Supplementary Material

Multi-view Non-rigid Refinement and Normal Selection for High Quality 3D Reconstruction

Sk. Mohammadul Haque, Venu Madhav Govindu
Indian Institute of Science
Bengaluru, India
{smhaque, venu}@ee.iisc.ernet.in

1 Generation of synthetic data

We generate our input data used in the synthetic experiments as follows. For each scan, we generate a deformation map of same size as the corresponding input depth map. A uniformly random 2D location is selected and assigned a deformation value sampled from a uniformly distribution with absolute maximum of 1cm. We then blur the deformation map with a Gaussian kernel of size 1/4 of its minimum dimension. We then apply this deformation to the ground truth depth map to generate the synthetically deformed scan.

2 Additional results

In addition to the results shown in the main paper, we present here some more detailed results on synthetic datasets and real datasets to show the effectiveness of our method. As in the main paper, we compare our results with three approaches, namely **AVG** which is a simple averaging of all available normals for a given mesh vertex, **MERGE2-3D** of [2] and the Priority Selection scheme **PS** for normal selection described in [9].

2.1 Synthetic datasets

In Figure 1, we show comparative results on a synthetic dataset of the Armadillo from the Stanford 3D scanning repository for 12 views. The first column shows overall views of the 3D reconstructions obtained. The second column shows the normal selection labels for **PS** and **OURS**. The methods **AVG** and **MERGE2-3D** do not select normals and hence are indicated in grey. The zoomed-in views of the highlighted regions of the right side of the neck-region (red box) of the reconstructions are respectively shown in the third column. From the third column in Figure 1, it can be observed that our method retains the sharp edges (marked with arrows) in the neck region of the Armadillo better than all the other three methods. The fourth column shows the rear-views of the respective reconstructions. From the fourth column, it can be observed that our method minimises the non-rigid discrepancies (marked with arrows) across different views at the rear side of the Armadillo while at the same time preserves the high frequency details on the reconstructed surface. However non-rigid artefacts are observed in the reconstruction of the other methods.

In Figure 2, we show comparative results on a synthetic dataset of a sphere with multiple ridges for 12 views. The first column shows overall views of the 3D reconstructions obtained. The second column shows the normal selection labels for **PS** and **OURS**. The methods **AVG** and **MERGE2-3D** do not select normals and hence are indicated in grey. The views of the obverse side of the reconstructions are respectively shown in the third column. From the first and third columns in Figure 2, it can be observed that our method retains the sharp edges (marked with arrows) of the ridges better than all the other three methods by minimising the non-rigid discrepancies across different views. However non-rigid artefacts are observed in the reconstructions of the other methods. Since our adaptive normal selection step considers normals coming from regions that have non-rigid deformations as unreliable, it can be observed from the second column in Figure 2, that in the normal selection step of our method, the normals in the region marked with an arrow are automatically selected from a view (pink coloured) that has lesser amount of non-rigid deformation than the more frontal view (violet coloured). This ensures that the final selected normals are most reliable not only considering the viewing direction but also considering the non-rigid deformations.

2.2 Real datasets

In Figure 3, we show comparative results on the clay Buddha figurine obtained from 3 views. The first row shows overall views of the 3D reconstructions obtained. The second row shows the normal selection labels for **PS** and **OURS**. The methods **AVG** and **MERGE2-3D** do not select normals and hence are indicated in grey. The zoomed-in views of the highlighted regions of the left rib-region (red), face (black) and right hand (pink) of the reconstruction are respectively shown in Figure 4. From the first row in Figure 4, it can be observed that our method minimises the non-rigid discrepancies across different views near the left rib-region while at the same time preserves the high frequency details on the reconstructed surface. In the second row of Figure 4, it can be observed that the non-rigid artefacts on the two sides of the nose are observed in the reconstruction of the other methods. However, such artefacts are not present in our result. Similarly, in the third row of Figure 4, it can observed that our method preserves most of the details while the other methods fail to do so.

In Figure 5, we show comparative results on a clay elephant model obtained from 5 views. The first row shows overall views of the 3D reconstructions obtained. The second row shows the normal selection labels for **PS** and **OURS**. The methods **AVG** and **MERGE2-3D** do not select normals and hence are indicated in grey. The zoomed-in views of the highlighted regions of the face (apple-green) and side (red) are respectively shown in the third and fourth rows. From the third row, it can be observed that the averaging methods **AVG** and **MERGE2-3D** fail to recover the high frequency details and results in blurring of details. However, **PS** and our method **OURS** are able to recover the details. On the other hand, from the fourth row, it can be observed that while the method **PS** fails to select good normals resulting in poor reconstruction, only our method recovers most of the details.

All the citation numbers are as in the main paper.

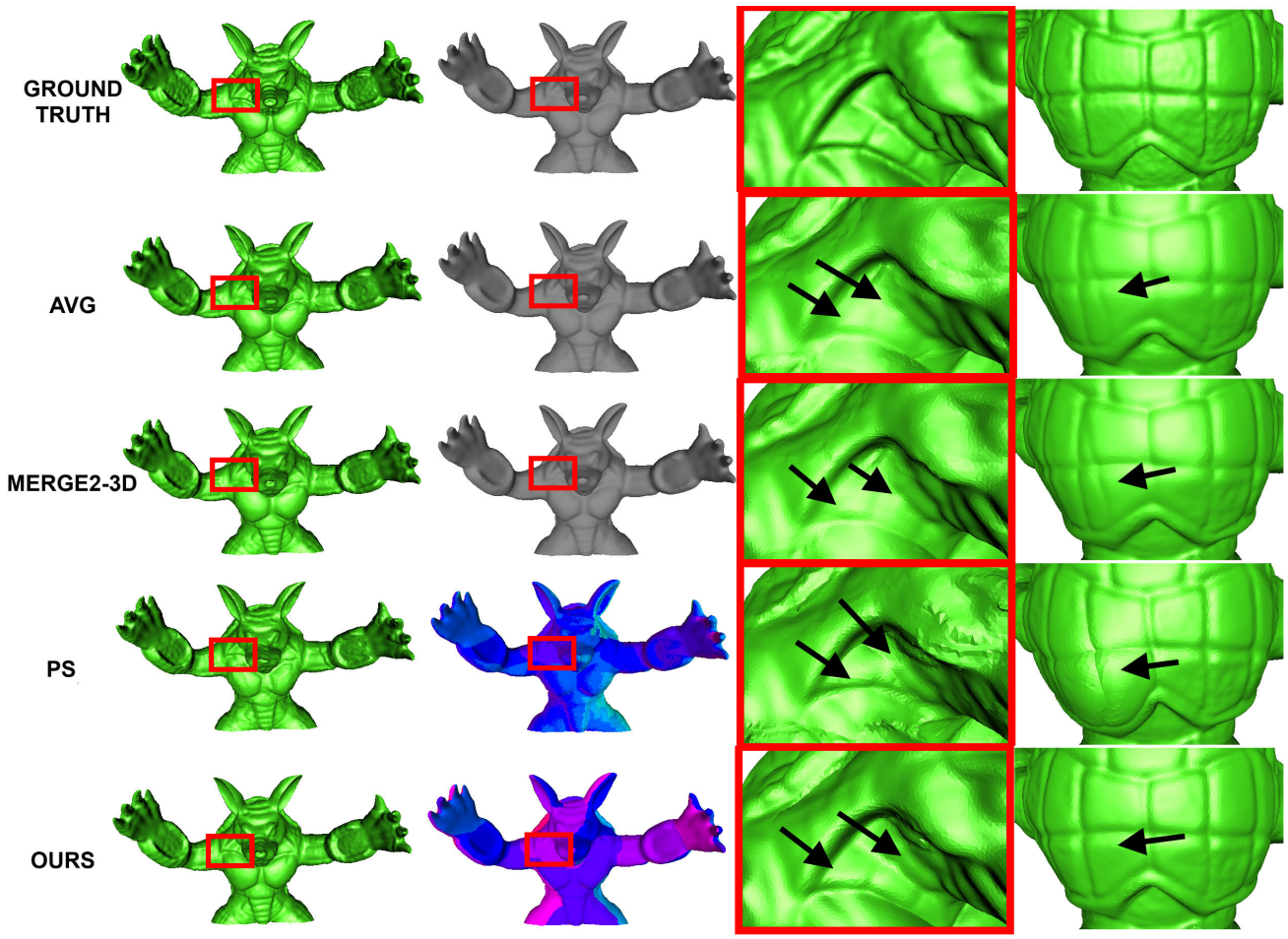

Figure 1: Comparative results on the Armadillo from the Stanford 3D scanning repository obtained for 12 views. The rows correspond to ground truth, the methods **AVG**, **MERGE2-3D** of [2], **PS** of [9] and our method **OURS** respectively. The first column shows overall views of the 3D reconstructions obtained. The second column shows the normal selection labels for **PS** and **OURS**. The methods **AVG** and **MERGE2-3D** are indicated in grey since they average all available normals and do not select normals. The zoomed-in views of the highlighted regions of the right side of the neck-region (red box) of the reconstructions are respectively shown in the third column. The fourth column shows the rear-views of the respective reconstructions. Please see this figure in colour and text for details.

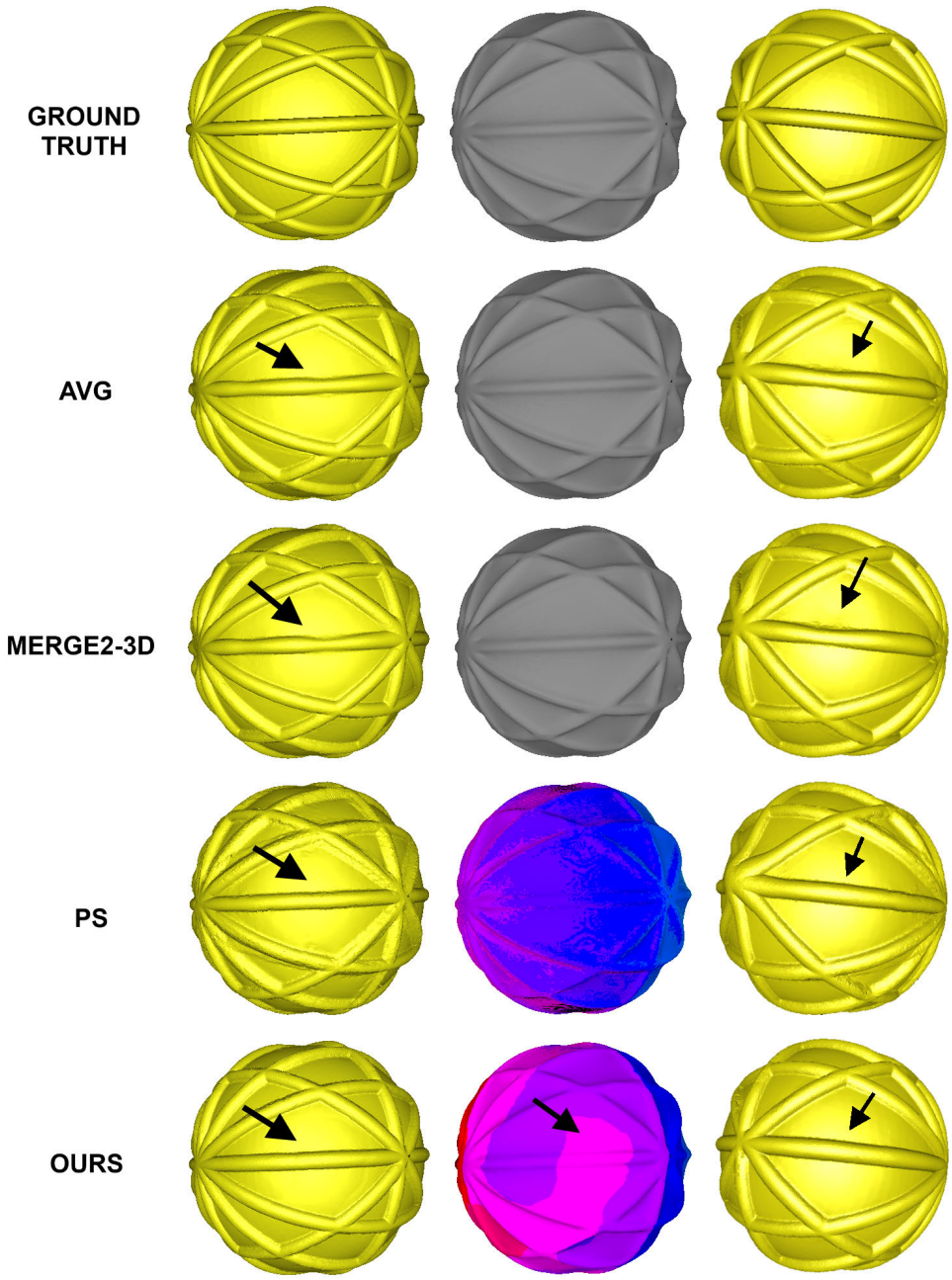

Figure 2: Comparative results on a synthetic dataset of a sphere with multiple ridges for 12 views. The rows correspond to ground truth, the methods **AVG**, **MERGE2-3D** of [2], **PS** of [9] and our method **OURS** respectively. The first column shows overall views of the 3D reconstructions obtained. The second column shows the normal selection labels for **PS** and **OURS**. The methods **AVG** and **MERGE2-3D** are indicated in grey since they average all available normals and do not select normals. The views of the obverse side of the reconstructions are respectively shown in the third column. Please see this figure in colour and text for details.

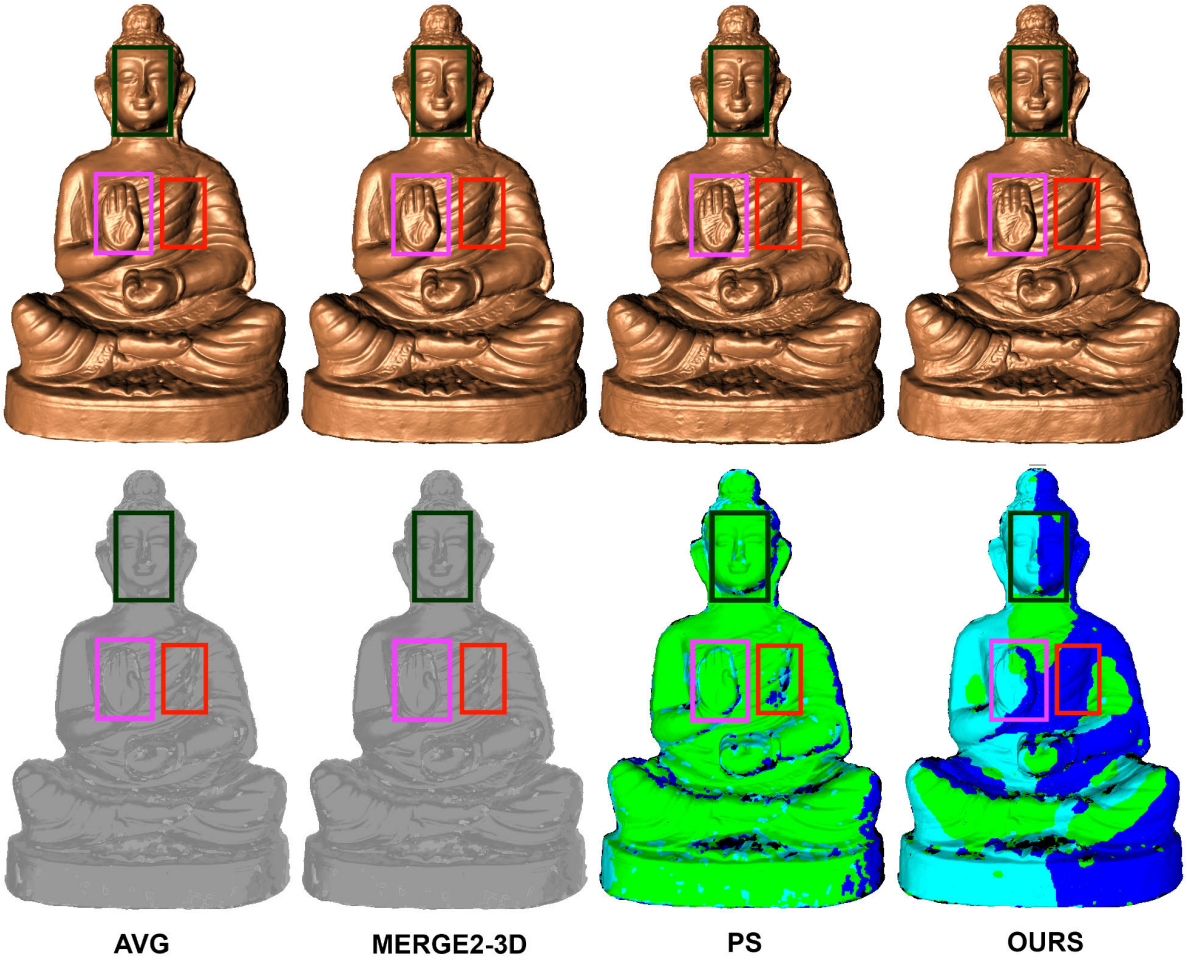

Figure 3: Comparative results on the clay Buddha figurine obtained from 3 views. The columns correspond to the methods **AVG**, **MERGE2-3D** of [2], **PS** of [9] and our method **OURS** respectively. The first row shows overall views of the 3D reconstructions obtained. The second row shows the normal selection labels for **PS** and **OURS**. The methods **AVG** and **MERGE2-3D** are indicated in grey since they average all available normals and do not select normals. The zoomed-in views of the highlighted regions of the left rib-region (red), face (black) and right hand (pink) of the reconstruction are respectively shown in Figure 4. Please see this figure in colour and text for details.

Figure 4: Comparative results on zoomed-in regions (highlighted in Figure 3) on the clay Buddha figurine obtained from 3 views. The columns correspond to the methods **AVG**, **MERGE2-3D** of [2], **PS** of [9] and our method **OURS** respectively. The first, second and third rows show zoomed-in views of the left rib-region, face and right hand of the reconstruction respectively. Please see text for details.

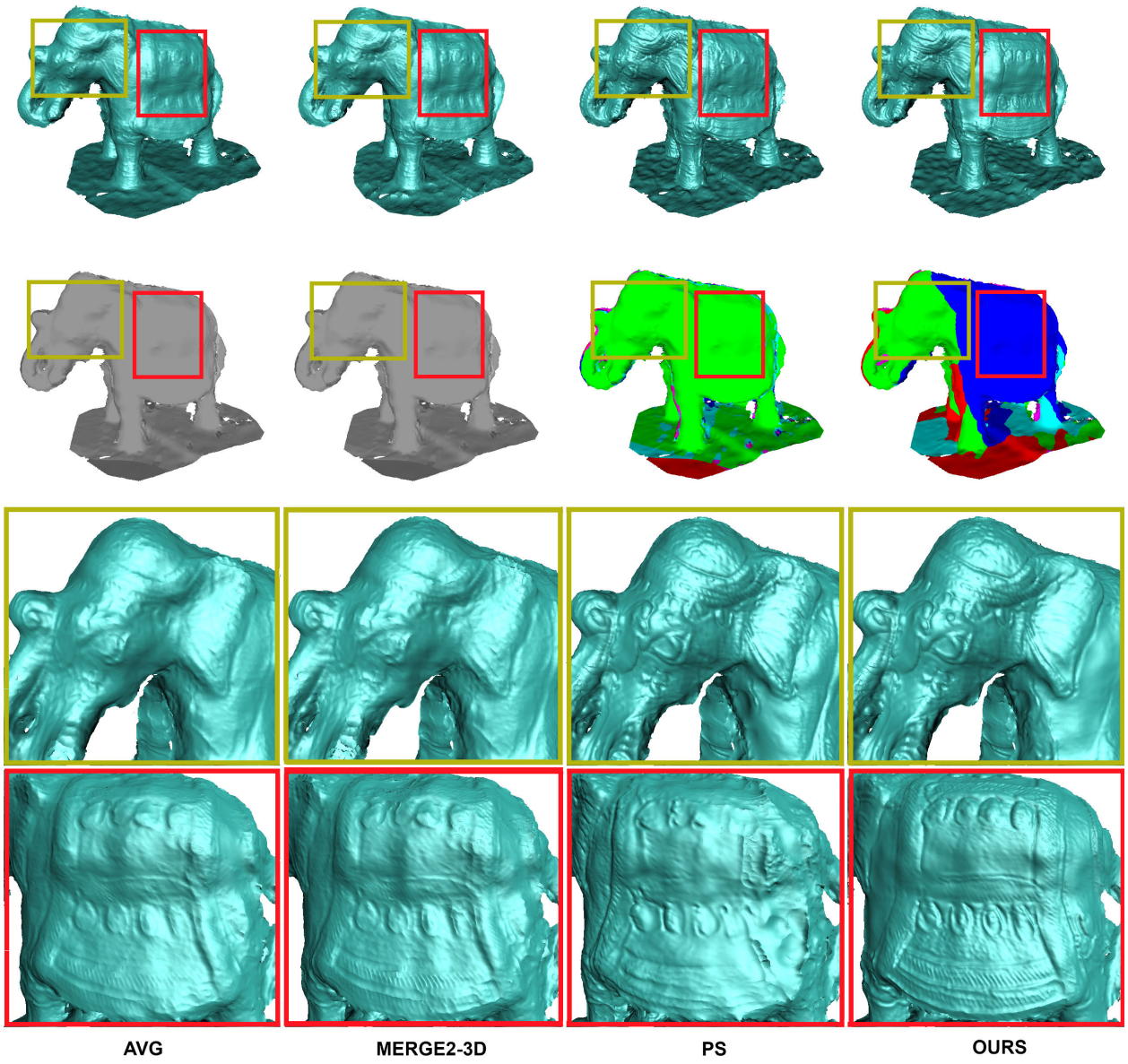

Figure 5: Comparative results on a clay elephant model obtained from 5 views. The columns correspond to the methods **AVG**, **MERGE2-3D** of [2], **PS** of [9] and our method **OURS** respectively. The first row shows overall views of the 3D reconstructions obtained. The second row shows the normal selection labels for **PS** and **OURS**. The methods **AVG** and **MERGE2-3D** are indicated in grey since they average all available normals and do not select normals. The third and fourth rows show zoomed-in views of the face (apple-green) and side (red) of the reconstruction. Please see this figure in colour and text for details.

3 Relationship between non-rigid refinement step and normal selection step

The roles of our two contributions of non-rigid refinement and normal selection are crucially linked together. We make the following observation. Even if the other methods (**AVG**, **MERGE2-3D** of [2], **PS** of [9]) were to use non-rigid refinement (which they do not), one will still need to carefully account for the reliability of the resulting alignment. We take into consideration the fact that larger deformations are less reliable and this should play an important role in normal selection. In other words, our novelty is not merely in putting non-rigid refinement and normal selection together. Our method considers the estimated amount of deformation and appropriately weights the normals for normal selection.

We demonstrate this argument using the sphere with ridges dataset of the paper by showing the reconstruction errors without and with non-rigid refinement for all the methods below in Table 1.

Method ———>	AVG	MERGE2-3D	PS
WITHOUT Non-Rigid Refinement	0.451	0.630	0.450
WITH Non-Rigid Refinement	0.458	0.607	0.432
Our method: 0.418			

Table 1: MSE Reconstruction errors

With the non-rigid refinement, all the methods, except AVG, performed better than without any non-rigid refinement. AVG performed poorer after non-rigid refinement as it blindly averages the normals in the unreliable deformed regions that are oblique to the view.

But we carefully account for the fact that deformed regions have less reliable normals and that the normal selection is better than averaging. As can be seen, our proposed method gives a lower reconstruction error of 0.418.

4 Run-times

Run-times in Matlab 64bit on a desktop computer with Core i7 3770K 16GB RAM are given below in Table 2.

Dataset	#Views	#Vertices	Run-time (secs)		
			Ours (Non-rigid+Selection)	AVG	MERGE2-3D
Sphere Ridges	12	350562	(1062+27) = 1089	8	31
Armadillo	12	122099	(362+11) = 373	4	11
Horse	6	540715	(258+24) = 282	8	32
Clay Pot	4	691503	(113+22) = 135	6	35
Buddha	3	481202	(118+14) = 132	4	14
Elephant	5	309249	(218+18) = 236	5	11

Table 2: Run-times for different methods

N. B. Priority Selection (PS) run-time is not applicable due to manual selection.