Topology-aware Reconstruction of Thin Tubular Structures

Tobias Martin[1] Juan Montes[2] Jean-Charles Bazin[1] Tiberiu Popa[2]
[1]ETH Zürich [2]Concordia University

Figure 1: *Given a sequence of color images of a tubular structure as input, our method computes the geometry of the tubular structure. While visual hull and visual structure-from-motion techniques can provide an estimate of the geometry, our method returns, in addition to the geometry, the topological information in the form of a 1D path.*

Abstract

This paper is dedicated to the 3D reconstruction of thin tubular structures, such as cables or ropes, from a given image sequence. This is a challenging task, mainly because of self-occlusions of the structure and its thin features. We present an approach that combines image processing tools with physics simulation to faithfully reconstruct jumbled and tangled cables in 3D. Our method estimates the topology of the tubular object in the form of a single 1D path and also computes a topology-aware reconstruction of its geometry. We evaluate our method on both synthetic and real datasets and demonstrate that our method favourably compares to state-of-the-art methods.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Curve, surface, solid, and object representations

Keywords: tubular structures, surface reconstruction, topology

1 Introduction

Tubular structures occur in a variety of instances, such as electric cables, fire and garden hoses, and ropes, among many others. While these examples share the property of being a deformed tube, the way they are arranged to themselves, i.e., the way they bend, overlap, twist, or self-occlude, is different. It is thus difficult to correctly reconstruct the geometry of these varying topological arrangements.

While the field of 3D surface reconstruction has made impressive progress over the last few years [Wu 2013; Wu et al. 2011], conventional reconstruction methods are challenged in this context as tubular objects can be relatively thin. Video based reconstruction methods such as structure from motion provide limited quality even when many images are being accumulated as shown in Figure 1. Emerging color and depth cameras such as the Kinect device have paved the way for a more detailed reconstruction compared to conventional color cameras [Izadi et al. 2011; Zhou and Koltun 2013]. However, while significant progress has been made, the depth quality of current sensors is not sufficient to reconstruct thin features. Moreover, an important limitation common to all methods mentioned is that they do not take the topology of the object into consideration. However, understanding the topology is an important prior in the reconstruction process as illustrated in Figure 1. It shows that our method is able to compute a topology-aware reconstruction of a tubular object from a given set of images.

In this paper we propose a reconstruction method that combines image processing techniques with physics simulation for high fidelity reconstruction of thin, tubular objects. Our method robustly fuses the data from multiple images to identify the different segments of the cable as well as their 3D topological structure. While this can be achieved using fairly standard image processing techniques for individual images, robust fusion of the 2D information from many images that yields correct, consistent and complete 3D reconstruction is a challenging problem. Our method connects the segments via visual information and establishes the cable topology. Finally, it applies physics simulation for a complete, accurate and topology-aware reconstruction.

2 Related Work

The problem of reconstructing a surface from points, images, volumetric data, etc., has received major attention in graphics and vision over the past two decades. Most popular methods fall into one of the following main categories: structure from motion [Wu 2013; Wu et al. 2011], structured light [Valkenburg and McIvor 1998], volumetric fusion [Curless and Levoy 1996], and point-based methods [Berger et al. 2014]. While these methods are general and able to produce high-fidelity 3D surfaces, they are challenged when they are used to reconstruct thin features. This is due to the lack of a sufficient sampling of the thin structures, making it difficult to reconstruct a topologically correct surface. For instance, in KinectFusion [Izadi et al. 2011] and its derivatives, e.g., [Zhou and Koltun 2013], the quality of the reconstructed objects depends on depth map quality, in terms of image resolution and quantization of the error pattern.

An additional challenge is that tubular structures often consist of self-occlusions, as the sensor cannot acquire images from every view on the object. This makes reconstruction of a watertight object

SIGGRAPH Asia 2014, December 03 – 06, 2014, Shenzhen, China.

ACM 978-1-4503-2895-1/14/12
http://dx.doi.org/10.1145/2669024.2669035

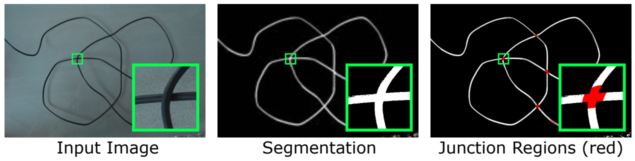

Figure 2: *Segmentation of the input image (left) into cable and non-cable pixels (middle), and identification of the junction regions.*

difficult. Because of these challenges, application specific methods have been proposed. For instance, [Hu et al. 2014] propose a system to reconstruct strands of hair from an input point set acquired from a multi-view stereo setup. The method proposed in [Livny et al. 2010] employs a series of global optimizations to consolidate a point cloud representing one or more tree objects into skeletal structures, and uses a graph-based approach to reconstruct tree branches. However, tree structures have a simpler topology compared to more general graphs, thus, their optimization requires a relatively dense set of points. Similarly, the method of [Tagliasacchi et al. 2009] extracts the medial axis from a point cloud and uses it to reconstruct the surface in a topologically correct way. [Li et al. 2010] introduce a new 1D primitive, called arterial snake, that is used to reconstruct 1D structures such as rods. Unfortunately, all these methods require a relatively dense set of sample points, much denser than can be obtained from one moving camera using state-of-the-art structure from motion algorithms, as demonstrated in the experiments section.

3 Proposed Approach

Our method takes as input a sequence of color images of a tubular structure. Our reconstruction pipeline consists of the following steps:

1. Segment the tubular structure pixels and identify the junction regions in the 2D images (Section 3.1).
2. Fuse images and junction information from the 2D images into a 3D occupancy grid (Section 3.2).
3. Estimate 1D curve skeleton segments that connect junction regions or end-points in 3D (Section 3.3).
4. Reconstruct the geometry of the tube segments combining the visual information from the images with rod physics (Section 3.4).
5. Join the 1D segments to establish the topology of the tubular structure (Section 3.5).
6. Reconstruct the final geometry of the tubular structure (Section 3.6).

3.1 2D Segmentation and Junction Regions

In the following we describe how to segment the 2D images and identify junction regions. Each image is converted to HSV colorspace and then thresholded for one frame of the sequence with an intuitive slider. All the other images of the sequence are then automatically processed. Pixels corresponding to the tubular object are refered to as cable pixels and are marked white, otherwise they are refered to as non-cable pixels and marked black (Figure 2-middle). Junction regions are collections of pixels where the cable overlaps with itself. Intuitively, these regions contain more cable than non-cable pixels. Hence, we detect these regions by making the assumption that a pixel belongs to a junction region when the number of cable pixels in a local window around this pixel is larger than in a window around

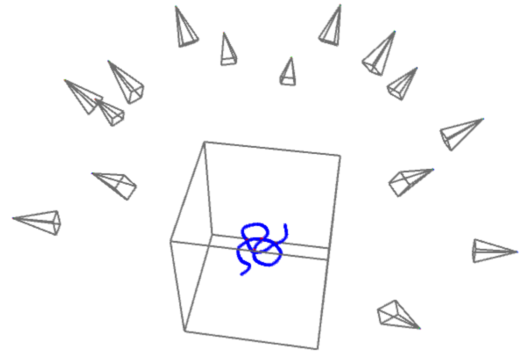

Figure 3: *Images from the cameras are embedded into a 3D volumetric grid (only a subset of the cameras are shown). Voxels marked in blue are occupied by the tubular structure.*

a pixel away from the junction region. If that number exceeds a threshold, the pixel is considered a junction pixel (Figure 2-right). While this is a relatively simple approach, it is efficient to compute and works well for all the examples shown in this paper.

3.2 3D Occupancy Grid Construction

We apply VisualSFM [Wu 2013; Wu et al. 2011], based on structure from motion, to the input image sequence to estimate the intrinsic and extrinsic parameters of the camera. The visual information from all the images is fused by embedding the scene into a regular 3D grid placed in the area of interest. Grid voxels are labeled as inside the cable and marked as occupied if they project into the segmented areas in at least half of the images. Same is applied to junction cells. Figure 3 illustrates the grid resulting from this classification.

3.3 3D Segments Extraction

Although the 3D occupancy grid provides some rough information regarding the location and geometry of the tubular structure, it has no knowledge of its topology. Our method extracts 1D curves that connect two junctions or a junction to an end-point (segments) by traversing through the occupied regions of the grid. Although junctions are identified, extracting the segments is non-trivial because there can be several segments joining the same two junctions, or there could be loops. To solve this problem, it is important to not only identify the junction regions (Figure 4a), but also to determine how many segments emanate from a given junction (Figure 4b). For this, we grow the front of a junction region (the front is the grid nodes that are on the boundary of the junction region) until they are no longer connected. The number of components of the subgraph induced by the front nodes is the number of segments emanating from a given junction region. Once these segments are identified, we proceed to grow these fronts until they connect with other fronts thus identifying the segments as shown in Figure 4c.

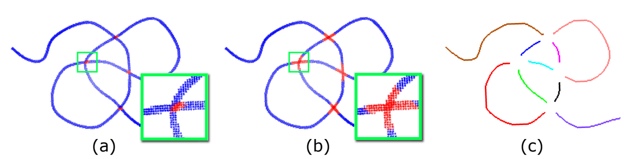

Figure 4: *Segmentation of the tubular structure in the 3D grid. (a) 3D junction regions (red) computed from the images. (b) 3D junction regions after growing. (c) Extracted 3D segments.*

3.4 3D Segments Reconstruction

The geometry of the tube is computed by sweeping a circle along the reconstructed segments. However, while the segments identified in the previous step provide accurate topological information regarding the tube, they are generally not geometrically accurate (Figure 5b). We improve accuracy by executing a physics based rod simulation based on [Bergou et al. 2008]. The external forces of the simulation are computed based on the occupancy grid. Each vertex in the occupancy grid which is sufficiently close to the rod exerts a force onto the closest vertex of the rod. This has the effect of naturally placing the rod inside the occupied grid voxels while at the same time pressuring its physical properties. The result is shown in Figure 5c.

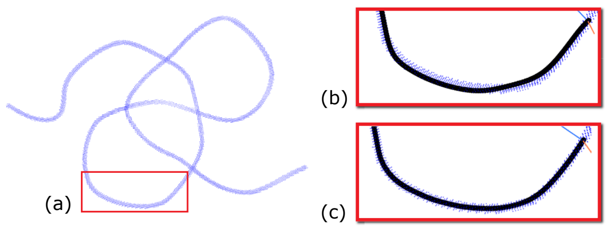

Figure 5: *Top view of the volumetric grid. (a) Reconstruction grid. (b) Reconstructed 3D segments without physics. (c) Reconstructed 3D segments with physics. The physics forces constrain the segments to stay within the occupied regions.*

3.5 3D Segments Connection

The reconstructed tubular segments need to be connected into a single tube. Determining how segment end points are connected to other segment end points at the junctions regions is a combinatorial problem. Once again, we use visual information from the images to solve this. We make the observation that when cables cross, the cable that goes on top has no sharp edges while the cable at the bottom generally exhibits two sharp edges due to the ambient occlusion as shown in Figure 6. Therefore, among the available images, we select one that looks down directly at the junction and we project the segment end-points to the image. Then, in image space, the shortest paths between each pair of end points are computed, using the image gradient as graph weights. The overall shortest path defines the connection between the segments. Finally, the topology of the whole cable is obtained by connecting all the segments into a 1D curve.

One additional problem is that tubular structures such as cables are often made out of plastic material that may exhibit specular highlights as shown in Figure 6. This can interfere with the computations as edges that cross such a highlight will still have a large weight. However, we observe that specular highlights tend to be aligned along the cable, therefore the shortest path is computed by choosing the brightest pixel in a small vicinity of the end points. This method works in most cases. However, in the rare case when it does not work, the system has a further combinatorial check. If the connections are made incorrectly, the result might be two tubular structures instead of one. This case is automatically detected by the system and the user is prompted for high level assistance.

3.6 Reconstruction of the Whole Tubular Structure

Finally, given the computed topology, rod-based physics simulation (see Section 3.4) is now performed on the entire structure, which smoothes out wriggles at the connection points. This provides the topology-aware geometry reconstruction (Figures 8 and 9).

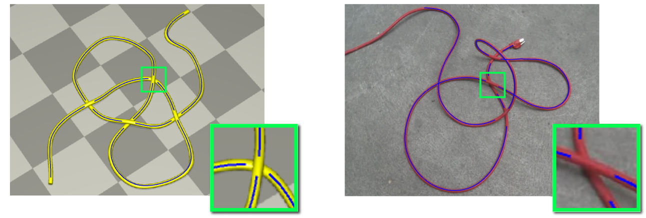

Figure 6: *Junction regions and segments (blue) for synthetic data (left) and video data (right).*

4 Results and Discussion

We demonstrate our method on synthetic datasets as well as datasets captured from real data to evaluate the accuracy of our method. The synthetic datasets are generated using physics simulation and a set of virtual views. Figure 8 shows two synthetic datasets. We overlay the simulated cable (blue) and the result from our reconstruction (red) to illustrate the accuracy of our method. For quantitative evaluation, we also measure the distance between the simulated cable and our reconstruction and we obtain a sub-millimetre precision. Furthermore, Figure 8-left illustrates an example where the reconstructed cable has a clear 3D shape that cannot be captured from a single image. Figure 8-right shows a more complex cable topology that is correctly reconstructed by our method. Figure 9 shows two real datasets acquired with a regular video camera. For both datasets, the complex topology is correctly reconstructed. In addition, the second dataset in Figure 9-right shows a reconstruction of a cable with more complex 3D geometry that cannot be captured by a single image, as deducing such 3D information requires multiple images. The geometry of the bench that supports the cable was created by hand. For additional comparison, we apply the state-of-the-art skeleton extraction method [Huang et al. 2013] to the data used in Figure 1 and Figure 9-left. Figure 7 shows the resulting output. It can be seen that overall the method [Huang et al. 2013] produces 1D curve skeletons of high quality. However, it does not capture the correct topology everywhere, contrary to our method, as indicated in the figure. In the accompanying video we provide more views of the reconstructed cables as well as side by side comparisons between the real video footage and our reconstruction.

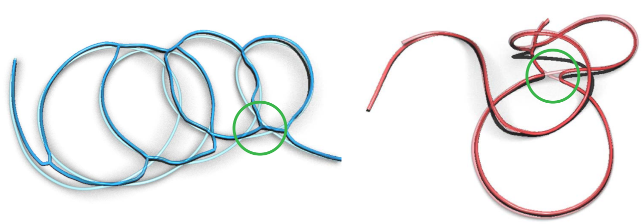

Figure 7: *Comparison with 1D curve skeletons computed using the L1-medial Skeleton method [Huang et al. 2013]. As highlighted in green, some parts in the data are not captured by that method. Our result is superimposed (in light colors) for an easier comparison.*

5 Conclusion

In this paper we propose a method that reconstructs the geometry of tubular structures and establishes their topology from a given set of input images. First, for each image we reconstruct the 2D topology of the tubular structure identifying the segments and the junctions. This is achieved by a low-level image analysis. Then, the topological information from all images is fused into a volumetric

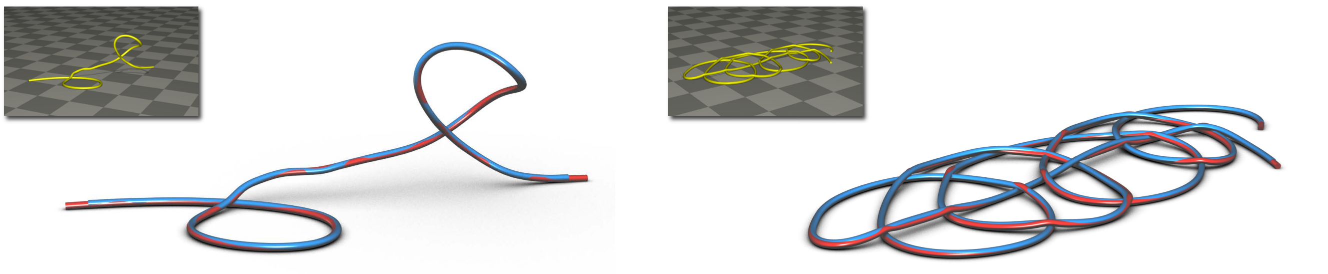

Figure 8: *Results of our method applied to two synthetic examples. To illustrate the accuracy of our reconstruction, we render our result (red) and the ground truth (blue) together in the same view.*

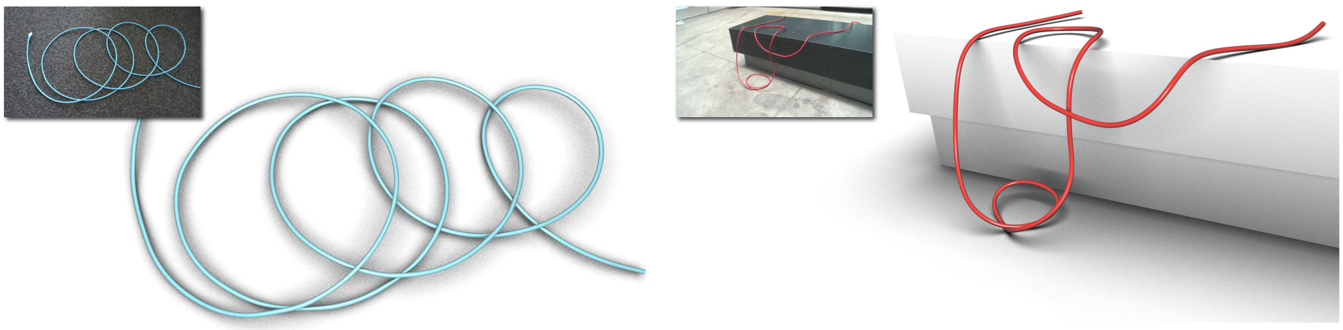

Figure 9: *Results of our method applied to two real datasets acquired using a conventional video camera.*

grid which is the basis for the reconstruction of the 3D structure. The final reconstruction is obtained by physics simulation. Our method is a first step towards robust reconstruction of tubular objects from images, as those structures are notoriously difficult to reconstruct due to their thin geometry and complex topology. While many tubular structures such as cables have uniform color that can be used to solve the topological inconsistencies, certain objects may exhibit a strong texture. These cases are left for future work. As shown in the results section, our method performs robustly if the segments are clearly visible in at least some views. We would like to overcome this limitation using a more robust analysis of the images and a better handle of occluded areas.

References

BERGER, M., TAGLIASACCHI, A., SEVERSKY, L. M., ALLIEZ, P., LEVINE, J. A., SHARF, A., AND SILVA, C. 2014. State of the art in surface reconstruction from point clouds. *Eurographics STAR (Proc. of EG'14)*.

BERGOU, M., WARDETZKY, M., ROBINSON, S., AUDOLY, B., AND GRINSPUN, E. 2008. Discrete elastic rods. *ACM Transactions on Graphics (SIGGRAPH) 27*, 3.

CURLESS, B., AND LEVOY, M. 1996. A volumetric method for building complex models from range images. In *SIGGRAPH*, 303–312.

HU, L., MA, C., LUO, L., AND LI, H. 2014. Robust hair capture using simulated examples. *ACM Transactions on Graphics (SIGGRAPH) 33*, 4.

HUANG, H., WU, S., COHEN-OR, D., GONG, M., ZHANG, H., LI, G., AND CHEN, B. 2013. L1-medial skeleton of point cloud. *ACM Transactions on Graphics 32*, 4.

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 *ACM UIST*, 559–568.

LI, G., LIU, L., ZHENG, H., AND MITRA, N. J. 2010. Analysis, reconstruction and manipulation using arterial snakes. *ACM Transactions on Graphics (SIGGRAPH Asia) 29*, 6.

LIVNY, Y., YAN, F., OLSON, M., CHEN, B., ZHANG, H., AND EL-SANA, J. 2010. Automatic reconstruction of tree skeletal structures from point clouds. *ACM Transactions on Graphics (SIGGRAPH Asia) 29*, 6.

TAGLIASACCHI, A., ZHANG, H., AND COHEN-OR, D. 2009. Curve skeleton extraction from incomplete point cloud. *ACM Transactions on Graphics (SIGGRAPH) 28*, 3.

VALKENBURG, R., AND MCIVOR, A. 1998. Accurate 3D measurement using a structured light system. *Image and Vision Computing 16*, 2, 99 – 110.

WU, C., AGARWAL, S., CURLESS, B., AND SEITZ, S. 2011. Multicore bundle adjustment. In *IEEE CVPR*, 3057–3064.

WU, C. 2013. Towards linear-time incremental structure from motion. In *International Conference on 3D Vision*, 127–134.

ZHOU, Q.-Y., AND KOLTUN, V. 2013. Dense scene reconstruction with points of interest. *ACM Transactions on Graphics (SIGGRAPH) 32*, 4.