Accurate 3D Footwear Impression Recovery From Photographs

Fernanda A. Andaló*, Fatih Calakli+, Gabriel Taubin+, and Siome Goldenstein*

* Institute of Computing, University of Campinas (Unicamp), Campinas, SP, Brazil
+ School of Engineering, Brown Univesity, Providence, RI, USA

Keywords: Computational Forensics, Footwear impression recovery, Shoe print, Multiview stereo.

Abstract

The recovery of footwear impression in crime scenes plays an important role in investigations to corroborate or refute information, or to narrow down the number of suspects. Casting 3D footwear impressions is a long-standing standard to obtain the 3D models of the prints, slowly being replaced by a less invasive method, 3D scanning. In this paper, we present an alternative method based on multiview stereo that yields an accurate 3D model and provides some benefits over existing methods. We evaluate the results comparing our reconstructed 3D models with the ones acquired by 3D scanning. We also examine the advantages and drawbacks of each method.

1 Introduction

An efficient crime investigation depends on the collection and analysis of various kinds of evidence – items or information gathered at the crime scene, or related locations, that are relevant to the investigation – which include DNA, tire tracks, fingerprints, shoe prints, bloodstains, among others.

Impression evidence, such as footprints, tire tracks, and tool marks, are an important and common source of physical evidence that can be used to corroborate or refute information provided by witnesses or suspects. According to a study conducted in Switzerland, shoe prints can be found in approximately 35% of all crime scenes [1].

Shoe prints can indicate whether a person was walking or running, was carrying something heavy or was unfamiliar with the area or unsure of the terrain [2]. They can provide additional information about the wearer, such as weight, height, and wear patterns that can be compared with a suspect's shoes. The location of the impressions at the scene can also often help in the reconstruction of the crime [1].

Shoe prints can be classified in three categories, based on how they are found at the crime scene: patent, plastic, or latent [3]. Patent shoe prints are those that are clearly visible at the crime scene; plastic or three-dimensional (3D) prints occur when the shoe sinks in the material that is being stepped on, leaving marks; and latent prints are invisible to the naked eye and need to be exposed using different forensic techniques.

Plastic or 3D footwear impressions have depth in addition to length and width, and are most commonly found outdoors in soil, sand, and snow [1]. The details that can be retained and captured depend on the material texture, composition, and conditions, and these attributes can largely vary.

In recent years, the standard method for capturing these 3D prints is by casting using materials such as dental stone [1] or plaster [2], and photographing the print to provide additional details which are taken into consideration later. The produced cast can be compared with manufactures' shoes [2] or analyzed in search of minutiae that can provide information about the wearer.

Just as shoe prints, each type of evidence requires a different forensic technique to be revealed, captured and analyzed. These techniques have been improving over the last years due to reliability of modern technology and the greater use of computational forensics. For example, pattern recognition and other computational methods can reduce the bias inherent in traditional criminal forensics [4]. In this sense, an ever growing system to collect evidence is 3D scanning. It is useful not only in collecting, but also in organizing evidence and providing an analysis tool.

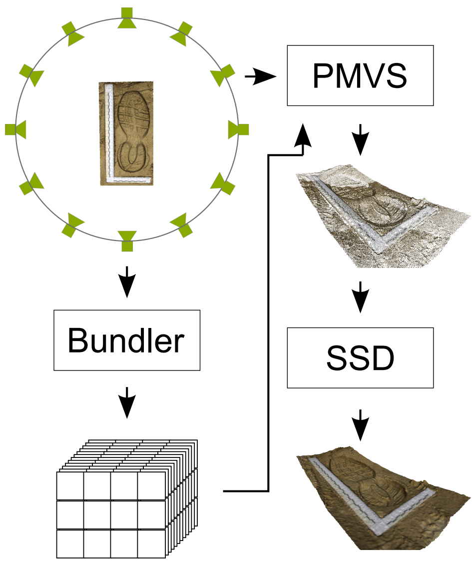
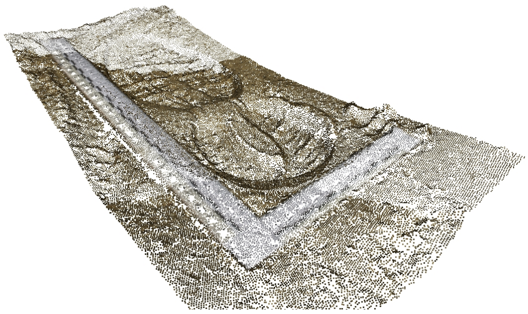
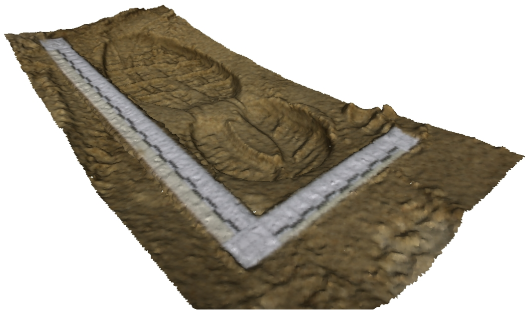


Figure 1. 3D footwear impression recovery pipeline. Photographs taken at different viewpoints are used as input. Bundler recovers camera parameters for each image. PMVS generates a dense point cloud. SSD reconstructs the surface.

(a) Examples of input images.



(b) 3D point cloud.



(c) Reconstructed 3D model.

Figure 2. 3D footwear impression recovery pipeline: (a) capture several photographs of the shoe print with a digital camera; (b) reconstruct a dense point cloud; (c) reconstruct a surface.

In this paper, we propose an alternate solution to the problem of capturing 3D footwear prints at crime scenes. We compare it to the existing solutions: casting and 3D scanning, and we consider their advantages and drawbacks. Our solution includes a pipeline (Figure 1) to obtain the 3D reconstruction using only digital photographs taken from the footwear print at the crime scene. The pipeline consists of three previously proposed methods that together reconstruct a complete 3D model from a collection of images taken at different camera viewpoints (Figure 2(a)). The first step is to recover a set of camera parameters and 3D locations for keypoints in each image using *Bundler* [5], a method proposed to perform structure from motion (SfM) on unordered image collections. The second step is to generate a dense point cloud using *PMVS* [6], a patch-based multiview stereo method (Figure 2(b)). The last step is to reconstruct the surface using a new method called *Smooth Signed Distance* (SSD) [7] (Figure 2(c)).

The main contributions of this paper are the pipeline to obtain 3D models of footwear prints from pictures taken at different angles, and the analysis of the results in respect to accuracy and its advantages over the existing methods.

2 Background on 3D footwear impression recovery

Several decades ago, casting was the main method for recovering 3D footwear impression evidence. Although the impressions were also photographed, the less sophisticated equipment made photography not convenient and often less successful than casting. The casting material at the time, *plaster of Paris*, also induced a time consuming and messy casting procedure [1].

From the 60s to the 80s, photography equipment and film improved, allowing photography to became a much more popular option than casting. Many departments completely discontinued plaster casts. However, in recent years, many quality casting materials and more simplified procedures changed casting into an easier and convenient way of recovering 3D impression evidence [1].

Together with this recent casting trend, High Definition Surveying (HDS), or 3D scanning, became popular for surveying buildings, terrain, and other architectural features in a fast and detailed manner [8]. There are two types of 3D scanners employed in forensics: crime scene scanners that can capture a large overview map of the scene; and close-up 3D scanners that can capture individual objects in full color and high resolution [9]. Tire tracks, footprints, shoe prints, and bones, for example, can be scanned in place with this last type of scanner [8].

The use of 3D scanning over casting is beneficial. Casting, in some cases, can destroy the original evidence in the process, while 3D scanners often use lasers that can scan the object without touching or affecting it. Some scanners also capture the color surface, producing a visually accurate replica that would not be possible with casts. Furthermore, casts are physical objects that are difficult to share across locations and take up physical storage space [9]. It is also important to highlight that creating a complete and highly detailed 3D model of a footprint takes less than 15 minutes, which is 125% faster than traditional methods of casting [9].

Aside from acquisition speed, 3D scanners are particularly well suited for scanning organic shapes and highly curved surfaces that would otherwise be difficult to measure [8]. The main disadvantage is that 3D scanners are not always appropriate to scan all kinds of surfaces and materials. Introducing a new kind of expensive equipment as a standard can also take time and it will not always be available at all locations.

The use of 3D scanners in practice and in works related to forensics is growing together with a field called *Computational Forensics*. CF is an emerging interdisciplinary research domain and it is understood as the hypothesis-driven investigation of a specific forensic problem using computers, with the primary goal of discovery and advancement of forensic knowledge [10].

Several works contribute to the field of Computational Forensics, such as automatic shoe print image retrieval [11, 12], latent palmprint matching [13], estimation of heights of objects and persons in a single image [14], 3D visualization of crime scenes [15], digital image and video forensics [16], and many others. We can see by these works that Computer vision and Forensics are a powerful combination to the recovery and examination of evidences, and it tends to grow as more scientific evaluation is available.

In this work, we provide a new insight for footwear print capturing that uses a Computer vision method, *multiview stereo*, and that has some advantages over the current solutions. Multiview stereo methods require photographs of an object at different camera viewpoints. They compute correspondences between image pairs and a depth estimation for each camera viewpoint. By combining these estimates, multiview stereo methods can provide a final 3D model of the object. A similar approach has been applied successfully to recover dinosaur footprints [17]. Table 1 summarizes the three discussed methodologies: casting, 3D scanning, and multiview stereo.

3 Photo to 3D Pipeline

The proposed pipeline uses as input a set of photographs of the footwear impression, taken in different viewpoints around the evidence (Figure 2(a)), then generates a 3D point cloud (Section 3.1) which is used later to obtain a 3D surface (Section 3.2).

3.1 From photos to point cloud

The goal of multi-view stereo (MVS) is to reconstruct a 3D model from images taken from known camera viewpoints. To learn the viewpoints, we use *Bundler* [5] which takes a set of images as input and accurately estimates the camera viewpoint per each image. Bundler first finds feature points (keypoints) in each input image. Each keypoint is associated with a local descriptor. The method then matches keypoint descriptors between each pair of images, using an approximate nearest neighbors approach, then robustly estimates a fundamental matrix for the pair. After finding a set of consistent matches between each image pair, the method organizes these matches into tracks of connected matching keypoints across multiples images, then recovers a set of camera parameters and a 3D location for each track by minimizing the sum of distances between the projections of each track and its corresponding image features.

We then use the Patch-based MVS (PMVS) [6] algorithm which takes the same set of images and the estimated camera parameters as input and produces 3D models with accuracy nearly comparable with laser scanners [18]. The algorithm is based on the idea of correlating measurements from several images at once to derive 3D surface information. It reconstructs a global 3D model by using all the images available simultaneously.

3.2 Surface reconstruction

PMVS is also able to estimate a surface normal vector associated for each 3D point measurement, and as a result, produce a collection of so-called *oriented* points. Dense oriented point clouds have become a pervasive surface representation [19] due its simplicity and storage efficiency. However, since they do not constitute surfaces, they cannot be used to make certain measurements required by many applications in forensics.

The problem of reconstructing surfaces from oriented points has a long history [20]. Poisson Surface Reconstruction [21] has become a leading contender due to its high quality reconstructions. However, it is reported in a recent benchmark study [22] that this method tends to oversmooth the data. Alternatively, we use Smooth Signed Distance (SSD) Surface Reconstruction [7] which does not suffer from oversmooting, and still produces good quality surfaces. In fact, in many cases presented in [7], SSD constructs surfaces with less error than some prior art methods including Poisson Surface Reconstruction.

In SSD, oriented data points are regarded as samples of a smooth signed distance field f, and the surface S is defined by an implicit equation $S = \{x : f(x) = 0\}$. The implicit function f is estimated by minimizing the following energy

$$E(f) = \sum_{i=1}^{N} f(p_i)^2 + \lambda_1 \sum_{i=1}^{N} \|\nabla f(p_i) - n_i\|^2 + \lambda_2 \int_V \|Hf(x)\|^2 dx \quad (1)$$

where $\nabla f(x)$ is the gradient of $f(x)$, $Hf(x)$ is the Hessian of $f(x)$, $\{(p_1, n_1), \ldots, (p_N, n_N)\}$ are point-normal data pairs, V is a bounding volume, and $\{\lambda_1, \lambda_2\}$ are regularization parameters. A simple finite-difference discretization reduces the problem to solving sparse linear systems of equations.

In our experiments, we efficiently converted oriented point clouds into accurate polygon meshes using SSD reconstruction method (Figure 2(c)).

4 Experimental results

The experiments consisted on the comparison of our proposed pipeline with one of the methods used in practice, 3D scan-

	Casting	3D Scanning	Multiview stereo
Special materials	Dental stone or other quality material	3D scanner	A camera. It can be the same camera already used in crime scene investigations
Time for acquisition	$\sim$ 30min [9]	$\sim$ 15min [9]	$\sim$ 5min
Post acquisition work	Transportation to the analysis location	Depends on the scanner. The model can be ready to be analyzed or it may need to be processed	Pipeline presented in this work
Intrusiveness	It may destroy the evidence	None	None
Good scenarios	The material cannot be too soft	Not all materials are suitable for being scanned	There is no limitation, as long as the examiner can take pictures at different viewpoints
Accuracy in these scenarios	High	High	Medium to high*

Table 1. Comparison between footwear prints recovery. (*) The accuracy is analyzed in Section 4.

ning, with respect to the recovery of various footwear prints in sand. We set up a sandbox with an attached 3D scanner, where we produced footwear marks with four different types of shoes (Figure 3(a)). We used two pieces of equipment in our experiments:

- NextEngine 3D Laser Scanner with a point and texture density on target surface of 150 DPI, and the dimensional accuracy of ± 0.015".

- Digital camera Canon EOS Rebel XSi with 12.2-megapixel sensor and Canon EF-S 18-55mm f/3.5–5.6 IS lenses, without the use of flash and with automatic focus.

We first set up the scanner at a distance of 31.5" from the sand in the box. For each footwear print, we first scanned it and then we took twelve photographs rotating around it (Figure 3(b)), approximately in equal sized angles. For comparison purposes, we also scanned the four shoe soles corresponding to the prints (Figure 3(c)).

The twelve images of each shoe print were used as input to our pipeline, generating complete 3D reconstructions. The scanned shoe prints and our generated 3D models were registered with the corresponding shoe soles using *Meshlab*[1]. To evaluate the quality of the generated 3D models, we adopted the *Haursdoff distance* computed by *Metro tool* [23].

We consider as groundtruth the distance map d_g that represents the computation of *Haursdoff distance* for the vertices of the shoe print scan in relation to the vertices of the shoe sole scan. Note that d_g is subjective to the curvature of the shoe sole while the shoe is not being used and consequently not flat with respect to the ground. In the same fashion, we computed the distance map d_m between our 3D model and the shoe sole scan. The value $(\overline{d_m} - \overline{d_g})$ measures the quality of our model compared to 3D scanning, where $\overline{d_g}$ and $\overline{d_m}$ are the mean values of the distances in d_g and d_m, respectively[2].

Figure 4 shows the two obtained 3D models for the first shoe print (with the 3D scanner and with our pipeline) and also the proposed evaluation. Figures 4(c) and 4(d) illustrate the distance maps d_g and d_m, where red represents the minimum distance and blue represents the maximum distance. Figure 5 shows the same results for the third shoe print.

By analyzing Figures 4 and 5, we can see that our method is able to capture comparable amount of details, and the same fact can also be deduced by analyzing the value $(\overline{d_m} - \overline{d_g})$ computed for each shoe print (Table 2).

Shoeprint #	$\overline{d_g}$	$\overline{d_m}$	$(\overline{d_m} - \overline{d_g})$
1	9.996	10.002	**0.006**
2	8.157	8.660	**0.503**
3	8.715	9.480	**0.765**
4	8.816	9.114	**0.298**

Table 2. Evaluation of obtained 3D model for each shoe print using *Haursdoff distance* in millimeters.

5 Conclusion

In this work, we presented a pipeline to recover footwear impression from crime scenes based on a well known technique in Computer Vision, *multiview stereo*, which has not been consider or analyzed for this kind of application in the literature until now. Despite the simplicity for set up and acquisition, the reconstructed surfaces proved to be comparable with 3D scan-

[1] http://meshlab.sourceforge.net/

[2] In most works that adopt *Haursdoff distance*, researchers consider not the mean but the maximum value of the map. In our work, the maximum value does not represent a good value for comparison, since it always appears outside the print, where we don't expect the method to be accurate. To illustrate this, please refer fo Figure 4(d).

(a) (b) (c)

Figure 3. (a) Sandbox with attached 3D scanner. (b) Twelve photographs taken from the first shoe print. (c) Shoe soles scans.

ning, a high-end technology used in practice, providing accurate 3D models of the shoe prints. A digital camera is the only equipment required to recovery the evidence, which makes the process convenient and fast.

Acknowledgements

This work is primarily supported by CNPq grant 201238/2010-1, with additional funding from NSF grants IIS–0808718, CCF–0729126, and CCF–0915661.

References

[1] W. Bodziak, *Footwear impression evidence: detection, recovery, and examination*. USA: CRC Press, 2nd ed., 2000.

[2] K. Hess and C. Orthmann, *Criminal Investigation*. USA: Delmar Cengage Learning, 9th ed., 2010.

[3] Australian School Innovation in Science, Technology and Mathematics Project, “Impressions – introduction to tracks, footprints and plaster casts.” Course Lecture.

[4] S. Srihari, “Beyond CSI: The Rise of Computational Forensics,” *IEEE Spectrum*, pp. 38–43, 2010.

[5] N. Snavely, S. Seitz, and R. Szeliski, “Photo tourism: exploring photo collections in 3D,” in *ACM Transactions on Graphics*, vol. 25, pp. 835–846, 2006.

[6] Y. Furukawa and J. Ponce, “Accurate, dense, and robust multi-view stereopsis,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 8, 2009.

[7] F. Calakli and G. Taubin, “SSD: Smooth Signed Distance Surface Reconstruction,” *Computer Graphics Forum*, vol. 30, no. 7, 2011. http://mesh.brown.edu/ssd/.

[8] E. Liscio, “A Primer on 3D Scanning in Forensics,” *Forensic Magazine*, 2009.

[9] P. DeLaurentis, “3D Scanning: A New Tool for Cracking Tough Cases,” *Forensic Magazine*, vol. 6, no. 1, pp. 37–40, 2009.

[10] K. Franke and S. Srihari, “Computational forensics: An overview,” *Computational Forensics*, pp. 1–10, 2008.

[11] L. Zhang and N. Allinson, “Automatic shoeprint retrieval system for use in forensic investigations,” in *Proc. of the 2005 UK Workshop on Computational Intelligence*, pp. 137–142, 2005.

[12] G. AlGarni and M. Hamiane, “A novel technique for automatic shoeprint image retrieval,” *Forensic science Inter.*, vol. 181, no. 1-3, pp. 10–14, 2008.

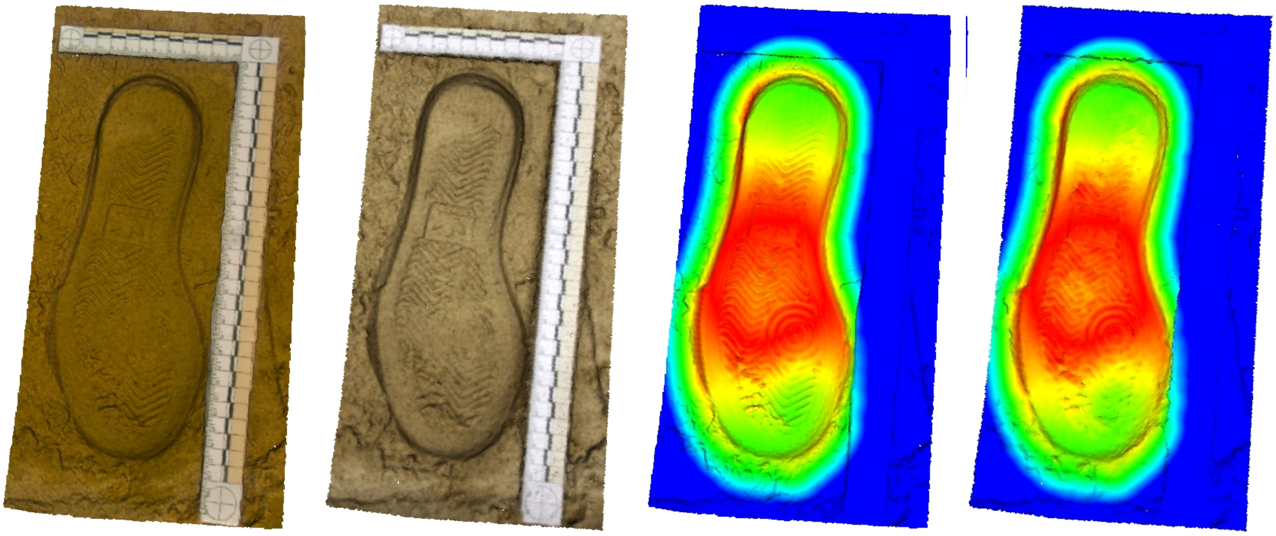
[13] A. Jain and J. Feng, “Latent palmprint matching,” *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1032–1047, 2008.

[14] F. A. Andaló, G. Taubin, and S. Goldenstein, “Detecting vanishing points by segment clustering on the projective plane for single-view photogrammetry,” in *IEEE International Workshop on Information Forensics and Security*, pp. 1–6, 2010.

[15] M. Ma, H. Zheng, and H. Lallie, “Virtual Reality and 3D Animation in Forensic Visualization,” *Journal of forensic sciences*, 2010.

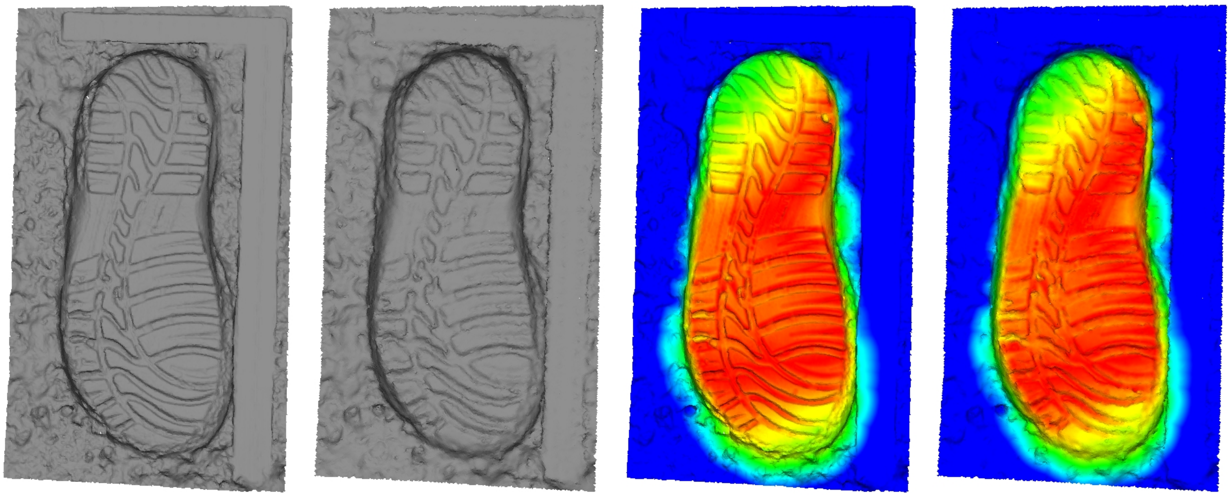
[16] A. Rocha, W. S. T. Boult, and S. Goldenstein, “Vision of the unseen: Current trends and challenges in digital image and video forensics,” *ACM Computing Surveys*, 2011.

[17] F. Remondino, A. Rizzi, S. Girardi, F. Petti, and M. Avanzini, “3D ichnology – recovering digital 3D models of dinosaur footprints,” *The Photogrammetric Record*, vol. 25, no. 131, pp. 266–282, 2010.



(a) Shoe print scan. (b) Our 3D model. (c) Groundtruth d_g. (d) Our result d_m.

Figure 4. Acquired data and evaluation for the first shoe print. In (c) and (d), red represents the minimum distance and blue represents the maximum distance.



(a) Shoe print scan. (b) Our 3D model. (c) Groundtruth d_g. (d) Our result d_m.

Figure 5. Acquired data and evaluation for the third shoe print. The 3D models in (a) and (b) are displayed without texture for better visualization. In (c) and (d), red represents the minimum distance and blue represents the maximum distance.

[18] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in *Proc. of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 519–528, 2006.

[19] L. Kobbelt and M. Botsch, "A Survey of Point-Based Techniques in Computer Graphics," *Computers & Graphics*, vol. 28, no. 6, pp. 801–814, 2004.

[20] O. Schall and M. Samozino, "Surface from scattered points: A brief survey of recent developments," in *1st International Workshop on Semantic Virtual Environments*, pp. 138–147, 2005.

[21] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson surface reconstruction," in *Proc. of the 4th Eurographics Symposium on Geometry Processing*, pp. 61–70, 2006.

[22] M. Berger, J. Levine, L. Nonato, G. Taubin, and C. Silva, "An End-to-End Framework for Evaluating Surface Reconstruction," SCI Technical Report, University of Utah, Jan. 2011.

[23] P. Cignoni, C. Rocchini, and R. Scopigno, "Metro: measuring error on simplified surfaces," in *Computer Graphics Forum*, vol. 17, pp. 167–174, 1998.