to track and register multiple frames of a deformable object. Similarly, Wand et al. [2007] reconstructs both the shape and its trajectories through a 4D shape optimization algorithm and in [Wand et al. 2009] authors formulate surface motion in terms of a displacement field and compute the common shape that deforms and matches the data. Süssmuth et al. [2008] and Sharf et al. [2008] explicitly model and reconstruct a 4D surface from an implicit space-time motion representation. Nevertheless, these techniques are time consuming, requiring a dense sampling of the motion. In our work, we detect and explicitly remove the large non-rigid motions out and perform reconstruction on the effective data.

Methods that assume a specific motion type or an available template success in reducing the problem complexity and narrowing down solution space. Pekelny and Gotsman [2008] present a method for tracking and registration of piece-wise rigid motions of articulated bodies. Their method assumes a given segmentation of the data into rigid parts and an underlying skeletal structure. Similarly, Chang and Zwicker [2008] register scanned pairs of articulated shapes by searching for an optimal set of piecewise rigid transformations. The utilization of a deformable template that can register scans with missing data was introduced in [Chang and Zwicker 2011]. In their work they optimize a set of parameters controlling a reduced deformable model to align and register scans together. In [Li et al. 2009], a smooth template is used to guide the reconstruction process allowing to recover fine geometric details and in [Tevs et al. 2012] consistent trajectories are computed by finding isometrically consistent scene parts in time. Nevertheless, our algorithm does not assume a piecewise rigid motion, nor a predefined template and local isometries. Instead, we focus on tracking and reconstructing a general scene where templates are not available.

To perform non-rigid scan registration, Li et al. [2008] utilize a deformation graph [Sumner et al. 2007]. They compute a graph from the scan points which controls the registration transformations. Similarly, Huang et al. [2008] utilize a graph approximation of geodesic distances in order to extract a set of geodesic consistent correspondences under the assumption of isometric deformations. Nevertheless, we avoid building such a graph construct on top of the point cloud as it is not always practical (for example a sparse scans of a foliage of a plant with many small unconnected parts).

Izadi et al. [2011] use Kinect to explore dynamic user interaction with a reconstructed 3D static scene. Their method focus on scanned motion tracking through foreground/background subtraction. Our problem is significantly harder, as we integrate together 3D interaction, acquisition and reconstruction on-the-fly.

Our work was inspired by the recent work of Li et al. [2013] intro-

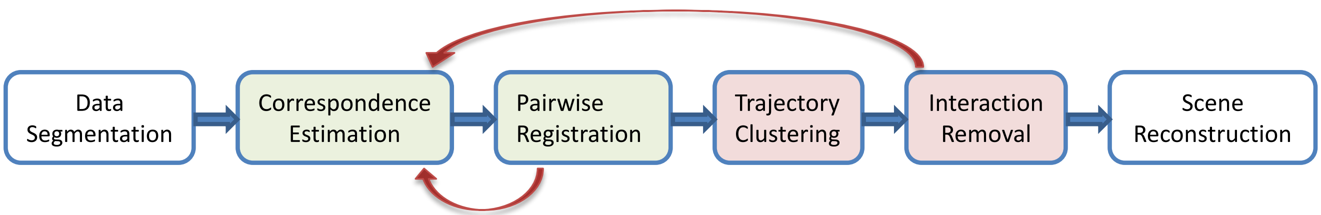


Figure 3: *Algorithm flowchart.*

ducing an end-to-end solution for dynamic scan reconstruction and registration. With their system a user can freely capture a 3D scene in an uncontrolled environment, allowing small deformations and motions. Our work follows in the same path, further simplifying the scanning process, allowing greater ease-of-use and leveraging its methodology and robustness.

3 Overview

Our algorithm process on-the-fly a stream of scanned 3D frames taken with a hand-held scanner. In the acquisition process, the user holds the scanning device and moves around to cover the whole scene from a set of views (denoted as dynamic scanning). Along this process, the user also interacts with the scene in order to modify it, remove occlusions and reveal parts that are hidden or completely inaccessible (see Figure 1(b)).

Initially, we oversegment the scanned frames into smooth patches to reduce data complexity and remove small outliers and noise (Figure 2(a)). Pairwise registration between consecutive frames performs by computing a sparse correspondence and deforming the frames to minimize the distance defined by this correspondence. To tackle the matching-registration problem, we devise a variant of the iterative closest (ICP) point algorithm. We control correspondence density and deformation rigidity and gradually adjust them, while repeating correspondence and deformation process until scans fully align (Figure 2(b)).

In the next step, we compute trajectories using the registered frames and prune out outliers and noisy results. We cluster together trajectories using a trajectory similarity metric and segment out large trajectories which belong to scene modifications (Figures 2(b-c)).

Modification motion is typically large compared to camera movement and scene deformations, hence it introduces significant noise and deviation in the scan registration . A key point to our algorithm is the detection of the user modification interaction and its accurate segmentation from the 3D scan. We refer to interaction intervals as periods in which the user performs a large scene modification to reveal occluded parts. We analyze characteristic features which allow to detect and separate the interaction from other dynamic motions. Thus, we remove large trajectory clusters, and repeat the whole registration process again (Figure 3). This allows us to increase the accuracy of our registration.

Finally, we register the whole scene together. We use the motion registration of the modification motion, to guide registration of whole scene. We demonstrate the plausible surface reconstruction of the fully registered 3D scan (Figures 2(d-e)).

4 Technical Details

Piecewise-smooth segmentation. Given a sequence of scanned 3D frames $\{f_0, f_2, ..., f_n\}$, our aim is to robustly register them into a complete static 3D scene $\mathcal{S}$. In scene completeness we refer to achieving a sufficiently dense sampling of the whole physical object including hidden and occluded regions. Our algorithm performs on-the-fly, processing the stream of frames as they arrive. Thus we process at each time only a limited number of buffered frames.

For a scanned frame f_i we initially cluster points into smooth patches using a hierarchical agglomerative clustering. This step reduces both data complexity and noise, allowing to remove small outliers. Starting from the whole point set, we group together neighboring points within a close distance and with similar normals. More specifically, two points $(p, q) \in f_i$ with normals (n_p, n_q) are grouped together iff:

$$\|p - q\|_2 < \varepsilon_D \ \ , \ \ (n_p \cdot n_q) > \varepsilon_A \tag{1}$$

where $||_2$ denotes the L_2 Euclidean norm and $\cdot$ is the dot product between two normal vectors. Following our scanner's precision and frame-rate specifications we set thresholds to $\varepsilon_D = 2cm$ and $\varepsilon_A = 0.8$. Our underlying assumption here is that motions in the scene do not exceed reasonable speeds. We use the Flann library [Muja and Lowe 2009] to perform fast approximate k-nearest neighbors searches for each point in the input.

Pairwise frame registration. Given two consecutive frames f_k, f_{k+1} we compute the deformation $\mathcal{D}$ that aligns the two frames together. The deformation between consecutive frames consists of a global rigid component $\mathcal{D}_R$ due to camera movement and a non-rigid component $\mathcal{D}_E$ due to local modifications and small scene motions, e.g. small movement of a plant leaves in Figure 1 .

Our registration algorithm performs iteratively from global to local by computing an initial global rigid transformation which is refined by additional local non-rigid deformations. Initially, we compute a sparse set of point correspondences, as the set of closest points sharing similar normal vectors. Formally a point $q \in f_{k+1}$ corresponds to $p \in f_k$ iff:

$$\arg\min_{q \in f_{k+1}} (\|p - q\|_2) \ \ s.t. \ \ (n_p \cdot n_q) > \varepsilon_A \tag{2}$$

We use Horn [Horn 1987] to compute the optimal global rigid transformation $\mathcal{D}_R$ between the two sets of matching points. We apply the transformation only halfway and recompute correspondence and rigid transformation iteratively. This rigid ICP steps an approximates the global rigid transformation component in the scene. Naturally, this approximation depends on the non-rigid motions in the scene which alter our solution.

To compute the non-rigid component of the registration, we take a local approach. Essentially, our scene consists of a mixture of small, local deformations together with few large deformations caused by interactive modifications. This calls for devising a local solution rather than attempting to globally register scans together. Thus, we use our piecewise-smooth patch segmentation and compute a set of local non-rigid transformations which align patches together.

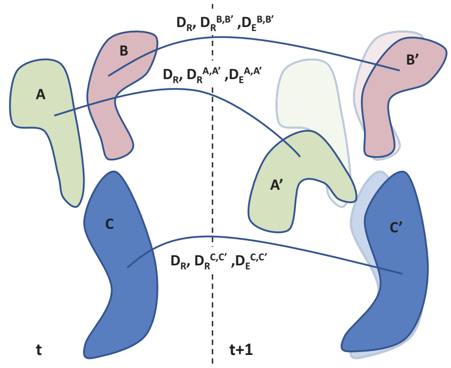


Figure 4: *Registration of two consecutive frames segmented into three patches (colored shapes). Initially we globally align frames using a rigid transformation $\mathcal{D}_R$, followed by patch-wise rigid $\mathcal{D}_R^{p,q}$ and non rigid $\mathcal{D}_E^{p,q}$ transformations.*

Similar to point correspondence, we first compute a patch-based dense correspondence. We replace point with patch in equation 2, taking p as the patch center and n_p as the average patch normal at p, (considering the one-ring neighbors). For each pair of corresponding patches p, q we compute a denser correspondence and compute their rigid transformation in an iterative manner denoted $\mathcal{D}_R^{p,q}$. Finally, for each pair of patches, we replace the rigid transformation with a Thin Plate Splines (TPS) deformation denoted $\mathcal{D}_E^{p,q}$ and repeat the same iterative process to obtain a perfect piecewise patch alignment (see Figure 4).

Trajectory clustering. Given pairwise registered frames, we compute their dense correspondences using equation 2 for the aligned points. Next we generate motion trajectories simply by aggregating pairwise correspondences together along the scan sequence yielding 3D+time polygonal vectors. We formally denote by $\langle\phi^i_{(0,\ldots n)}\rangle$ the ith trajectory in the scene consisting of $n+1$ 3D points, one for each of the n frames (see Figure 5).

Resulting trajectories are noisy since objects are not sampled equally along time and some shifts typically occur. We smooth trajectories using a space-time Laplacian operator:

$$\triangle\phi^i = \frac{\partial^2\phi^i}{\partial x^2} + \frac{\partial^2\phi^i}{\partial t^2} \tag{3}$$

We approximate the discrete Laplacian, by the finite difference method considering the neighboring vertices in the 4D ring of a vertex ϕ^i_j

$$\triangle\phi^i_j = \phi^i_j - \frac{1}{k}\sum_{\phi^l_m \in 1ring(\phi^i_j)} \phi^l_m \tag{4}$$

We then search for clusters of long trajectories using an agglomerative hierarchical clustering algorithm similar to our patch clustering. More specifically, we group together long trajectories iff they are within a close distance and share similarity in their direction. We define our grouping criteria for two trajectories (ϕ^i, ϕ^j) as:

$$\begin{aligned} \|\phi^k\|_2 > \varepsilon_L \ \ , \ \ \forall k \in (i,j) \\ \|\phi^i - \phi^j\|_2 < \varepsilon_D \\ (\phi^i \cdot \phi^j) > \varepsilon_A \end{aligned} \tag{5}$$

where $\|\phi^k\|_2$ is the trajectory length defined as the sum of its part lengths in L_2, the distance between two trajectories is measure as the average distance between their points and their dot product is the average angle between their segments. We set $\varepsilon_L = 7cm$ and repeat the trajectory clustering process as long as there remain unclustered long trajectories.

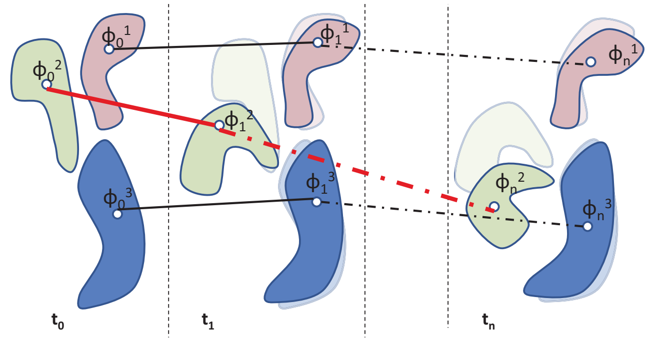


Figure 5: *Trajectories between frames f_0 and f_n. Scene consists of small local deformations of red and blue patches and a large non-rigid deformation of the yellow patch. Our method clusters the large trajectories corresponding to this motion (red line).*

Our method is on-the-fly, processing each time only a constant number of buffered frames from the input (in our experiments we set the buffer size to $n = 200$). Nevertheless it is possible that the interactive modification may be longer than n frames. In order to capture long trajectories as a whole, we adaptively resize the buffer size once a long cluster is detected. Thus, we rescale the buffer by powers of 2 until no more long trajectories are existent.

Modification motion removal Given trajectory clusters, we compute a 4D mask in order to remove long trajectories from the scene. In the simple case, we compute the 4D convex-hull [Barber et al. 1996] of the clustered trajectories and remove its contents, thus removing the modification motion from the scene in a conservative manner.

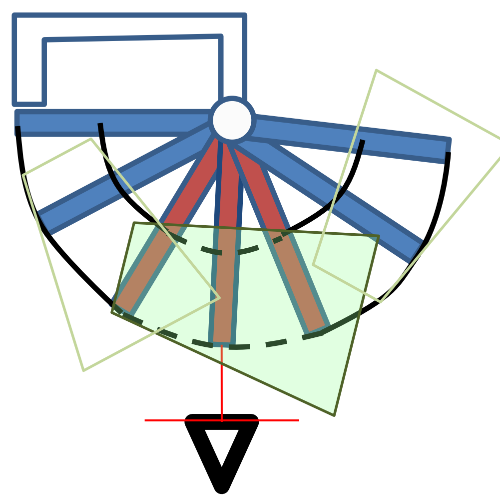
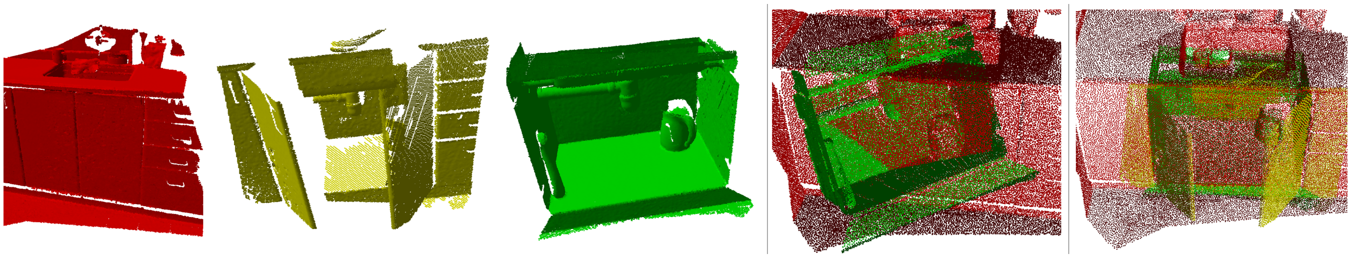


Figure 6: *Illustration of the trajectory convex hull scaling. As a rotating door becomes orthogonal to the camera (black triangle) plane, it disappears (and may reappear again). We define the trajectory convex hull (green trapezoid) to search and collect missing trajectories.*

Nevertheless, since modification motion is large, it may introduce



(a) (b) (c) (d) (e)

Figure 7: *Modification motion guides registration between exterior and interior parts. Given an exterior (a), an interior (c) and a modification (b) frame, we demonstrate incorrect registration of (a)+(c) in (d) vs. correctly registering (a)+(b)+(c) in (e). Note that in (b), motion of the doors is segmented out by our algorithm, leaving only static parts to guide full registration.*

additional deformations in its spatio-temporal neighborhood and may not be accurately scanned. Furthermore, the moving part may completely disappear while becoming orthogonal to the camera plane. In Figure 6, the rotating door (blue) is not captured while aligning with the camera direction (red).

To overcome these problems, we scale the spatio-temporal 4D convex-hull to capture the full modification trajectories. This operation defines a larger space in which we search for similar trajectories that can additionally cluster together. In spatial domain, we scale the convex-hull vertices in outward direction by 10% of the bounding box. In temporal domain, we scale the 4D convex hull by 10% following its principal direction.

In the final step, we remove long trajectory clusters belonging to modification motions. Next, we recompute pairwise frame registration yielding a more accurate global registration. By removing the large modification motions, we now faster converge to the optimal frame registration, since handling mainly globally rigid and small local deformations.

The remaining spatio-temporal neighborhood of the modification provides an important guidance for the registration of exterior and interior parts. Due to little and even no overlap between these regions, their registration is hard. In Figure 7 (d), we demonstrate incorrect registration of an exterior (a) and interior (c) where no overlap exist. A frame belonging to the modification interaction (b) bridges between the exterior and interior by providing an overlap, thus guiding towards the correct registration (e).

Motion processing robustness Our motion removal may not be perfect. Since our piecewise-smooth patches are not be compatible with motion trajectories, the tracked geometry may be discontinuous. E.g. a patch in frame f_k may split in two different patches in frame f_{k+1} and vice versa (Figure 8(b)). Thus, static parts near interaction regions may be classified as moving points and moving points near static regions may be classified as static. We show that this phenomena is rare, of small scale and does not affect our motion removal correctness.

- Piece-wise smooth patches are used for non-rigid registration of adjacent frames, while trajectories are computed independently at point level. Therefore, even if patches and motion are not coherent, it can affect only the registration step.
- Deformations between adjacent frames are assumed to be small and therefore local registration error cannot be large (Figure 8(a-b)).
- Interaction motions are typically few, while patch resolution is typically higher than motion clusters. Hence, they may only introduce very local misalignment errors (Figure 8(c-d)).
- Static points are robust. Static regions are naturally scanned multiple times from different views, as the scanner moves around. Thus, if points are incorrectly labeled as dynamic and removed, points covering the same region in other frames will fill in.
- Motion removal is robust. If points are incorrectly labeled as static, the trajectory convex hull scaling removes the full interaction motion in a non-conservative manner.

5 Results and discussion

We demonstrate our proactive scanning technique on various scanning scenarios and with a wide range of objects and scenes. In order to demonstrate the effectiveness of our method, we allow the user to scan around the scene and freely interact with it by modifying movable parts (see accompanying video). In all our experiments we kept the same parameter values as specified in the paper.

We used a MantisVision™mid-range active-light scanner with $1-6$ meters range, below $1mm$ accuracy, at $15-20$ fps, capturing roughly $40-70k$ points per frame. We have also experimented with an Artec Eva™short range scanner with similar accuracy and frame rates, and a Kinect®scanner with lower resolution and accuracy.

The stream of frames is sent from the scanner to a PC through a firewire connection as they are being scanned. The method was implemented in C++ without any GPU optimizations, running on a Lenovo Intel®Core i5-3210M CPU, 2.5GHz with 8GB RAM. Our processing time is nearly real-time with 2 fps on average. Since our camera typically performs at $15-20$ fps, we use a large buffer to store streaming data.

In our experiments we proactively scanned piecewise rigid objects such as furniture, household items and cars. In these scenes, modification interaction consisted of moving occluding parts through rigid transformations. In Figure 9 we demonstrate our algorithm on four different piecewise rigid scenes. In Figure 9(a) we proactively scan the back of a car, then lift its trunk and acquire its interior. In 9(b), we scan a full car and its interior. In this process, we freely open and close the front and back doors while scanning the car. Figure 9(c) shows a kitchenette with occluded interior. The interactive modification consists of opening two doors one after the another and reconstructing the full scene. Finally, in 9(d) we move two sliding doors of a cabinet through a translational motion and capture the interior.

We have also tested our proactive scanning method with non-rigid objects. Thus, modifications consisted of non-rigid motions to reveal occluded regions in the scene. In Figures 1 and 10(a) we scan two plants with dense foliage which occludes the interior branches and trunks. We proactively move leaves and branches aside to access the interior. In Figure 10(b) we scan a table completely covered by a table-cloth which we move aside to capture the legs underneath the table. Similarly, in Figure 10(c) we interactively lift (one-by-one) the two cloth parts covering the entrance of a tent and reveal its interior. (See supplemental material for many other results).

Our method was able to capture and reconstruct a significant amount of detail by interacting with the scene, moving parts and accessing hidden interiors. as can be seen in the accompanying video, proactive scanning allows to continuously capture and access hidden regions in a straightforward manner.

For visual purposes, we reconstruct the mesh from the point-cloud using Poisson surface reconstruction [Kazhdan et al. 2006]. In rare cases when the global Poisson did not provide a plausible reconstruction, we manually separated the interior and exterior and then we reconstruct them independently and stitch them together (in our experiments, the full car and kitchenette in Figure 9(b,d))

Limitations. Our patch-based scan segmentation is coupled with motion analysis. Thus, a segmentation which is incompatible with the motion may yield inaccurate trajectories and clusters. Nevertheless, motions generated by user interactions are typically few and thus trajectory clusters are of much lower resolution than patches.

Our motion analysis technique was designed to handle a specific interaction of occluder removal, characterized by piecewise rigid or non-rigid deformations that are local in the static scene. In Figure 11(a), the interior of a drawer is moving as the drawer is pulled (left) and thus, is classified as part of the modification motion and removed (right).

To correctly compute motion trajectories we register adjacent frames together assuming a small deformation between them. In this case, a simple global-to-local non-rigid ICP approach can correctly register frames together. Thus, a quick modification motion will not be tracked correctly yielding incorrect registration. In Figure 11(b), the quick lifting of the table cloth from the table (mid) is not correctly tracked and clustered yielding an inaccurate registration (right).

Finally, our method assumes a discrepancy between modification, camera and scene motions, allowing to separate the modification from other motions in the scene. In cases where camera moves abruptly or scene consisting of large self deformations (e.g. high wind in a tree, and etc.) our method fails to segment and classify the modification motion. In Figure 11(c), the scanner moves abruptly while opening the bottom doors of a cabinet. Hence, parts of the static scene are clustered and classified as modification motion and removed. This yields a partial and incorrect registration (rightmost).

6 Conclusion and future work

In this paper we presented the *proactive scanning* paradigm. Its novelty is that it allows the user actively modify the scene while scanning it, in order to reveal and access occluded regions. This presents a holistic approach where both the 3D acquisition of the static scene as well as the modification motion are integrated into one continuous scanning process.

Our results show a progress over traditional acquisition pipeline. Our method allows to capture, register and reconstruct complete scenes revealing their hidden interiors. Furthermore, our accompanying video demonstrates that hidden interiors and occluded parts

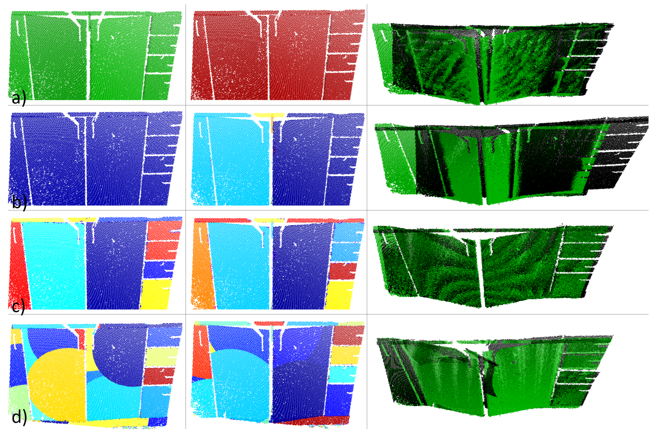


Figure 8: *Segmentation to registration relation. Rows (top-to-bottom) show the effect of increasing the number of segments starting from* 1 *(topmost) on registration accuracy. Columns (left-to-right) show initial segmentation of two frames and their mutual registration (rightmost) colored black and green respectively.*

in the scene can be captured by our method in a relatively simple and straight-forward process.

In the future, we plan to further expand the proactive scanning paradigm and investigate additional interaction possibilities. One immediate direction is incorporating a richer set of gestures to allow an extended control on the scanning process. In fact, the user may guide the acquisition and reconstruction processes on-the-fly. Another interesting venue is in the field of robotics where our proactive scanning can join forces with an autonomous robot creating an enhanced process where the robot scans and interacts with the scene while our method robustly registers the data together.

Acknowledgments

We thank Prof. Daniel Cohen-Or for insightful discussions and reviewers for their valuable comments. This work is supported in part by the NSFC (61232011, 61379090, 61161160567), National 863 Program (2013AA01A604), CAS Young Scientists Plan (2013Y1GA0007), Shenzhen Innovation & Technology Program (CXB201104220029A, ZD201111080115A, KC2012JSJS0019A, JSGG20130624154940238, JCYJ20120617114842361), the Israel Science Foundation and the European FP7.

References

AMBERG, B., ROMDHANI, S., AND VETTER, T. 2007. Optimal step nonrigid icp algorithms for surface registration. In *CVPR*.

BARBER, C. B., DOBKIN, D. P., AND HUHDANPAA, H. 1996. The quickhull algorithm for convex hulls. *ACM TRANSACTIONS ON MATHEMATICAL SOFTWARE 22*, 4, 469–483.

BESL, P. J., AND MCKAY, N. D. 1992. A method for registration of 3-d shapes. *IEEE Trans. Pattern Anal. Mach. Intell. 14*, 2 (Feb.), 239–256.

BOUAZIZ, S., TAGLIASACCHI, A., AND PAULY, M. 2013. Sparse iterative closest point. In *Proc. Eurographics Symp. on Geometry Processing*, 113–123.

BROWN, B. J., AND RUSINKIEWICZ, S. 2007. Global non-rigid alignment of 3-d scans. In *ACM SIGGRAPH 2007 Papers*, ACM, New York, NY, USA, SIGGRAPH '07.

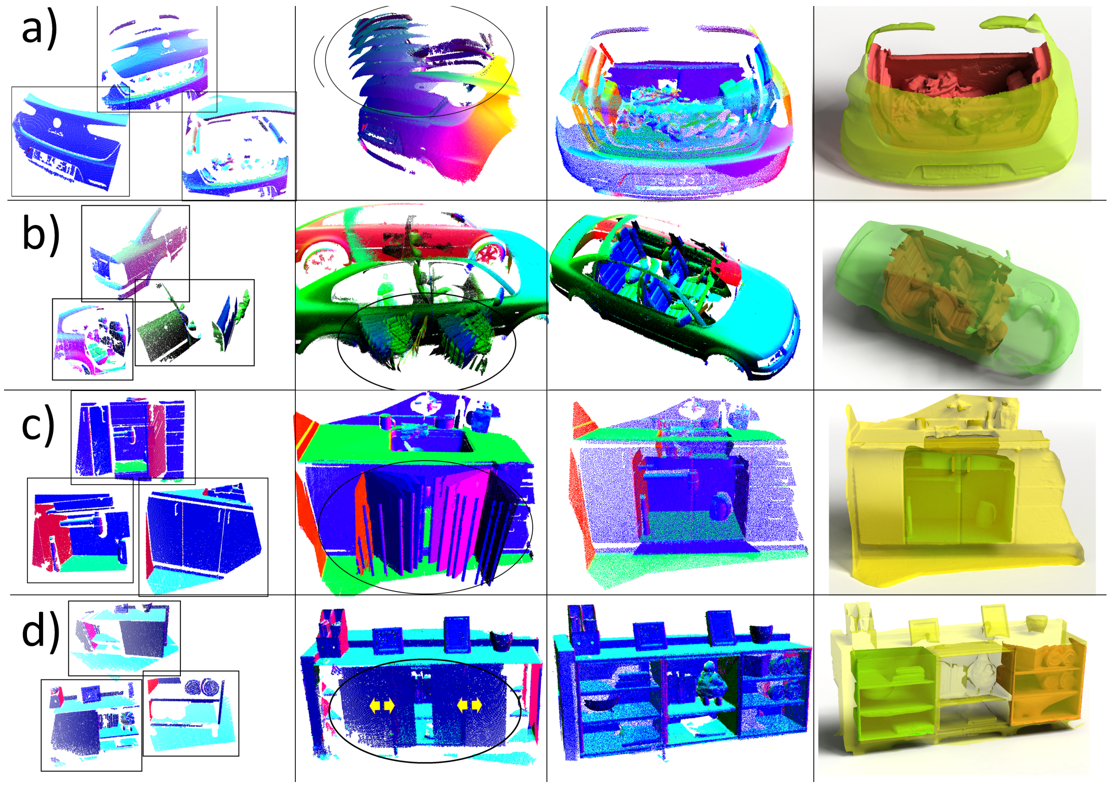


Figure 9: *Results of our proactive scanning with piecewise rigid modifications of scene parts. Left-to-right are three individual frames, the modification motion, full 3D registration and reconstruction.*

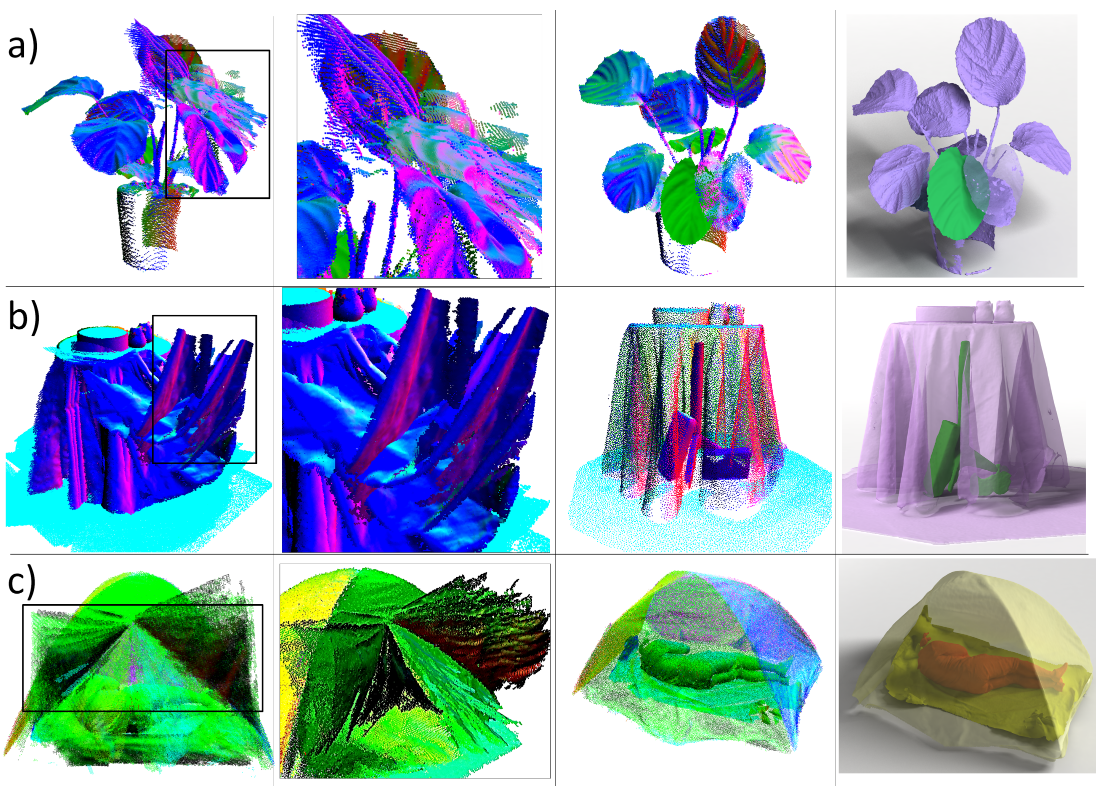


Figure 10: *Proactive scanning with non-rigid modifications of scene parts. Left-to-right is the modification motion, a zoom-in into motion, complete 3D registration and reconstruction.*

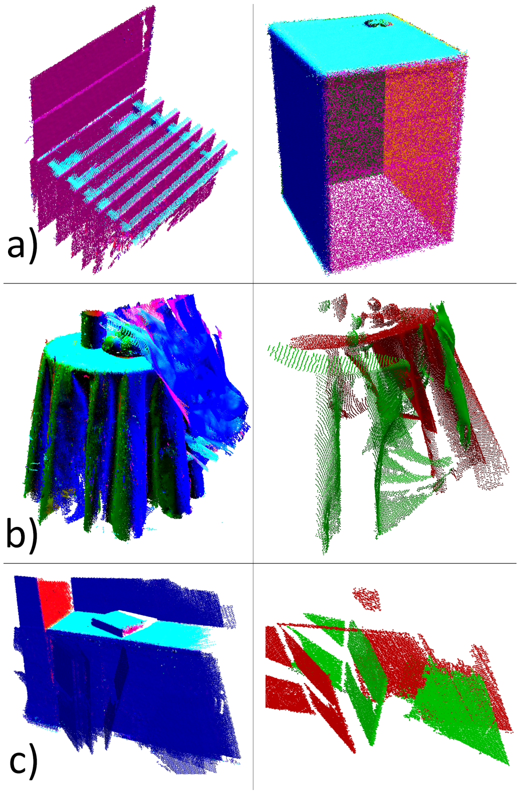


Figure 11: *Proactive scan limitations. Left-to-right is the modification motion and registration result. A sliding drawer interior (a) is removed, large cloth motion (b) and scanner motion (c) yield incorrect registrations.*

CHANG, W., AND ZWICKER, M. 2008. Automatic registration for articulated shapes. *Computer Graphics Forum (Special Issue of Symposium on Geometry Processing) 27*, 5, 1459–1468.

CHANG, W., AND ZWICKER, M. 2011. Global registration of dynamic range scans for articulated model reconstruction. *ACM Trans. Graph. 30*, 3 (May), 26:1–26:15.

CHEN, Y., AND MEDIONI, G. 1992. Object modelling by registration of multiple range images. *Image Vision Comput. 10*, 3 (Apr.), 145–155.

CHUI, H., AND RANGARAJAN, A. 2003. A new point matching algorithm for non-rigid registration. *Comput. Vis. Image Underst. 89*, 2-3 (Feb.), 114–141.

DE AGUIAR, E., STOLL, C., THEOBALT, C., AHMED, N., SEIDEL, H.-P., AND THRUN, S. 2008. Performance capture from sparse multi-view video. In *ACM SIGGRAPH 2008 papers*, ACM, New York, NY, USA, SIGGRAPH '08, 98:1–98:10.

HORN, B. K. P. 1987. Closed-form solution of absolute orientation using unit quaternions. *J. Opt. Soc. Am. A 4*, 4 (Apr), 629–642.

HUANG, Q.-X., ADAMS, B., WICKE, M., AND GUIBAS, L. J. 2008. Non-rigid registration under isometric deformations. In *Proceedings of the Symposium on Geometry Processing*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, SGP '08, 1449–1457.

IZADI, S., KIM, D., HILLIGES, O., MOLYNEAUX, D., NEWCOMBE, R., KOHLI, P., SHOTTON, J., HODGES, S., FREEMAN, D., DAVISON, A., AND FITZGIBBON, A. 2011. Kinectfusion: Real-time 3d reconstruction and interaction using a moving depth camera. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, 559–568.

KAZHDAN, M., BOLITHO, M., AND HOPPE, H. 2006. Poisson surface reconstruction. In *Proceedings of the Fourth Eurographics Symposium on Geometry Processing*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, SGP '06, 61–70.

LI, H., SUMNER, R. W., AND PAULY, M. 2008. Global correspondence optimization for non-rigid registration of depth scans. In *Proceedings of the Symposium on Geometry Processing*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, SGP '08, 1421–1430.

LI, H., ADAMS, B., GUIBAS, L. J., AND PAULY, M. 2009. Robust single-view geometry and motion reconstruction. In *ACM SIGGRAPH Asia 2009 papers*, 175:1–175:10.

LI, H., VOUGA, E., GUDYM, A., LUO, L., BARRON, J. T., AND GUSEV, G. 2013. 3d self-portraits. *ACM Transactions on Graphics (Proceedings SIGGRAPH Asia 2013) 32*, 6 (November).

MITRA, N. J., FLORY, S., OVSJANIKOV, M., GELFAND, N., GUIBAS, L., AND POTTMANN, H. 2007. Dynamic geometry registration. In *Proceedings of the fifth Eurographics symposium on Geometry processing*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, SGP '07, 173–182.

MUJA, M., AND LOWE, D. G. 2009. Fast approximate nearest neighbors with automatic algorithm configuration. In *International Conference on Computer Vision Theory and Application VISSAPP'09)*, INSTICC Press, 331–340.

PEKELNY, Y., AND GOTSMAN, C. 2008. Articulated object reconstruction and markerless motion capture from depth video. *Computer Graphics Forum 27*, 2 (Apr.), 399–408.

SHARF, A., ALCANTARA, D. A., LEWINER, T., GREIF, C., SHEFFER, A., AMENTA, N., AND COHEN-OR, D. 2008. Space-time surface reconstruction using incompressible flow. In *ACM SIGGRAPH Asia 2008 papers*, ACM, New York, NY, USA, SIGGRAPH Asia '08, 110:1–110:10.

SUMNER, R. W., SCHMID, J., AND PAULY, M. 2007. Embedded deformation for shape manipulation. In *ACM SIGGRAPH 2007 Papers*, ACM, New York, NY, USA, SIGGRAPH '07.

SÜSSMUTH, J., WINTER, M., AND GREINER, G. 2008. Reconstructing animated meshes from time-varying point clouds. In *Proceedings of the Symposium on Geometry Processing*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, SGP '08, 1469–1476.

TEVS, A., BERNER, A., WAND, M., IHRKE, I., BOKELOH, M., KERBER, J., AND SEIDEL, H.-P. 2012. Animation cartography - intrinsic reconstruction of shape and motion. *ACM Trans. Graph.*, 12:1–12:15.

WAND, M., JENKE, P., HUANG, Q., BOKELOH, M., GUIBAS, L., AND SCHILLING, A. 2007. Reconstruction of deforming geometry from time-varying point clouds. In *Proceedings of the fifth Eurographics symposium on Geometry processing*, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, SGP '07, 49–58.

WAND, M., ADAMS, B., OVSJANIKOV, M., BERNER, A., BOKELOH, M., JENKE, P., GUIBAS, L., SEIDEL, H.-P., AND SCHILLING, A. 2009. Efficient reconstruction of nonrigid shape and motion from real-time 3d scanner data. *ACM Trans. Graph. 28*, 2 (May), 15:1–15:15.