Realistic Walkthrough of Cultural Heritage Sites-Hampi

Uma Mudenagudi, Syed Altaf Ganihar, Shreyas Joshi, Shankar Setty, Rahul G., Somashekhar Dhotrad, Meera Natampally, Prem Kalra

B. V. Bhoomaraddi College of Engineering and Technology-Hubli, NIAS Bangalore, IIT-Delhi

Abstract. In this paper we discuss the framework for a realistic walkthrough of cultural heritage sites. The framework includes 3D data acquisition, different data processing steps, coarse to fine 3D reconstruction and rendering to generate realistic walkthrough. Digital preservation of cultural heritage sites is an important area of research since the accessibility of state of the art techniques in computer vision and graphics. We propose a coarse to fine 3D reconstruction of heritage sites using different 3D data acquisition techniques. We have developed geometry based data processing algorithms for 3D data super resolution and hole filling using Riemannian metric tensor and Christoffel symbols as a novel set of features. We generate a walkthrough of the cultural heritage sites using the coarse to fine 3D reconstructed models. We demonstrate the proposed framework using a walkthrough generated for the Vittala Temple at Hampi.

1 Introduction

In this paper we describe the framework for the generation of realistic digital walkthrough of cultural heritage sites. The advent of digital technology has resulted in a great surge in interest to digitally restore heritage sites [1][2]. A large number of cultural heritage sites are deteriorating or being destroyed over a period of time due to natural weathering, natural disasters and wars. The heritage sites at Hampi, India are largely composed of rock structures which are in a grievous situation as can be seen in Fig. 1 and this necessitates the digital preservation of the sites at Hampi. Digital preservation of the heritage sites can be accomplished using modern techniques in computer vision and graphics.

Digital restoration of cultural heritage sites has been in the purview of computer graphics and vision research since a long time. The notable works reported in the literature are Modeling from Reality [3], The Great Buddha Project [2], Stanford University's Michelangelo Project [1], IBM's Pieta Project [4] and Columbia University's French cathedral project [5] to mention a few. Modeling from Reality [3] discusses the modeling of cultural heritage sites in a precise manner using laser range scanners. The Great Buddha Project [2] describes the pipeline for the digital preservation and restoration of Great Buddhas using a pipeline, consisting of acquiring data, aligning data, aligning multiple range images and merging of range images. The Stanford University's Michelangelo Project [1] describes a hardware and software system for digitizing the shape and color of large fragile objects under non-laboratory conditions. Columbia University's French cathedral project [5] describes building of a system which can automatically acquire 3D range scans and 2D images to build 3D models of urban environments.

The acquisition of the 3D data is an integral step in the digital preservation of the cultural heritage sites. The classic 3D modeling tools are often derisory to accurately portray the complex shape of sculptures found at cultural heritage sites. The advent of inexpensive 3D scanning devices like Microsoft Kinect and ToF (Time of Flight) cameras have simplified the 3D data acquisition process. The state of the art 3D laser scanning devices generate very accurate 3D data of the objects. However the scanning of large outdoor objects at the cultural heritage sites invite a lot of tribulations due to the generation of partial meshes. The image based methods like SFM (Structure from Motion) [6] and PMVS (Patch based Multi-View Stereo) [7] consolidate the 3D data acquisition process but do not generate high resolution 3D data to accurately depict the art work at the heritage sites. The occlusions during the scanning process result in the occurrence of missing regions in the 3D data (holes) and generation of partial meshes. This warrants the need for efficient data processing techniques for the digital preservation of the cultural heritage sites.

Fig. 1. The ruins at the Vittala Temple - Hampi, India.

Fig 2 shows the comparison of the rendered scene of the Vittala Temple at Hampi with the original image of the scene. Our framework generates a realizable digital walkthrough of the cultural heritage sites using a coarse to fine 3D reconstruction of the cultural heritages sites. We put into service several data processing algorithms like noise filtering, 3D super resolution, 3D hole filling and texture mapping for the fine level 3D reconstruction of the objects. The fine level 3D reconstructed models at the cultural heritage sites are registered with the coarse level models to generate a coarse to fine 3D reconstructed model. The coarse to fine 3D reconstructed models are subsequently rendered to obtain a

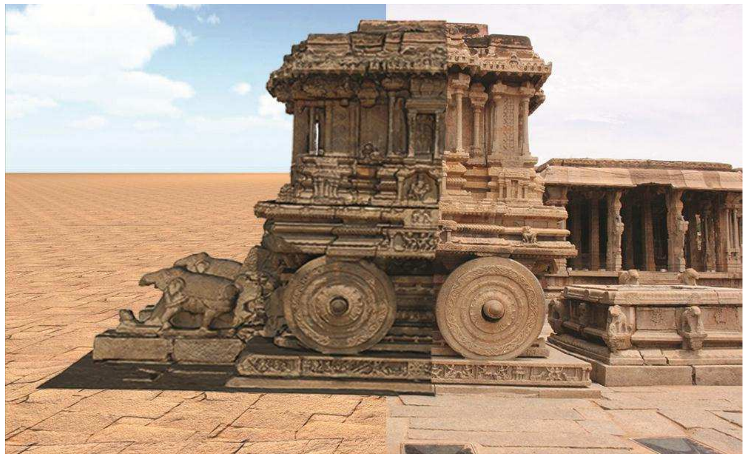


Fig. 2. Comparison of the rendered scene and the original image of the Stone Chariot at Vittala Temple - Hampi, India: Left half is the rendered image and the right half is the original image.

digitally realizable walkthrough of the heritage site. Towards this we make the following contributions:

1. We propose a framework for the generation of realistic walkthrough of cultural heritage sites with coarse to fine 3D reconstruction.
2. We propose 3D super resolution and hole filling algorithms for efficient 3D data processing using concepts of Riemannian geometry with metric tensor and Christoffel symbols as a novel set of features.
3. We demonstrate the proposed framework for Vittala Temple at Hampi, India.

The rest of the paper is organized as follows. In Section 2 we describe the 3D data acquisition techniques employed for the generation of the digital walkthrough. In Section 3 we discuss the 3D super resolution and hole filling algorithms. In Section 4 we explain the coarse to fine 3D reconstruction and rendering of the heritage sites. In Section 5 we demonstrate the results of the proposed framework and provide the conclusion in Section 6.

2 Proposed Framework and Data Acquisition

The proposed framework of coarse to fine 3D reconstruction is as shown in Fig 3. The data acquisition step includes acquisition of 3D data for different modalities like CAD model, Single-view model, Kinect model and Multi-view model. The 3D point cloud data generated during acquisition is fed to data processing. Hole filling and 3D super resolution is performed to refine the point cloud data in the data processing step. In the rendering step, the refined data is fed to coarse to fine 3D reconstruction stage. Finally, rendered view is generated using rendering engine.

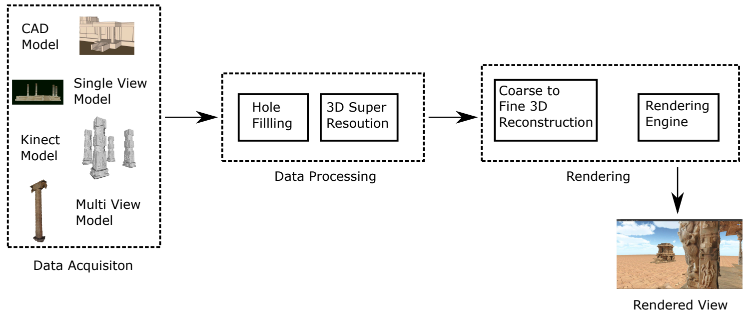


Fig. 3. Framework of coarse to fine 3D reconstruction and rendering to generate realistic walkthrough.

The 3D data acquisition of the cultural heritage sites, is the process of capturing 3D models from the on-site real world objects and is an important part in the digital restoration process. The coarse level models are obtained either using CAD modeling tools or using single view reconstruction. The CAD models obtained do not accurately depict the geometry of the artwork at the cultural heritage sites. The CAD models while modeling are recreated or restored in order to incorporate some of the missing, withered or prophesied part of the cultural heritage site. The CAD models or the single view reconstructed models do not accurately portray the artworks at the cultural heritage sites. The fine level models are hence required to precisely represent the artworks. We acquire the fine level models at the cultural heritage sites in the following ways depending upon the location, size and feasibility of the method.

1. The Microsoft Kinect 3D sensor consisting of a depth and a RGB camera is employed to scan the 3D models. Under appropriate lighting conditions, scanning is done on a 3D model and we use the Kinect Fusion (KinFu) [8] to generate a dense point-cloud or a mesh of the scanned model.
2. A set of images of a object to be reconstructed are captured under appropriate lighting conditions. The images are then fed to dense reconstruction algorithms like SFM [6] or PMVS [7] to generate point cloud models.

3 Data Processing

The data processing algorithms are a vital component in the digital restoration of cultural heritage sites. The obtained data is in the form of a point-cloud which is filtered using Statistical Outliers filter in order to eliminate any noisy data acquired during the scanning process. The data acquired using scanners like laser scanners, Microsoft Kinect or image based methods comprise of certain missing regions (holes), partial meshes or is of low resolution. To address these issues we propose geometry based data processing algorithms for 3D data super resolution and hole filling. The pipeline for the generation of fine level models is shown in Fig 4.

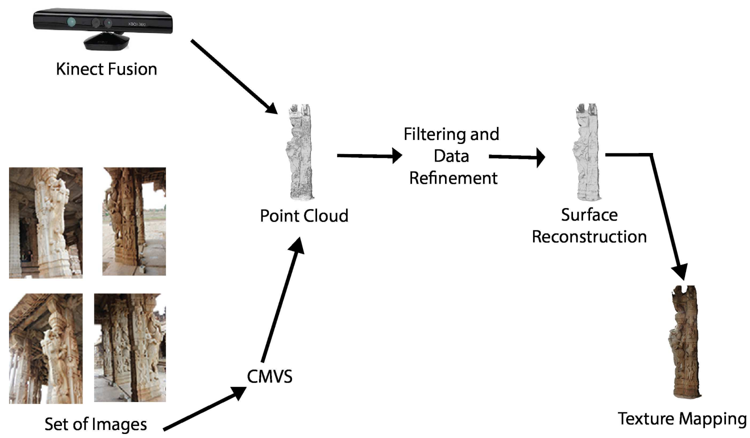


Fig. 4. Overview of the data acquisition and data refinement pipeline for detailed reconstruction.

3.1 Super Resolution

The point-cloud data obtained from the low-resolution 3D scanner like the Microsoft Kinect or from sparse reconstruction algorithms usually fail to capture the accurate geometric properties and detailed structure of the 3D object either due to the presence of occlusions during the scanning process, non-feasibility of the sparse reconstruction algorithm or adverse scanning environment. As a result, these techniques fail to portray all the details in a model's surface resulting in a low-resolution point-cloud data. The generation of high resolution 3D data is important for the realistic rendering of cultural heritage sites. Hence there is an immense requirement to produce a high-resolution point-cloud data from a given low-resolution point-cloud data. Authors in [9] proposed decision framework for super resolution. The decision framework facilitates to obtain the comparatively best fit interpolation curve based on the voting parameters obtained from the point cloud thus producing super-resolved point cloud. However, we propose a learning based super resolution. The overview of the proposed learning based super resolution framework is shown in Fig 5. Given 3D model is modeled as a set of Riemannian manifolds [10–13] in continuous and discretized space. A Kernel based SVM learning framework [14] is employed to decompose a given 3D model into basic shapes viz., sphere, cone and cylinder using metric tensor and Christoffel symbols as a set of novel geometric features. The decomposed models are then independently super-resolved using selective interpolation techniques for example the spherically decomposed model is super resolved using spherical surface interpolation technique. The independently super resolved algorithms are merged to obtain the final super resolved model.

The metric tensor [10–13] $g_{\mu\nu}$ is a symmetric tensor and in 3-dimensions consists of 6 independent components. The metric tensor gives the quantitative measure for the deviation in the manifold from the Euclidean space. The

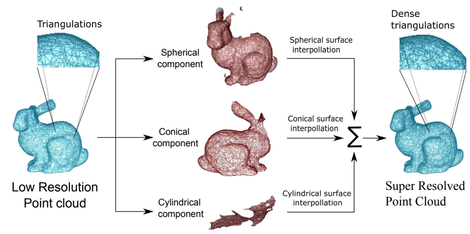


Fig. 5. Proposed learning based super resolution framework.

Christoffel symbols [10–13] give a measure of the deviation of the metric tensor as a function of position. The Christoffel symbols in 3-dimensions consists of 18 independent components.

The features used for the decomposition consist of 24 independent components which are in turn dependent on the geometrical position of the point over which the features are calculated. The decomposition of 3D model into basic shapes is carried out using a SVM [14] framework. The training data consists of unit sphere, cone and cylinder which are learned in the SVM framework. The spherical decomposed part of the 3D model is interpolated using spherical surface interpolation method. Similarly the conical and the cylindrical decomposed parts are interpolated using conical surface interpolation and cylindrical surface interpolation method respectively. The interpolated decomposed parts are then fused to generate a super-resolved point-cloud of the 3D model. The algorithm achieves better result than reported in the literature.

3.2 Hole filling

The 3D data acquired using the proposed techniques consists of missing regions or holes due to occlusions in the surface to be scanned. To address this issue we propose a hole filling algorithm using metric tensor and Christoffel symbols as features. The holes are identified by using the boundary detection algorithm used in [15]. The neighborhood of the hole is decomposed into basic shapes using a kernel based SVM learning framework with metric tensor and Christoffel symbols as features. The overview of the proposed hole filling algorithm is shown in Fig 6. The decomposed regions in the neighborhood of the hole are interpolated using selective surface interpolation techniques. The centroid of the hole region is computed and the selective surface interpolation is carried out along the directional vector.

The point-cloud is surface reconstructed using Poisson surface reconstruction [16] or Ball-pivoting surface reconstruction algorithm [17]. The surface reconstructed model is texture mapped using image alignment with mutual information [18] and parameterization of the registered rasters for the surface reconstructed model.

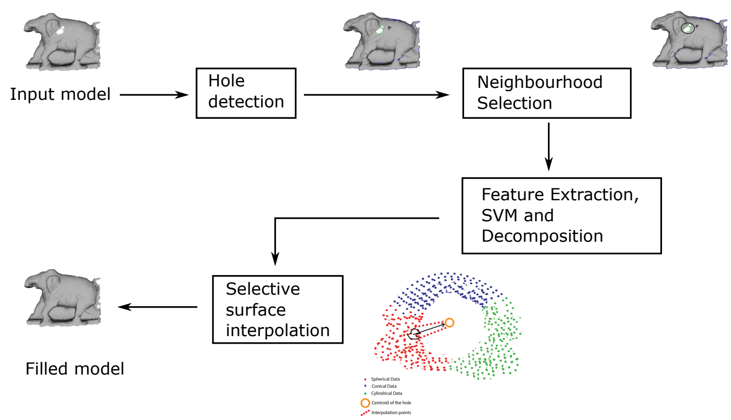


Fig. 6. Proposed hole filling algorithm.

4 Coarse to Fine 3D Reconstruction and Rendering

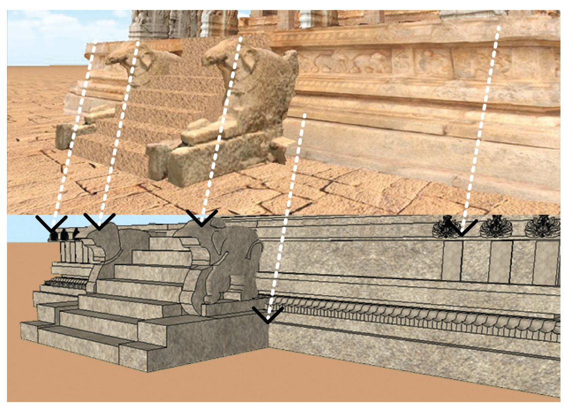


Fig. 7. Coarse to fine reconstruction of 3D objects using ICP registration with corresponding points in the coarse model and the fine model. Upper part of the image shows the fine reconstruction model and lower part of the image shows the coarse reconstruction model.

In this section we present the coarse to fine 3D reconstruction and the rendering of the reconstructed models for the generation of digital walkthrough. We carry out coarse level 3D reconstruction using methods such as single-view 3D reconstruction [19] or from modeling tools. The models generated using modeling tools and single view reconstruction do not accurately portray the geometrical complexities of the artwork at the cultural heritage sites. However, the fine level

3D reconstruction of large scale outdoor objects is not feasible using the techniques discussed in the previous section. To resolve this issue we propose a coarse to fine level 3D reconstruction of the cultural heritage sites. The coarse to fine level 3D reconstruction is achieved by registering the coarse level 3D models with the fine level 3D models. The fine level 3D models are superimposed on the coarse level 3D models by interactively selecting the correspondence points in the model. The coarse and fine level 3D models are subsequently registered using the ICP (Iterative Closest Point) algorithm [20] for the corresponding points as shown in Fig. 7.

The coarse to fine 3D reconstructed models are rendered for the generation of the digital walkthrough. The rendering of the reconstructed models is carried out using either a rendering engine like OGRE 3D or a gaming engine like Unity 3D.

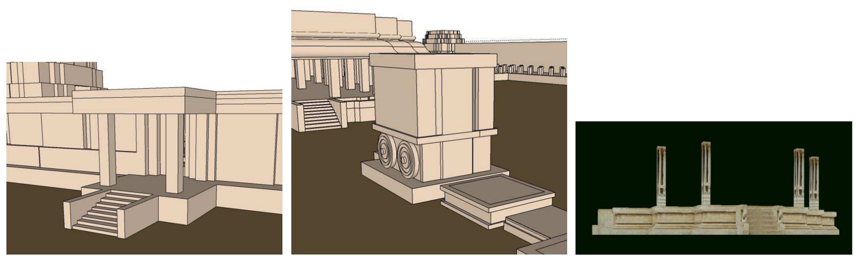


Fig. 8. Coarse level models obtained using CAD and single view reconstruction for Maha Mantapa, Stone Chariot and kalyan mantap at Vittala Temple - Hampi

Fig. 9. Figure shows the 3D reconstructed models from Laser scanner for stone chariot, Kinect model for a pillar at main mantapa and CMVS model for a pillar at kalyan mantapa.

5 Results and Discussion

We demonstrate the proposed framework for Vittala Temple at Hampi, India. The data processing algorithms are implemented on Intel(R) Xeon(R) CPU E5-

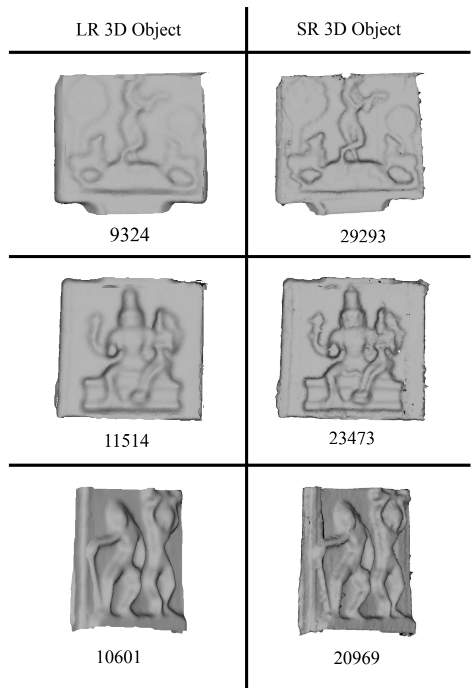
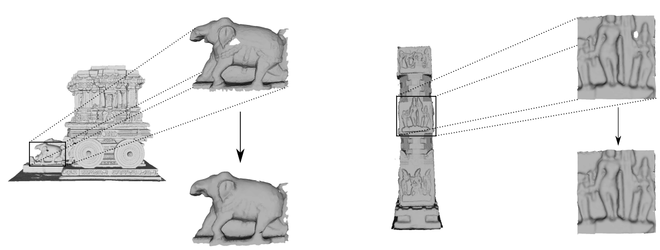


Fig. 10. Results for the proposed super resolution algorithm. Left column shows the 3D Objects of Low Resolution (LR) point cloud data. Right column shows the 3D Objects of Super Resolved (SR) point cloud data.



(a) (b)

Fig. 11. Results for the proposed hole filling algorithm.

2665 0 @2.40GHz (16 CPU's) and 64GB RAM with NVIDIA Quadro K5000 graphics, 4GB DDR3 graphics memory.

5.1 3D Reconstruction

The coarse level 3D models at the cultural heritage site are obtained either using single view reconstruction or using modeling tools. The CAD model for the Vittala Temple and the single view reconstruction of the Kalyan Mantap at Vittala Temple is as shown in Fig. 8. The fine level models are obtained using 3D scanning devices like laser scanner, Microsoft Kinect and image based methods like SFM and PMVS as shown in Fig. 9.

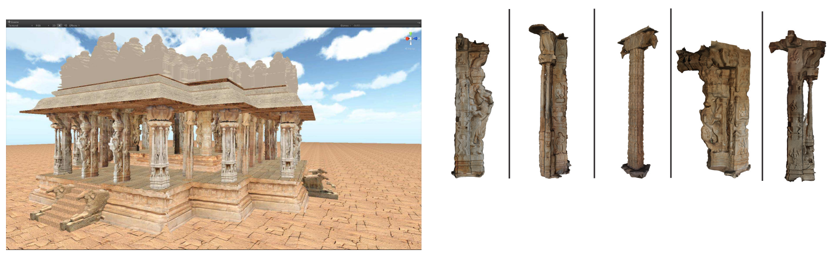


Fig. 12. Coarse to Fine 3D reconstructed model of Kalyan Mantapa along with Reconstructed models of the five variants of the pillars at Kalyan Mantapa, Vittala Temple - Hampi.

Fig. 13. Rendered views of Kalyan mantap and stone chariot using Unity 3D gaming engine and OGRE 3D rendering engine.

5.2 3D Super Resolution and Hole Filling

The fine level 3D models are processed using the proposed 3D super resolution and 3D hole filling algorithm. The processed 3D models are then surface reconstructed using the Poisson surface reconstruction algorithm with the following parameter values *Octree depth* $= 12$, *Solver divide* $= 10$, *Samples per node* $= 2$, $Surface\ offsetting = 1$. The surface reconstructed models are subsequently textured mapped using image alignment mutual information and registration of rasters. Fig. 10 shows super resolution models generated for different artifacts of one of the pillers at Main Mantap - Hampi with magnification factor of approx 2. Fig. 11 (a) shows hole filling for a part of the Stone Chariot at Vittala Temple - Hampi and Fig. 11 (b) shows hole filling for a part of the artifact of one of the piller's at Main Mantap - Hampi.

5.3 Coarse to Fine 3D Reconstruction

The coarse level models and fine level models are registered using ICP algorithm [20]. The coarse to fine level reconstruction of Kalyan Mantapa, Vittala Temple is shown in Fig 12. The pillars at Kalyan Mantap can be classified into five different types. The fine level models for the five variants of the pillars are obtained using the proposed pipeline and are as shown in Fig 12. The fine level models of the pillars comprise of roughly $\approx 300,000$ vertices and $\approx 600,000$ triangles.

5.4 Realistic Rendering

The coarse to fine level reconstructed models are rendered using OGRE 3D rendering engine and Unity 3D gaming engine and the rendered views are shown

Fig. 14. Comparison of rendered views and the original images at Vittala Temple. Left: Rendered scene, Right: Original image.

in Fig 13. The closeups of the rendered scene and the original images at the Vittala Temple are shown in Fig 14.

6 Conclusion

In this paper we have proposed a framework for the realization of digital walkthrough of cultural heritage sites. Digital restoration and preservation of cultural heritage sites is an important area of research due to the availability of techniques in data acquisition, data processing and rendering. The main goal of the paper is to create a framework for the generation of digital walkthrough of cultural heritage sites. To accomplish this we have proposed a framework for coarse to fine level 3D reconstruction using coarse level 3D reconstruction of the cultural heritage sites and fine level 3D reconstruction of the artworks at the cultural heritage sites. We also proposed data processing algorithms like 3D super resolution and 3D hole filling using concepts of Riemannian geometry with metric tensor and Christoffel symbols as a novel set of features. We have demonstrated the proposed framework for Vittala Temple - Hampi, India.

Acknowledgement. This work is supported by the Department of science and technology (DST), India, under grant NRDMS/11/201/Phase-III/ as a part of India Digital Heritage project. We thank DST, NIAS Banglore and IIT Delhi for the support.

References

1. Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., Fulk, D.: The digital michelangelo project: 3d scanning of large statues. In: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. SIGGRAPH '00, New York, NY, USA, ACM Press/Addison-Wesley Publishing Co. (2000) 131–144
2. Ikeuchi, K., Oishi, T., Takamatsu, J., Sagawa, R., Nakazawa, A., Kurazume, R., Nishino, K., Kamakura, M., Okamoto, Y.: The great buddha project: Digitally archiving, restoring, and analyzing cultural heritage objects. International Journal of Computer Vision **75** (2007) 189–208
3. Ikeuchi, K., Sato, Y.: Modeling from reality. Kulwer Academic Press (2001)
4. Wasserman, J.: Michelangelo's florence peita. Princeton University Press (2003)
5. Stamos, I., Allen, P.K.: Automatic registration of 2-d with 3-d imagery in urban environments. In: ICCV. (2001) 731–737
6. Snavely, N., Seitz, S.M., Szeliski, R.: Photo tourism: Exploring photo collections in 3d. ACM Trans. Graph. **25** (2006) 835–846
7. Furukawa, Y., Ponce, J.: Accurate, dense, and robust multi-view stereopsis. IEEE Trans. on Pattern Analysis and Machine Intelligence **32** (2010) 1362–1376
8. Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A., Fitzgibbon, A.: Kinectfusion: Real-time 3d reconstruction and interaction using a moving depth camera. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. UIST '11, New York, NY, USA, ACM (2011) 559–568

9. Ganihar, S., Joshi, S., Patil, N., Mudenagudi, U., Okade, M.: Voting-based decision framework for optimum selection of interpolation technique for 3d rendering applications. In: Students' Technology Symposium (TechSym), 2014 IEEE. (2014) 270–275
10. Jost, J.: Riemannian Geometry and Geometric Analysis. Springer Universitat texts. Springer (2005)
11. Kumaresan, S.: A course in differential geometry and Lie groups. Texts and readings in mathematics. Hindustan Book Agency (2002)
12. Weinberg, S.: Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity. Wiley, New York, NY (1972)
13. Misner, C., Thorne, K., Wheeler, J.: Gravitation. W.H. Freeman and Company (1973)
14. Burges, C.J.: A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery **2** (1998) 121–167
15. Liepa, P.: Filling holes in meshes. In: Proceedings of the 2003 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing. SGP '03, Aire-la-Ville, Switzerland, Switzerland, Eurographics Association (2003) 200–205
16. Kazhdan, M., Bolitho, M., Hoppe, H.: Poisson surface reconstruction. In: Proceedings of the Fourth Eurographics Symposium on Geometry Processing. SGP '06, Aire-la-Ville, Switzerland, Switzerland, Eurographics Association (2006) 61–70
17. Bernardini, F., Mittleman, J., Rushmeier, H., Silva, C., Taubin, G.: The ball-pivoting algorithm for surface reconstruction. IEEE Transactions on Visualization and Computer Graphics **5** (1999) 349–359
18. Corsini, M., Dellepiane, M., Ponchio, F., Scopigno, R.: Image-to-geometry registration: a mutual information method exploiting illumination-related geometric properties. Computer Graphics Forum **28** (2009) 1755–1764
19. Koutsourakis, P., Simon, L., Teboul, O., Tziritas, G., Paragios, N.: Single view reconstruction using shape grammars for urban environments. In: Computer Vision, 2009 IEEE 12th International Conference on. (2009) 1795–1802
20. Rusinkiewicz, S., Levoy, M.: Efficient variants of the ICP algorithm. In: Third International Conference on 3D Digital Imaging and Modeling (3DIM). (2001)