What Do Neural Radiance Fields Offer for 3D Human Reconstruction?

Kamil Opalach[1]✉[0000−0001−6550−1859], Joanna Porter-Sobieraj[1][0000−0002−1411−475X], and Robert Sitnik[1,2][0000−0002−8156−5462]

[1] Warsaw University of Technology, Warsaw, Poland
[2] Mnemosis, Krakow, Poland
kamil.opalach.dokt@pw.edu.pl

Abstract. High-fidelity 3D human scans are essential in various engineering applications, including entertainment and medicine. Recent advancements in neural scene representation techniques, such as Neural Radiance Fields (NeRF), offer promising alternatives to traditional 3D reconstruction methods like Structure from Motion (SfM). This paper contributes to Computational Science by presenting a comparative study of two computational paradigms for 3D human reconstruction — NeRF, implemented using Instant-NGP, and SfM, carried out with the use of Agisoft Metashape — with particular emphasis on their temporal stability. The analysis focuses on the reconstruction of a detailed human body by evaluating reconstructions carried out across three consecutive time frames for each test scene. The evaluation employs extensive point cloud analysis based on four quantitative criteria: reconstruction completeness, resolution, surface smoothness, and point cloud consistency. The results demonstrate that NeRF surpasses photogrammetry in capturing fine volumetric structures and reconstructing regions that are only visible in a limited subset of input views, yielding a reconstructed area approximately 9% larger in the analysed full-body scans. Conversely, SfM demonstrates superior performance in terms of surface smoothness, producing reconstructions with approximately 45% higher surface smoothness than those obtained with NeRF. Temporal analysis reveals that SfM maintains high geometric stability between consecutive frames, while NeRF exhibits discretisation effects due to its voxel-based representation, leading to greater frame-to-frame variability. The complementary strengths of these methods highlight opportunities for future research to leverage both artificial intelligence and traditional techniques for enhanced 3D reconstruction.

Keywords: 3D reconstruction · Neural Radiance Fields · Structure from Motion · photogrammetry · point clouds · temporal stability.

1 Introduction

Image-based 3D reconstruction is a widely used technique in computer graphics and computer vision that enables the recovery of 3D object properties, such as

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34

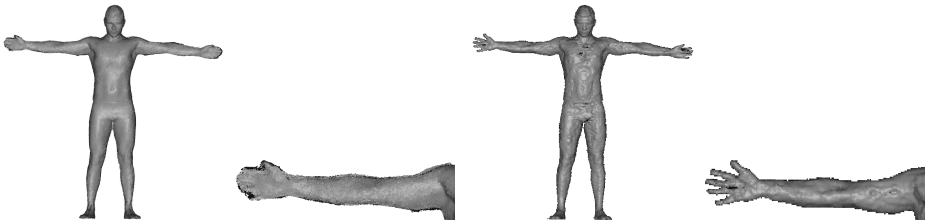
shape and colour, from multiple 2D images captured from different viewpoints [18]. The resulting representations are typically realised as point clouds containing spatial coordinates and additional attributes such as colour and normal vectors. When image sequences represent successive time frames of object motion, the process extends to spatio-temporal or 4D reconstruction, allowing the analysis of surface changes over time.

In recent years, 3D reconstruction has gained significant attention in human body scanning applications, particularly in entertainment and healthcare [22, 8]. Spatio-temporal reconstruction enables realistic motion capture and animation in entertainment [11], as well as posture assessment, rehabilitation monitoring, and diagnostic analysis in medicine [6, 9]. Insufficient temporal stability can introduce geometric artefacts and noise, leading to unnatural motion representation and reduced analytical reliability.

Among image-based reconstruction techniques, Structure from Motion (SfM) is widely used to generate dense 3D point clouds from multiple images with known or estimated camera parameters [25, 2]. SfM has been successfully applied in cultural heritage preservation, landscape modelling, medical imaging, and human body reconstruction [6]. However, SfM is sensitive to low-texture regions, transparency, and occlusions, often resulting in incomplete reconstructions [16].

Neural Radiance Fields (NeRF), introduced in 2020 [14], offer a more recent alternative by learning a continuous volumetric representation of a scene from overlapping 2D images. NeRF enables novel view synthesis and point cloud extraction, offering a promising approach to reconstruction.

Figure 1 compares SfM and NeRF reconstructions of a human body dataset using identical inputs. SfM produces smoother and denser surfaces but struggles with fine structures such as fingers, while NeRF better preserves small details at the cost of surface artefacts and discrete offsets in curved regions.



(a) SfM reconstruction (b) NeRF reconstruction

Fig. 1: Reconstruction results and a close-up of arm reconstruction for SfM and NeRF methods.

This paper provides a detailed comparison of SfM- and NeRF-based reconstruction techniques, focusing on their performance in reconstructing human body datasets. We assess the quality and temporal stability of the reconstructions using the same sets of input images with known camera poses. The evalua-

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34

tion includes challenging cases such as self-occlusions and highly detailed regions, the likes of which frequently appear in human body scans. By analysing how each method handles human reconstruction and maintains geometric coherence across time frames, we investigate the trade-offs between the computational paradigms of NeRF and SfM and evaluate their complementary use for high-fidelity human body reconstruction.

2 Related work

Structure from Motion: Structure from Motion (SfM) is a computational framework for reconstructing the three-dimensional structure of a scene from multiple two-dimensional images. The process typically involves feature extraction and matching (e.g., Scale Invariant Feature Transform [12]), followed by triangulation to obtain a sparse reconstruction. Dense geometry is then recovered using Multi-View Stereo (MVS) techniques [21]. Early work by Snavely et al. [23] established a foundational pipeline for recovering camera poses and scene structure from unordered image collections, while later efforts addressed scalability through parallelisation [1]. The COLMAP system [20] further advanced the state of the art by improving triangulation, bundle adjustment, and outlier rejection, leading to higher reconstruction quality.

Neural Radiance Fields: Neural Radiance Fields (NeRF), introduced by Mildenhall et al. [14], represent a major advancement in 3D reconstruction and novel view synthesis. NeRF uses a multi-layer perceptron to model a continuous function that maps 3D positions and viewing directions to volume density and colour. Training optimises photometric consistency across input images using differentiable volumetric rendering. Subsequent extensions have improved NeRF's robustness and efficiency. *Mip-NeRF* [3] addresses aliasing via multiscale representations, while *NeRF in the Wild* [13] adapts the method to unconstrained environments with variable illumination and transient objects. More recently, hash-based encodings have been integrated into the NeRF framework to further accelerate both training and inference while preserving reconstruction quality. A technique described by Müller et al. [15] replaces the traditional encoding with a hash table that efficiently stores and retrieves spatial feature representations at multiple scales. To improve geometric fidelity, NeuS [26] replaces volumetric density with a signed distance function (SDF), and Neuralangelo [10] combines neural SDFs with multiresolution hash encodings and curvature-aware regularisation for high-resolution surface reconstruction in large-scale scenes. Similarly, Zip-NeRF [4] combines multiscale NeRF formulations and grid-based models through multiscale sampling and filtering along rays, achieving improved rendering quality and efficiency.

Several studies have evaluated NeRF reconstruction in the context of traditional photogrammetry methods such as SfM and MVS. However, these works do not perform comparisons using real-world data that represent the human body. Furthermore, none of them provides an analysis of the temporal stability of NeRF-based reconstruction.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34

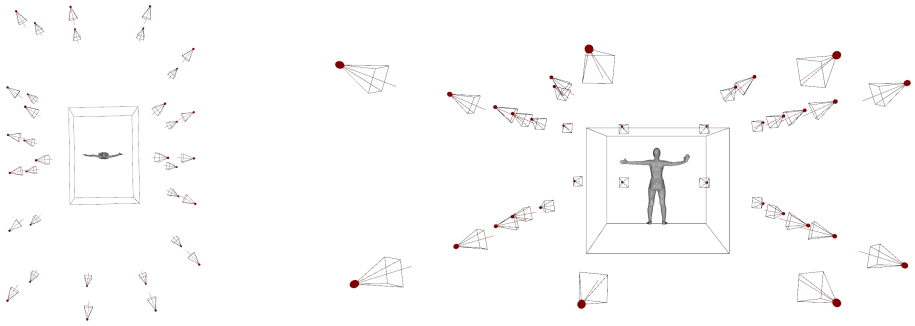
Croce et al. [7] evaluate NeRF and photogrammetry in the context of cultural heritage, analysing reconstruction quality under varying image quantities and resolutions. Their results show that NeRF achieves higher completeness and material preservation on datasets with fewer images and lower image resolution. Moreover, NeRF more accurately represents surfaces with view-dependent materials, suggesting the potential benefits of hybrid approaches. Similarly, Remondino et al. [19] compare NeRF and MVS across objects with diverse sizes and surface properties, finding NeRF more effective for transparent and metallic surfaces, while photogrammetry performs better on diffuse-like surfaces. Petrovska et al. [17] investigate the impact of occlusion, showing that although SfM achieves higher accuracy, NeRF provides significantly higher coverage of occluded objects.

3 Methodology

3.1 Test cases

The evaluation of SfM and NeRF reconstructions uses two real-world full-body datasets - Male T-pose and Female T-pose - depicting individual subjects captured in an empty room. Image sequences are recorded synchronously at 30 fps, with each test case consisting of three consecutive frames in which the actors maintain nearly identical poses. Minor inter-frame motion enables assessment of temporal stability and comparison between reconstruction methods. The scenes include regions with varying surface curvature, self-occlusions, and fine details.

Each scene comprises 34 images at a resolution of 4096 × 2160 pixels, captured simultaneously from multiple viewpoints using cameras placed at different heights around the subject. Camera positions and orientations are predefined, as illustrated in Figure 2.



(a) Top view (b) Front view

Fig. 2: Layout of the cameras for the full body scene.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34

3.2 Reconstruction process

The captured images and corresponding camera parameters are processed using two reconstruction pipelines to generate point clouds for comparative analysis:

- **Structure-from-Motion (SfM):** Dense point clouds are generated independently for each scene using Agisoft Metashape, producing 3D reconstructions directly suitable for comparison.
- **Neural Radiance Fields (NeRF):** Neural reconstruction is performed using the Instant-NGP implementation [15], as it provides superior performance compared to other NeRF methods [19, 17] while enabling fast, high-quality reconstructions. The volumetric representation encodes density values corresponding to the scene geometry. The final 3D point clouds are obtained by filtering high-density values from the sampled volumetric grid, where points of higher density correspond to the solid surface of the subject. As NeRF reconstructions may also include internal points within the object, filtering is followed by post-processing to remove internal points via multi-view depth testing.

All 3D reconstructions were conducted on a workstation equipped with an Intel Core i7-11700K CPU, 32 GB of RAM, and an NVIDIA GeForce RTX 3070 Ti GPU with 8 GB of VRAM.

3.3 Comparison criteria

This analysis directly compares the geometric properties of 3D point clouds produced by the two reconstruction methods. All point clouds are first aligned to a common coordinate system to enable consistent evaluation. The comparison is based on quantitative metrics that focus on different aspects of geometric quality:

- **Cloud-to-cloud comparison** evaluates geometric differences between two point clouds. The evaluated cloud P is aligned to the reference cloud Q using the Iterative Closest Point (ICP) algorithm [5]. Nearest-neighbour distances are then computed with a cutoff defined as $d_{\text{cutoff}} = 2 \cdot a$, where a is the voxel size of the NeRF point cloud. For each point $p_i \in P$, the nearest point $q_j \in Q$ is identified, and the Euclidean distance d_i is calculated as:

$$d_i = \min \left(\min_{q_j \in Q} \|p_i - q_j\|, d_{\text{cutoff}} \right). \tag{1}$$

  A histogram of d_i visualises the distribution of geometric discrepancies between the two point clouds providing further insight into the nature and magnitude of the differences. Additionally, the median distance from P to Q is computed as:

$$d_{\text{med}} = \text{median} \left(\{ \min_{q_j \in Q} \|p_i - q_j\| \mid p_i \in P \} \right) \tag{2}$$

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34

This metric is asymmetric, as differences in point density and reconstruction artefacts affect each comparison direction differently.

– **Reconstruction resolution** assesses variations in point density across a point cloud P. Two analyses are performed:
  - **Average distance to k-nearest neighbours:** For each point $p_i \in P$, the average distance to its k nearest neighbours is calculated as:

$$D(p_i) = \frac{1}{k} \sum_{j \in NN_k(p_i)} \|p_i - p_j\|, \tag{3}$$

    where $NN_k(p_i)$ is the set of the k-nearest neighbours of p_i.
    A histogram of $D(p_i)$ visualises the distribution of average distances across the entire point cloud and reveals variations in local point spacing.
  - **Number of neighbours in a fixed radius:** A histogram of the number of neighbours $|N_r(p_i)|$ within a radius r is used to examine the distribution of point densities.
– **Surface smoothness** quantifies local surface regularity. For each point p_i, a plane is fitted to neighbouring points within radius r using a covariance matrix and Singular Value Decomposition (SVD) [24]. The plane normal is defined by the eigenvector associated with the smallest eigenvalue of the covariance matrix. Smoothness is computed as the mean perpendicular distance of neighbouring points to the fitted plane:

$$s_i = \frac{1}{|N_r(p_i)|} \sum_{q_j \in N_r(p_i)} d_{\text{plane}}(q_j), \tag{4}$$

    where $N_r(p_i)$ is the set of neighbouring points within radius r of p_i, and $d_{\text{plane}}(q_j)$ represents the perpendicular distance of point q_j to the fitted plane. The global smoothness of the point cloud is given by:

$$S = \frac{1}{|P|} \sum_{p_i \in P} s_i, \tag{5}$$

    where P is the set of all points in the point cloud.
    To analyse scale-dependent behaviour, smoothness is evaluated for increasing neighbourhood radii $(r_1, r_2, \ldots, r_n)$.

These metrics support both spatial and temporal evaluation. Spatial analysis compares SfM and NeRF reconstructions within the same time frame, while temporal analysis evaluates consistency between consecutive frames for each method. Since motion data is captured at 30 Hz, where pose changes between frames are minimal, this approach enables assessment of temporal stability and identification of regions where reconstruction artefacts shift over time despite similar metric values.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34

4 Results and discussion

This section presents a comparative analysis of SfM and NeRF reconstructions. Spatial comparison is conducted using the first time frame from each test case, while temporal stability is assessed based on three consecutive frames by measuring differences between neighbouring time frames for each method. The evaluation follows the criteria outlined in Section 3.3.

Two full-body test cases are analysed: a Male and a Female T-pose, each consisting of three consecutive frames used to evaluate SfM and NeRF reconstructions. The NeRF volumetric representation is sampled at a resolution of 512, producing a $512 \times 512 \times 512$ grid. Since NeRF requires the reconstructed subject to be scaled to a unit cube, the voxel sizes of the resulting density grid differ between the Male and Female datasets, measuring 4.36 mm and 4.32 mm, respectively, to fit the reconstructed geometry. Sample source images and reconstruction results for each frame are shown in Figure 3.

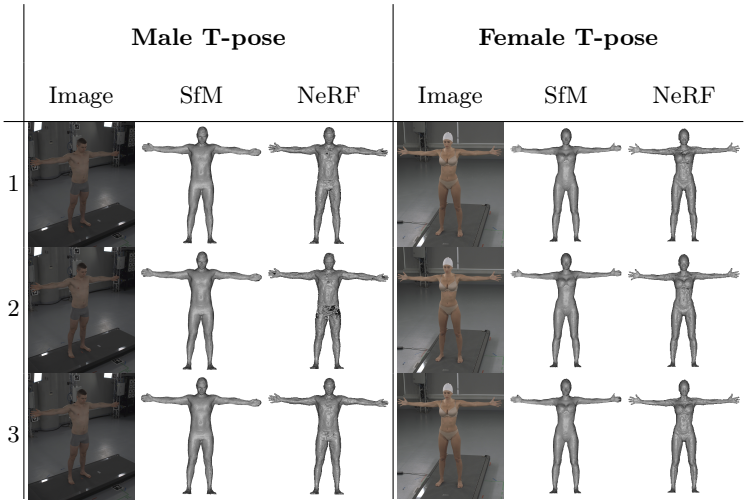
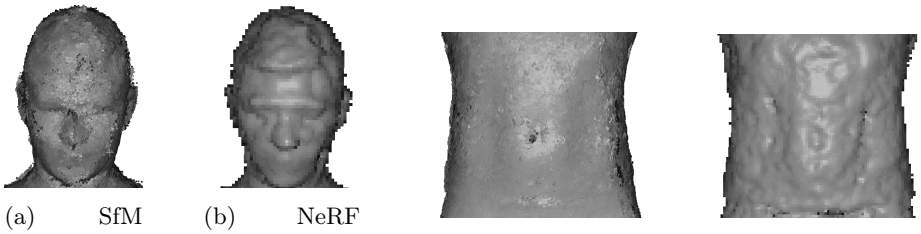


Fig. 3: Sample source images and reconstructions for three consecutive frames of Male and Female T-pose test cases.

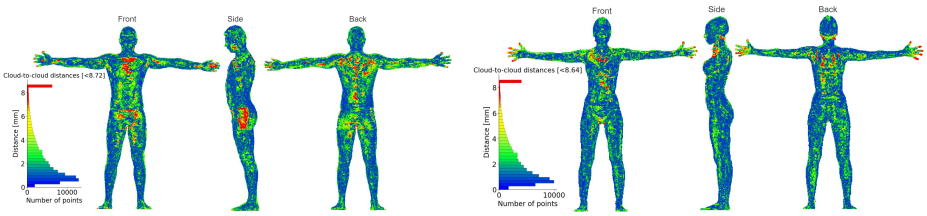
In both scenes, most body parts are well visible and fully reconstructed by both methods. However, qualitative differences are observed. For the male case, NeRF failed to reconstruct the shorts accurately in the second frame, resulting in visible surface holes. In the female case, no such cavities are present, though larger discrepancies are noted in the torso region. Additionally, close-ups of the most visually different areas in the first frame of the Male T-pose test case are presented in Figures 1 and 4. NeRF reconstruction better preserves the shape of small structures like fingers or ears. However, it is characterised by less smooth surfaces, which are especially noticeable in the torso area.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34



(a) SfM head reconstruction (b) NeRF head reconstruction (c) SfM torso reconstruction (d) NeRF torso reconstruction

Fig. 4: Male T-pose: Reconstruction results close-ups for SfM and NeRF methods.



(a) Male T-pose (b) Female T-pose

Fig. 5: Cloud-to-cloud distances comparison between SfM and NeRF reconstructions.

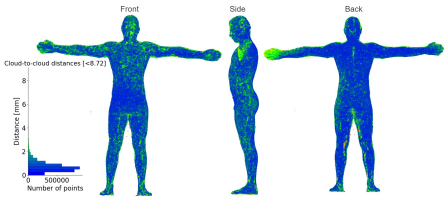
Reconstructed point cloud data for both methods are summarised in Table 1 for Male and Female T-pose test cases. SfM reconstructions consistently yield point clouds with an order of magnitude more points, lower average distances to the closest neighbour, and a higher surface density. These values remain temporally consistent across all three frames in both cases.

Cloud-to-cloud Cloud-to-cloud comparisons are used to assess reconstruction differences between the two methods, as well as temporal stability within each method. For both test cases, NeRF reconstructions match the SfM point clouds in most places. In the male case (Figure 5a), the largest discrepancies appear in the hands, upper chest, back, and hips. The average distance from the NeRF to the SfM point cloud is 3.23 mm, with 78.5% of points falling within the first half of the histogram bins and 6.7% exceeding the cutoff distance of 8.72 mm. In the female case (Figure 5b), differences are also concentrated in the fingers, upper chest, and back. NeRF again performs better in preserving fine structures like individual fingers, but exhibits larger reconstruction errors in the torso region. No holes are observed in the NeRF reconstructions of the female subject. In this case, 85.4% of NeRF points are within the first half of the histogram bins, indicating a concentration of smaller reconstruction errors.

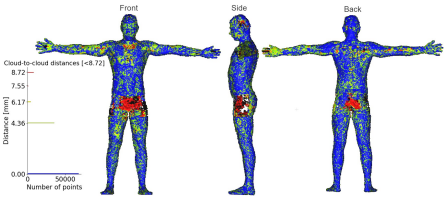
ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34

Table 1: SfM and NeRF reconstruction details for Male and Female T-pose test cases.

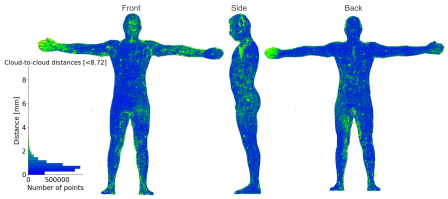
Method	Frame	Number of points	Average distance to the closest neighbour [mm]	Surface density [points/mm^2]
		Male T-pose		
SfM	1	3 203 629	0.65	1.63
	2	3 244 213	0.66	1.64
	3	3 271 682	0.6	1.66
NeRF	1	117 452	4.38	0.06
	2	115 771	4.39	0.06
	3	113 353	4.37	0.06
		Female T-pose		
SfM	1	2 255 242	0.62	1.66
	2	2 261 020	0.62	1.67
	3	2 259 960	0.62	1.67
NeRF	1	89 208	4.33	0.06
	2	89 387	4.33	0.06
	3	88 526	4.33	0.06



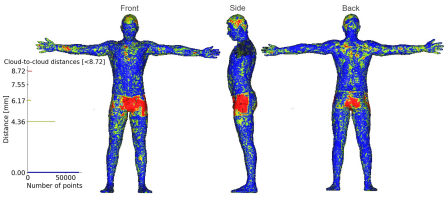
(a) SfM: Frame 2 to 1 comparison



(c) NeRF: Frame 2 to 1 comparison



(b) SfM: Frame 3 to 2 comparison

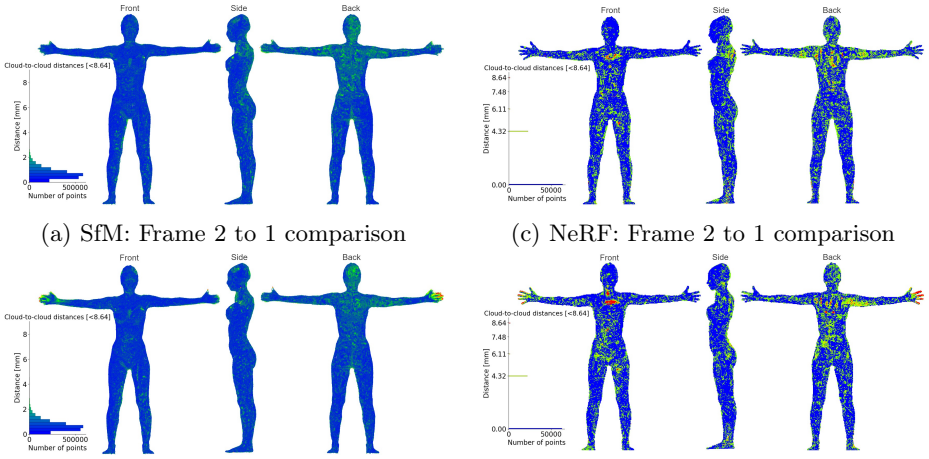


(d) NeRF: Frame 3 to 2 comparison

Fig. 6: Male T-pose: Temporal cloud-to-cloud comparison for the three consecutive reconstructions of the SfM and NeRF methods.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34

Temporal comparisons (Figure 6) show that SfM reconstructions had median cloud-to-cloud distances of 0.65 mm for the male and 0.65-0.64 mm for the female between consecutive frames. These values are close to the average distances between points for respective SfM clouds. The small median distances between frames indicate high temporal consistency for SfM reconstructions. In contrast, NeRF reconstructions showed distinct patterns reflecting voxel-based discretisation, where the voxel grid resolution inherently constrains point placements. In the male case, between frames 1 and 2, 58% of point distances are zero, and 30% matched the voxel size. For frames 2 and 3, these values are 59% and 31%. For the female subject, over 98% of points had distances smaller than twice the voxel size. Larger temporal deviations in NeRF occur primarily in poorly reconstructed regions such as the upper chest and, for the male subject, the shorts. As with SfM, increased distances are also observed in the hands, reflecting slight inter-frame motion. To provide a broader qualitative context, Figure 8 shows twelve consecutive frames, illustrating the temporal evolution of the reconstructions beyond the three-frame quantitative evaluation.



(a) SfM: Frame 2 to 1 comparison

(c) NeRF: Frame 2 to 1 comparison

(b) SfM: Frame 3 to 2 comparison

(d) NeRF: Frame 3 to 2 comparison

Fig. 7: Female T-pose: Temporal cloud-to-cloud comparison for the three consecutive reconstructions of the SfM and NeRF methods.

Reconstruction resolution Reconstruction resolution is assessed through histograms of distances to the 20 nearest neighbours (Figure 9) and neighbour counts within a 10 mm radius (Figure 10). SfM reconstructions show tighter distributions of nearest-neighbour distances accumulated in the first bins of the histogram, suggesting more evenly distributed points on the surface. NeRF point clouds exhibit broader, Gaussian-like distributions with more points hav-

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34

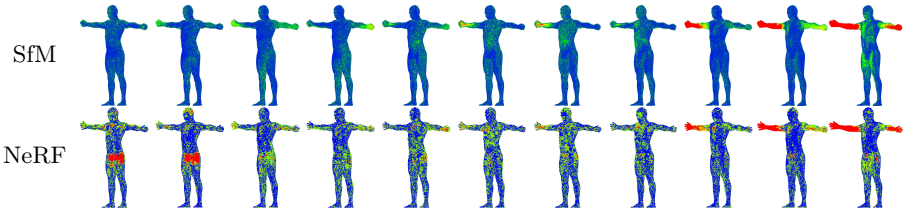


Fig. 8: Male T-pose: Temporal cloud-to-cloud comparison of twelve consecutive reconstructions.

ing higher and lower distances than the average. In both cases, SfM has a more even surface sampling, although some areas still exhibit reduced density.

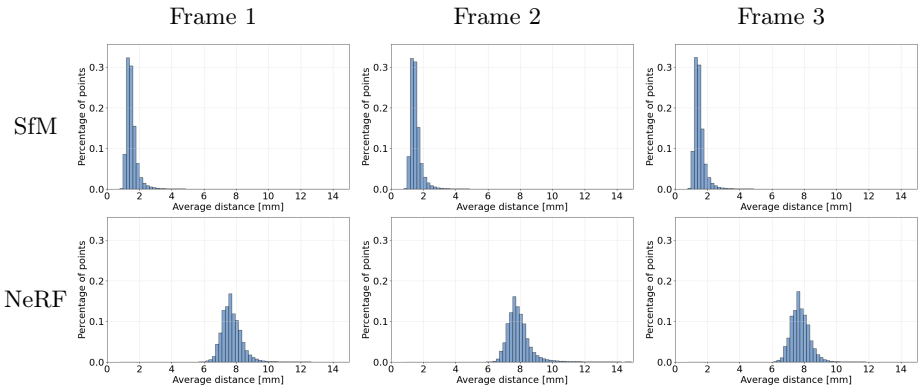


Fig. 9: Male T-pose: Histograms of distance to 20 nearest neighbours (<15 mm) for three consecutive time frames of SfM and NeRF reconstructions.

For the neighbour count analysis, SfM clouds are uniformly subsampled to match the number of points from the NeRF reconstruction result. Both methods show similar overall neighbour densities across the whole surface, with SfM point clouds having a narrower histogram distribution. In point clouds from both test cases, there is no significant number of points with extremely large or small numbers of neighbours, which indicates that most of the surface has good coverage of points. All of the histograms remain consistent across frames, showing good temporal stability when it comes to surface point distribution.

Surface smoothness Surface smoothness was evaluated using local neighbourhoods of increasing radii from 8 mm to 24 mm. Results for both cases are presented in Figure 11. Surface smoothness values increase for both reconstructions in a similar, nearly linear manner as the radius increases. SfM consistently

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34

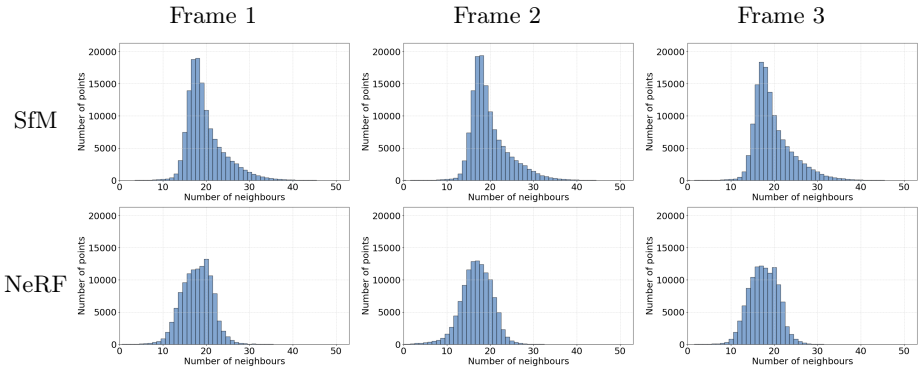
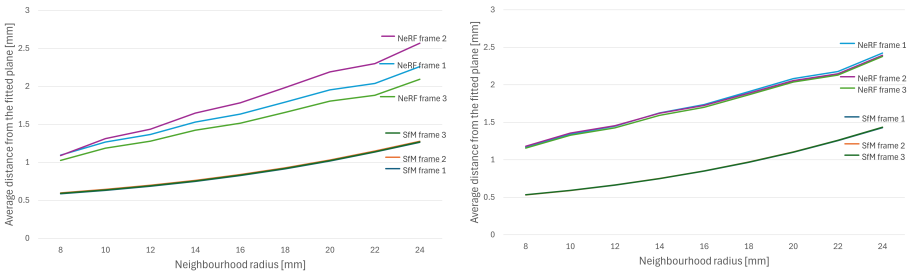


Fig. 10: Male T-pose: Number-of-neighbours histograms for three consecutive time frames of SfM and NeRF reconstructions.

produces smoother surfaces, as indicated by lower smoothness values across all radii. NeRF reconstructions, which contain approximately 1/30 of the points compared to SfM, exhibit higher surface roughness, especially in the male case's second frame, which has additional holes in the surface around the shorts.



(a) Male T-pose (b) Female T-pose

Fig. 11: Surface smoothness charts for increasing neighbourhood radius for three frames reconstructed with SfM and NeRF.

Temporal stability is high for SfM, with nearly identical smoothness curves across frames for both subjects. NeRF curves show greater variation, particularly with larger radii, indicating variability in local curvature over time.

5 Conclusions

This study provides a comprehensive comparison of SfM (carried out using Agisoft Metashape) and NeRF (implemented with Instant-NGP) reconstruction

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34

methods using real-world datasets of full-body scenes. The analysis evaluates reconstruction quality, differences between methods, and temporal stability across successive frames. While both methods successfully reconstructed the subjects, notable differences were observed in point cloud characteristics, surface smoothness, and the handling of fine structures, all of which are critical for accurately representing human body details.

SfM consistently produces denser point clouds with higher point counts, resulting in more detailed and smoother reconstructions of human body features. NeRF reconstructions, by contrast, give better geometry of small or occluded features but exhibit lower point densities and discrete offsets due to volumetric sampling, leading to rougher surfaces. Table 2 summarises the performance of both methods across four practical criteria relevant to general-purpose 3D human body reconstruction.

Table 2: Comparison of reconstruction criteria between SfM and NeRF.

Criterion	**SfM**	**NeRF**
Small, protruding structure preservation	Missed or merged fingers and ears	**Separate fingers and ears**
Surface smoothness	**45% higher than NeRF, consistent across frames**	Frame-dependent
Geometric temporal stability	**Median differences close to the point spacing**	10–27% of points exhibit larger variation than double point spacing
Reconstruction time	5 minutes	**3 minutes**

A summary of the performance of both reconstruction methods is presented below.

Cloud-to-cloud comparison. Discrepancies between methods are most apparent in complex or highly detailed regions. SfM demonstrates better temporal coherence, with differences primarily being due to subject movement rather than reconstruction instability.

Reconstruction resolution. SfM produces denser and more uniformly distributed point clouds, while NeRF has greater variation in resolution. Both methods maintain temporal stability in resolution across frames.

Surface smoothness. SfM reconstructions are smoother and more temporally stable. NeRF exhibits rougher surfaces due to its volumetric nature and shows more frame-to-frame variation in surface smoothness.

Temporal stability. Both methods are temporally stable, but SfM consistently outperforms NeRF in maintaining coherent reconstructions over time, making it more suitable for applications such as animating human motion.

Overall, the results show that Neural Radiance Fields offer complementary strengths rather than a direct replacement for SfM in 3D human reconstruction.

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34

While SfM remains superior in surface detail and temporal stability, NeRF surpasses it when it comes to reconstructing fine and partially occluded structures. However, the latter's potential edge regarding modelling semi-transparent or reflective materials was not assessed in this study because such properties were not present in the evaluated datasets. Future work could explore hybrid approaches that combine the strengths of both methods to improve reconstruction quality and robustness for human subjects.

Acknowledgments. The publication was financed in part by a subsidy from the Warsaw University of Technology. The funding source had no involvement in the design, execution, or scientific writing of the study. The authors wish to thank Mnemosis S.A. for providing the datasets used in this research.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Agarwal, S., Furukawa, Y., Snavely, N., Simon, I., Curless, B., Seitz, S.M., Szeliski, R.: Building rome in a day. Commun. ACM **54**(10), 105–112 (Oct 2011). https://doi.org/10.1145/2001269.2001293
2. Agarwal, S., Snavely, N., Simon, I., Seitz, S.M., Szeliski, R.: Building rome in a day. In: 2009 IEEE 12th International Conference on Computer Vision. pp. 72–79 (2009). https://doi.org/10.1109/ICCV.2009.5459148
3. Barron, J.T., Mildenhall, B., Tancik, M., Hedman, P., Martin-Brualla, R., Srinivasan, P.P.: Mip-NeRF: A multiscale representation for anti-aliasing neural radiance fields. In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 5835–5844 (2021). https://doi.org/10.1109/ICCV48922.2021.00580
4. Barron, J.T., Mildenhall, B., Verbin, D., Srinivasan, P.P., Hedman, P.: Zip-NeRF: Anti-aliased grid-based neural radiance fields. ICCV (2023)
5. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**, 239–256 (1992)
6. Choi, S., Ashdown, S.P.: 3D body scan analysis of dimensional change in lower body measurements for active body positions. Textile Research Journal **81**(1), 81–93 (2011). https://doi.org/10.1177/0040517510377822
7. Croce, V., Billi, D., Caroti, G., Piemonte, A., De Luca, L., Véron, P.: Comparative assessment of neural radiance fields and photogrammetry in digital heritage: Impact of varying image conditions on 3D reconstruction. Remote Sensing **16**(2) (2024). https://doi.org/10.3390/rs16020301
8. Haleem, A., Javaid, M.: 3D scanning applications in medical field: A literature-based review. Clinical Epidemiology and Global Health **7**(2), 199–210 (2019). https://doi.org/10.1016/j.cegh.2018.05.006
9. Krasowicz, K., Michoński, J., Liberadzki, P., Sitnik, R.: Monitoring improvement in infantile cerebral palsy patients using the 4DBODY system—a preliminary study. Sensors **20**(11) (2020). https://doi.org/10.3390/s20113232
10. Li, Z., Müller, T., Evans, A., Taylor, R.H., Unberath, M., Liu, M.Y., Lin, C.H.: Neuralangelo: High-fidelity neural surface reconstruction. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2023)

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34

11. Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., Black, M.J.: SMPL: a skinned multi-person linear model. ACM Trans. Graph. **34**(6) (Oct 2015). https://doi.org/10.1145/2816795.2818013
12. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60**, 91–110 (2004)
13. Martin-Brualla, R., Radwan, N., Sajjadi, M.S.M., Barron, J.T., Dosovitskiy, A., Duckworth, D.: NeRF in the wild: Neural radiance fields for unconstrained photo collections. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp. 7206–7215 (2021). https://doi.org/10.1109/CVPR46437.2021.00713
14. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: NeRF: representing scenes as neural radiance fields for view synthesis. Commun. ACM **65**(1), 99–106 (dec 2021). https://doi.org/10.1145/3503250
15. Müller, T., Evans, A., Schied, C., Keller, A.: Instant neural graphics primitives with a multiresolution hash encoding. ACM Trans. Graph. **41**(4), 102:1–102:15 (Jul 2022). https://doi.org/10.1145/3528223.3530127
16. Nielsen, M.S., Nikolov, I., Kruse, E.K., Garnæs, J., Madsen, C.B.: Quantifying the influence of surface texture and shape on structure from motion 3D reconstructions. Sensors **23**(1) (2023). https://doi.org/10.3390/s23010178
17. Petrovska, I., Jutzi, B.: Vision through obstacles—3D geometric reconstruction and evaluation of neural radiance fields (NeRFs). Remote Sensing **16**(7) (2024). https://doi.org/10.3390/rs16071188
18. Remondino, F., El-Hakim, S.: Image-based 3D modelling: A review. The Photogrammetric Record **21**(115), 269–291 (2006). https://doi.org/10.1111/j.1477-9730.2006.00383.x
19. Remondino, F., Karami, A., Yan, Z., Mazzacca, G., Rigon, S., Qin, R.: A critical analysis of NeRF-based 3D reconstruction. Remote Sensing **15**(14) (2023). https://doi.org/10.3390/rs15143585
20. Schönberger, J.L., Frahm, J.M.: Structure-from-motion revisited. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 4104–4113 (2016). https://doi.org/10.1109/CVPR.2016.445
21. Seitz, S., Curless, B., Diebel, J., Scharstein, D., Szeliski, R.: A comparison and evaluation of multi-view stereo reconstruction algorithms. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06). vol. 1, pp. 519–528 (2006). https://doi.org/10.1109/CVPR.2006.19
22. Sitnik, R., Witkowski, M.: Locating and tracing of anatomical landmarks based on full-field four-dimensional measurement of human body surface. Journal of Biomedical Optics **13**(4), 044039 (2008). https://doi.org/10.1117/1.2960017
23. Snavely, N., Seitz, S.M., Szeliski, R.: Photo tourism: exploring photo collections in 3D. ACM Trans. Graph. **25**(3), 835–846 (Jul 2006). https://doi.org/10.1145/1141911.1141964
24. Stewart, G.W.: On the early history of the singular value decomposition. SIAM Review **35**(4), 551–566 (1993). https://doi.org/10.1137/1035134
25. Ullman, S.: The interpretation of structure from motion. Proceedings of the Royal Society of London. Series B, Biological Sciences **203**(1153), 405–426 (1979)
26. Wang, P., Liu, L., Liu, Y., Theobalt, C., Komura, T., Wang, W.: NeuS: Learning neural implicit surfaces by volume rendering for multi-view reconstruction. NeurIPS (2021)

ICCS Camera Ready Version 2026
To cite this paper please use the final published version:
DOI: 10.1007/978-3-032-29921-5_34