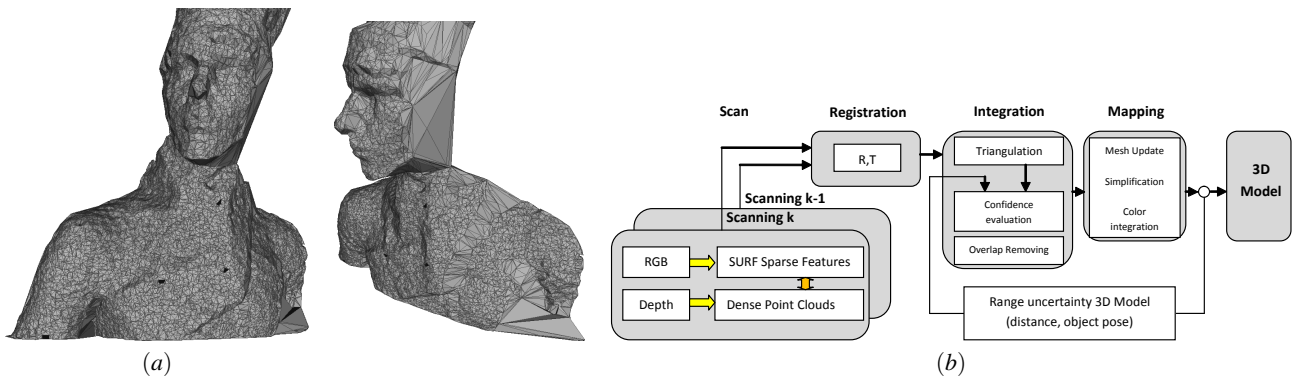


# 3D Modelling framework: an incremental approach

Luis Almeida<sup>1,2</sup> and Paulo Menezes<sup>1</sup> and Jorge Dias<sup>1</sup>

<sup>1</sup>Institute of Systems and Robotics, University of Coimbra, Portugal. <sup>2</sup>Institute Polytechnic of Tomar, Portugal.



**Figure 1:** (a) Mesh model using Crust triangulation. (b) The real-time 3D reconstruction system combines visual features and shape-based alignment between consecutive point clouds while the mesh model representation is updated incrementally using a new Crust based algorithm.

## Abstract

This paper presents a framework for on-line incremental 3D modeling useful for human computer interaction or telepresence applications. We aim a free viewpoint approach based on user's realistic representation to simulate a real face-to-face meeting. Our contribution includes a new adaptation of the Crust algorithm for incremental reconstruction purposes and, a confidence method that evaluates the fusion of new data into the reconstructed model, based on measure uncertainty and novelty. With depth and image information of a single RGB-D sensor, we incrementally reconstruct a mesh model by combining visual features and shape-based alignment.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality I.3.8 Applications,

## 1. Introduction

Realistic representations can enhance the presence feeling and immersion in netmeeting or interaction scenarios. Real-time full body 3D reconstructions based only on video cameras array present some problems due to the lack of accuracy in low-texture or repeated pattern regions, requiring high computational setups and are unsuitability for domestic use. A solution, is to perform RGB-D reconstructions involving alignment and integration of 3D data based on SLAM sparse methods [MBPF12].

Our proposal is a real-time 3D full reconstruction system that combines image invariant features and shape-based alignment between point clouds while the mesh model representation is updated incrementally using a new Crust based algorithm. The alignment differs from the majority of related works as it is based on a closed-form

solution. The paper is organized as follows: section 2 details the proposed reconstruction methodology, section 3 presents some experimental results and section 4, the conclusions and future work.

## 2. Mesh Modeling

We propose a modification of the Crust algorithm to incrementally integrate new acquired tridimensional points without recalculation of previous meshes. The mesh stitching procedure integrates the new mesh poles as new vertices and during the triangulation process, we only compute triangles in joint boundaries where the two surfaces share vertices. It is possible to determine a surface mesh  $S$  of an object from a set of registered 3D point  $X \in R^3$ . Our incremental adaptation of the Crust algorithm [ABK98], uses as input a set of points  $X$  and computes a set of Voronoi poles  $P$  that lie on

the surface medial axis. Then Delaunay triangulation of  $X \cup P$  is performed and the final mesh surface results from a subset of simplices whose vertices are with  $X$ . As adding vertices to Delaunay triangulation isn't costly we introduce an incremental characteristic to the Crust algorithm. Given that a set of points  $X_t$  were already processed by the Crust algorithm, the set of Voronoi poles  $P_t$  and the Delaunay triangulation of  $X_t \cup P_t$  are already available. To integrate a new set of sample points  $X_{t+1}$  to the surface mesh, without a complete mesh recalculation, we follow algorithm 1:

---

**Algorithm 1** Crust incremental algorithm
 

---

- 1:  $\mathbf{P}_{t+1}$  = poles of  $\mathbf{X}_{t+1}$
  - 2: Add  $\mathbf{P}_{t+1} \cup \mathbf{X}_{t+1}$  as new Delaunay triangulation vertices
  - 3: Extract triangles whose vertices belong to  $\mathbf{X}_t \cup \mathbf{X}_{t+1}$
- 

Figure 1(a) illustrates a mesh model obtained by the Crust approach and Figure 1(b) presents an overview of the algorithm: *multiview 3D scan, correspondence, registration, model mapping and integration*.

**Multiview 3D Scan:** The system acquires 3D data and RGB data using a commodity depth sensor device (e.g. Kinect).

**Correspondence:** To solve this correspondence problem, we take advantage of the fact that RGB-D sensor provides simultaneously scene 3D information and respective 2D image. We propose the use of "Robust Image Features" (*Speeded Up Robust Features* (SURF)), which enables the identification of one same point in consecutive images. The association of a visual feature with its 3D point, enables to establish a match between consecutive 3D point clouds. RANSAC algorithm removes false correspondent point pairs that wrongly biases the rigid body transformation estimation.

**Registration:** The registration process determine the optimal transformation to align several 3D point clouds into one same referential to create a global model. Considerer the existence of two corresponding 3D points sets  $\{\mathbf{x}_i^t\}$  and  $\{\mathbf{x}_i^{t+1}\}$ ,  $i = 1..N$ , from consecutive  $t$  and  $t + 1$  scans, which relationship is given by equation (1):

$$\mathbf{x}_i^{t+1} = \mathbf{R}\mathbf{x}_i^t + \mathbf{t} + \mathbf{v}_i \quad \epsilon^2 = \sum_{i=1}^N \left\| \mathbf{x}_i^{t+1} - \mathbf{R}\mathbf{x}_i^t - \mathbf{t} \right\|^2 \quad (1) \quad (2)$$

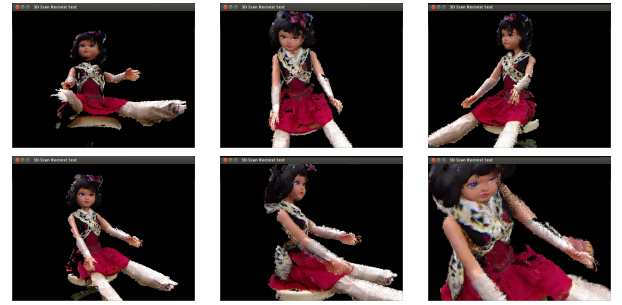
$\mathbf{R}$  represents a standard 3x3 rotation matrix,  $\mathbf{t}$  stands for a 3D translation vector, and  $\mathbf{v}_i$  is a noise vector. The optimal transformation  $\mathbf{R}$  and  $\mathbf{t}$  that maps the set  $\{\mathbf{x}_i^t\}$  on to  $\{\mathbf{x}_i^{t+1}\}$  can be obtained through the minimization of the equation (2) using a least square criterion. The singular value decomposition (SVD) of a matrix is used to minimize Eq. (2) and obtain the rotation (standard orthonormal 3x3 matrix) and the translation (3D vector) [AMD15].

**Model Mapping:** Suppose that the mapping from the world coordinates to one of the scans of the sequence, is known (ex: scan 0) and it is defined by the transformation  ${}^0\mathbf{T}_w$ . Thus, for any consecutive pair of scans ( $t, t+1$ ) from tracked points it is possible to estimate rotation and translation and combine them into a single homogeneous matrix 4x4,  ${}^{t+1}\mathbf{T}_t$ ,  $\mathbf{T} = [\mathbf{R}, \mathbf{t}]$ . Then, compute equation  ${}^i\mathbf{T}_0 = {}^i\mathbf{T}_{i-1} {}^{i-1}\mathbf{T}_{i-2} \dots {}^1\mathbf{T}_0$  and  ${}^i\mathbf{T}_w = {}^i\mathbf{T}_0 {}^0\mathbf{T}_w$  to update the reconstructed model. Deformable bodies can be segmented into parts and be treated as a set of rigid transformations.

**Integration:** To choose which information is relevant, we evaluate the data based on the uncertainty of range sensor. Sensor accuracy measures are dependent on the incident angle between the measuring ray and the surface distance. We associate to each triangle a confidence value based on the measure uncertainty of its 3D vertices:  $C_i = \left| \frac{1}{L\theta} \right|$  where,  $L$  is the distance between a 3D point and the range sensor's optical center and  $\theta$  represents the sensor's pose angle in relation to the surface. The overlapping region is determined by projecting the pre-built mesh vertices into the sensor 2D plane, once transformed for the referential of the newly scanned vertices and by checking the intersection area.

### 3. Results

Figure 2 shows a reconstructed 3D model. It results from several 3D point clouds fused in real time after applying successive 3D rigid body transformations, mesh refining integration and rendering.



**Figure 2:** Synthesized views of a on-line 3D reconstructed model dependent of observer point of view.

### 4. Conclusion

We presented a framework for on-line incremental 3D modeling useful for human computer interaction. Our approach explores virtual view synthesis through motion body estimation and a low cost RGB-D sensor device. Contributions include a new incremental version of Crust algorithm that efficiently adds new vertices to an already existing surface without having to recompute previous generated meshes, and integration of 3D data based on confidence measures that avoids redundant data information computation. Future work includes usability tests for telepresence meeting scenarios.

### References

- [ABK98] AMENTA N., BERN M., KAMVYSSELIS M.: A new voronoi-based surface reconstruction algorithm. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '98. ACM, New York, NY, USA, 1998, pp. 415–421. 1
- [AMD15] ALMEIDA L., MENEZES P., DIAS J.: Incremental reconstruction approach for telepresence or ar applications. In *SciTecIN'15 - Sciences and Technologies of Interaction, EPCGI'2015: The 22nd Portuguese Conf. on Computer Graphics and Interaction*. Instituto de Sistemas e Robótica, Coimbra, Portugal, 2015, pp. 47–54. 2
- [MBPF12] MAIMONE A., BIDWELL J., PENG K., FUCHS H.: Augmented reality: Enhanced personal autostereoscopic telepresence system using commodity depth cameras. *Comput. Graph.* 36, 7 (Nov. 2012), 791–807. 1