In-Situ Gaussian Splatting-generated 3D Thermal Mesh Visualization for Urban Trees in Augmented Reality

Chaimaa Delasse [1,2], Dario Billi [2,3], I Gede Mahendra Darmawiguna[4], Kadek Ananta Satriadi[4], Arnadi Murtiyoso [2], Hélène Macher [2], Vincent Lecomte[2], Rafika Hajji [1], Tania Landes [2]

[1] College of Geomatic Sciences and Surveying Engineering, Institute of Agronomy and Veterinary Medicine (IAV), Rabat 6202, Morocco - (c.delasse, r.hajji)@iav.ac.ma
[2] University of Strasbourg, INSA Strasbourg, CNRS, ICube UMR 7357 Laboratory, 67000 Strasbourg, France - (chaimaa.delasse, dario.billi, arnadi.murtiyoso, helene.macher, vincent.lecomte, tania.landes)@insa-strasbourg.fr
[3] Department of Civil and Industrial Engineering, ASTRO Laboratory, University of Pisa, Largo Lucio Lazzarino, 56122 Pisa, Italy; dario.billi@phd.unipi.it
[4] Faculty of Information Technology, Monash University, Australia – (i.darmawiguna, kadek.satriadi)@monash.edu

Keywords: Thermal Infrared (TIR) images, 3D Gaussian Splatting (3DGS), 3D tree modelling, Augmented Reality (AR).

Abstract

Urban trees provide critical ecosystem services in dense city environments, yet current workflows for monitoring their thermal behaviour remain confined to 2D desktop-based analysis with no three-dimensional spatial context or field-deployable visualization capability. This paper presents a complete pipeline for in-situ 3D thermal mesh visualization of urban trees in Augmented Reality (AR), combining Thermal InfraRed (TIR) image acquisition, Gaussian Splatting-based mesh reconstruction, quantitative validation, and mobile AR deployment. TIR images of a *Tilia tomentosa* acquired with a FLIR T560 camera are preprocessed with a standardized false-colour palette and fed into the MILo (Mesh-In-the-Loop Gaussian Splatting) framework to reconstruct a thermally attributed 3D mesh. Geometric evaluation against a Z+F IMAGER 5016 TLS reference using the M3C2 algorithm demonstrates that MILo recovers 13.5 times more canopy geometry than traditional multi-view stereo under thermal imagery, with a standard deviation of 4.0 cm. A colourmap inversion procedure recovers per-vertex temperature estimates from the GS-derived mesh colours, yielding a mean absolute difference of 0.7°C against direct T-Cam measurements (thermal camera mounted on the laser scanner), within the combined instrument accuracy of both sensors. The resulting thermal Gaussian Splat was deployed in a custom Android AR application supporting hybrid marker-based and GPS-based spatial anchoring for in-situ visualization. These results demonstrate the technical feasibility of GS-based thermal reconstruction and mobile AR as a medium for communicating three-dimensional canopy thermal information to educators and urban forestry practitioners.

1. Introduction

Urban trees play a critical role in regulating microclimates and enhancing thermal comfort, providing shade, evapotranspirative cooling, and supporting biodiversity in increasingly dense urban environments. Quantifying their thermal behaviour in three dimensions is essential for urban planners, arborists, and climate adaptation stakeholders who need to assess canopy health, detect early signs of water stress, and monitor the cooling performance of urban greenery. Thermal InfraRed (TIR) remote sensing, operating in the 8–14 µm wavelength range, is uniquely suited to this task: it captures surface temperature variations that reflect changes in transpirational cooling linked to stomatal conductance and plant water status (Smigaj et al., 2023). However, current operational workflows for urban tree thermal monitoring rely almost exclusively on 2D thermal imagery analysed on desktop workstations, offering no three-dimensional spatial context and no means for immersive, field-based exploration of canopy thermal patterns. This disconnect between data acquisition and communication represents a significant barrier to the uptake of thermal monitoring in urban forestry practice.

Recent advances in neural rendering, particularly 3D Gaussian Splatting (3DGS) (Kerbl et al., 2023), have opened new possibilities for photorealistic 3D reconstruction from multi-view imagery, enabling real-time rendering of complex scenes. The MILo (Mesh-In-the-Loop Gaussian Splatting) algorithm (Guédon et al., 2025) extends this framework, producing compact, high-quality surface meshes suitable for deployment in Augmented Reality (AR) environments. In parallel, the combination of Terrestrial Laser Scanning (TLS) with integrated thermal cameras has demonstrated the feasibility of generating ground truth georeferenced 3D thermal point clouds of trees. Despite these advances, three critical gaps remain unaddressed in the literature: (1) GS-based methods have not been applied to TIR imagery of natural vegetation, where the texture-poor nature of thermal images challenges feature-based reconstruction; (2) no evaluation of GS-derived thermal meshes against reference TLS thermal data exists for urban trees; and (3) no field-deployable AR application currently allows practitioners to visualize 3D thermal tree models in situ with perceptually calibrated temperature-to-colour mappings.

This paper addresses these gaps through a pipeline combining TIR image acquisition, GS-based thermal mesh reconstruction, quantitative validation against TLS reference data, and mobile AR visualization. Specifically, our contributions are: (1) we demonstrate for the first time the application of the MILo GS framework to TIR imagery of urban trees, generating thermal meshes with 3D canopy coverage significantly superior to those reconstructed from traditional multi-view stereo (MVS) approaches; (2) we provide a quantitative geometric and thermal evaluation of the reconstructed meshes against a Z+F IMAGER 5016 + T-Cam reference point cloud, including a colourmap inversion methodology for recovering per-vertex temperature estimates from GS-derived mesh colours; and (3) we implement the resulting thermal 3D models in a hybrid marker-based and GPS-georeferenced Android AR application, enabling dual-mode immersive field visualization of canopy thermal patterns for education, public outreach, and spatial analysis. The remainder of the paper is structured as follows: Section 2 reviews related work on TIR remote sensing of vegetation, GS-based surface reconstruction, and AR visualization in forestry contexts; Section 3 describes the multi-sensor data collection campaign conducted on an individual tree; Section 4 details the methodology covering TIR image preprocessing, MILo-based thermal mesh reconstruction, AR implementation, and the evaluation framework; Section 5 presents the geometric, thermal, and AR visualization results; Section 6 discusses the implications, limitations, and generalizability of the findings; and Section 7 concludes with directions for future work.

2. Related work

TIR remote sensing has proven valuable for detecting physiological stress responses in vegetation, as reviewed by Smigaj et al. (2023). While thermal imaging has been successfully applied in agriculture for indirect stress detection, its use in urban forest environments remains challenging due to canopy structural complexity, variable sun exposure, and confounding environmental factors. Recent UAV-based thermal studies have demonstrated the detection of disease-induced temperature increases, drought stress, and parasite infestations at the individual tree level, though interpretation of thermal differences requires careful consideration of factors such as tree density, canopy architecture, and microclimate effects.

In urban context, TIR measurements are increasingly used to characterise the role of street trees in microclimate regulation and heat mitigation, by mapping spatio-temporal variations in leaf and trunk temperatures and their influence on surrounding built surfaces. Lecomte et al. (2022) developed a methodology for combining TIR images with TLS data to create 3D thermal point clouds of trees. This work demonstrated that 3D thermal models can facilitate monitoring of spatial and temporal temperature variations across tree structures. Similar strategies have been pursued using UAV platforms, combining 3D geometric reconstruction and TIR payloads to enable high-resolution thermal mapping in precision viticulture (Santesteban et al., 2017). However, these approaches typically rely on dedicated multi-sensor platforms and post-processing pipelines that are difficult to deploy in operational urban forestry contexts, and they do not produce the compact mesh representations required for real-time AR rendering.

Recent advances in radiance fields, particularly 3DGS (Kerbl et al., 2023), have revolutionized 3D visualization from multi-view imagery. 3DGS represents scenes as collections of anisotropic 3D Gaussians with learnable attributes including position, covariance, opacity, and colour. Unlike implicit neural representations such as NeRF (Mildenhall et al., 2022), 3DGS enables real-time differentiable rendering through explicit Gaussian primitives, achieving state-of-the-art visual fidelity on standard benchmarks while remaining amenable to direct geometric manipulation. Several works have extended 3DGS toward surface reconstruction: SuGaR (Guédon and Lepetit, 2024) introduced regularization terms to flatten Gaussians onto surfaces and extract meshes via Poisson reconstruction, while 2DGS (Huang et al., 2024) enforced a collapse of Gaussians into 2D surfels, promoting geometrically consistent rendering. Most recently, MILo (Guédon et al., 2025) introduced a mesh-in-the-loop framework that directly extracts differentiable meshes from Gaussian parameters during optimization, enabling bidirectional consistency between volumetric and surface representations. By computing signed distance values from 3D Gaussians, MILo generates high-quality meshes with significantly optimized vertices. This capability is particularly relevant to AR applications that require efficient mesh representations. However, while GS-based methods have demonstrated impressive results for RGB scene reconstruction, they have only very recently been extended to thermal and multi-modal imagery (Chen et al., 2026) and have not yet been applied to the 3D modelling of individual trees from TIR imagery. This presents both technical challenges and opportunities for field-deployable AR visualization of vegetation thermal patterns.

Situated data visualization places both users and data within their physical context, allowing direct access to the data referent (White and Feiner, 2009; Willett et al., 2016). While it can be achieved in simulated environments such as virtual reality (Satriadi et al., 2023), a combination of AR and real environment is a common approach (Bressa et al., 2021). AR technologies are beginning to find applications in urban forestry, though predominantly in fully virtual environments. Zürcher et al. (2023) demonstrated an immersive VR system integrating lidar-derived point clouds with 360° imagery for forest monitoring in Germany's Eifel National Park. Their evaluation with domain experts revealed that while immersive visualization significantly enhanced spatial understanding of complex forest structures and facilitated stakeholder communication, users preferred these tools for training and education rather than operational forestry. The study highlighted the perceptual advantages of stereoscopic 3D-in-3D visualization over traditional desktop-based approaches for comprehending intricate canopy architectures. For AR specifically, marker-based tracking in frameworks such as Vuforia provides robust spatial anchoring in structured outdoor environments (Garg et al., 2025), while GPS-based georeferencing enables untethered field deployment at the cost of metric precision (Schall et al., 2009). Hybrid approaches that integrate global positioning methods (e.g., GPS) with vision-based tracking have been proposed to balance spatial precision and practical usability under varying field conditions (Reitmayr & Drummond, 2006). This hybrid strategy is adopted in the present work. To the best of our knowledge, no previous study has combined GS–based thermal representations with a hybrid AR tracking architecture for in-situ urban tree thermal visualization.

3. Data Collection

A multi-sensor acquisition campaign was conducted at the *Palais Universitaire* courtyard, University of Strasbourg, France, focusing on a *Tilia tomentosa* (silver linden) tree. This species and site were selected for their accessibility and the tree's well-developed canopy structure. The dataset comprises TIR images acquired with a FLIR T560 thermal camera (640×480-pixel TIR resolution, 2 °C accuracy) mounted on a tripod to ensure stable acquisition. Images were captured at multiple viewing angles around the tree to ensure comprehensive coverage of the canopy structure. The FLIR T560 is equipped with an integrated RGB camera, allowing for spatially aligned multispectral acquisitions from the same viewpoints. Camera interior orientation parameters and the relative offset between the RGB and TIR sensors were calibrated in advance using a bi-spectral aluminum checkerboard target detectable in both the visible and TIR domains. A total of 112 RGB images and 112 TIR images (Figure 1) were acquired.

At the same time, as reference data for validation, we acquired terrestrial laser scans using a Zoller+Fröhlich (Z+F) IMAGER 5016 equipped with an integrated thermal camera (T-Cam 382×288-pixel, 2 °C accuracy). The laser scanner provides geometric accuracy of 1-3 mm at 10 m range with a point spacing of approximately 6 mm at the same distance. Five scan stations were distributed around the tree to minimise occlusions. The thermal camera fixed on the laser scanner enables direct

https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-1021-2026 | © Author(s) 2026. CC BY 4.0 License.

generation of thermally attributed point clouds through synchronized acquisitions. All data were collected on Thursday, 12 May 2022, around solar noon under partially overcast conditions with no direct sunlight and light wind.

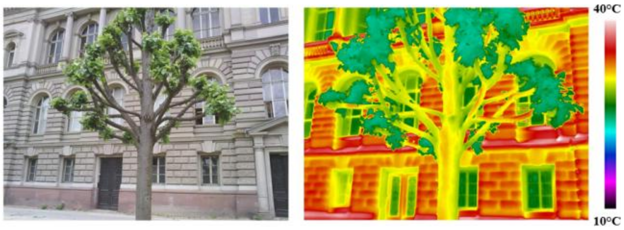


Figure 1. RGB (left) and TIR (right) image samples of the study tree.

4. Methodology

The pipeline consists of three main stages: (1) thermal image preprocessing, (2) thermal mesh reconstruction via MILo, and (3) AR implementation and visualization. All computational processing was performed on a consistent hardware configuration to ensure reproducibility: an NVIDIA GeForce RTX 4090 GPU (24 GB VRAM) and an AMD Ryzen 9 7950X 16-core CPU.

4.1 TIR Image Preprocessing

Raw TIR images from the FLIR T560 are stored as radiometric JPEG files encoding per-pixel temperature matrices. Prior to any photogrammetric processing, all 112 TIR images were standardized in FLIR Research Studio (Teledyne FLIR). Radiometric corrections were applied uniformly across the dataset by setting consistent object parameters (emissivity, reflected temperature, atmospheric temperature, relative humidity, and acquisition distance) to match the conditions at the time of acquisition. A fixed temperature scale of 10–40 °C was then applied to all images using the FLIR Rainbow High Contrast (RHC) colour palette, converting each temperature matrix into a false-colour RGB image with a consistent per-pixel colour-to-temperature mapping across the entire dataset. This standardization step is critical for MILo optimization because GS-based methods learn per-Gaussian colour attributes by minimizing a photometric loss with respect to input images, and radiometric consistency across viewpoints is a prerequisite for stable convergence and spatially coherent temperature encoding in the reconstructed mesh.

4.2 Thermal Mesh Reconstruction via MILo

Structure-from-Motion (SfM) camera orientation was computed in Agisoft Metashape using the multi-camera system option, which treats the FLIR T560 as a rigidly coupled RGB–TIR sensor pair (master/slaves). The RGB images served as the primary channel for automated tie-point detection and matching, owing to their superior spatial resolution and chromatic texture compared with the TIR images. The pre-calibrated interior orientation parameters of both sensors, along with the relative translation and rotation between the RGB and TIR optical centres, were exploited. Once RGB camera poses were estimated and compensated via bundle adjustment, the exterior orientation of each TIR image was deduced analytically from the corresponding RGB pose via the inter-sensor offset.

Starting from the SfM sparse point cloud and camera orientations, MILo initializes Gaussians at the sparse point positions and iteratively refines their parameters by minimizing a photometric loss against the input TIR images. Simultaneously, signed distance values are computed from the Gaussian parameters, and a differentiable mesh is extracted via Delaunay triangulation, with bidirectional consistency enforced between the volumetric Gaussian representation and the surface mesh throughout optimization.

The MILo training parameters were selected to account for the specific constraints of the TIR dataset. The RaDe-GS rasterizer was retained as the recommended option within the framework, given its ability to produce geometrically consistent surfaces from complex organic structures. High-resolution mesh configuration was adopted to preserve the fine branching and foliar detail of the canopy, while input images were downsampled to half their native resolution to reduce memory overhead without compromising Gaussian placement accuracy. The decoupled appearance model was enabled to handle the inter-frame radiometric variability inherent to TIR imagery, where non-uniformity correction cycles, gain drift, and ambient temperature fluctuations introduce exposure variations uncorrelated with scene geometry, preventing these artefacts from being absorbed into the geometric representation.

The TIR images fed into MILo encode temperature through the FLIR RHC palette, so the spherical harmonic coefficients learned by the optimization encode the false-colour thermal appearance of the canopy rather than a raw temperature scalar. The view-independent component of this colour (recovered from the zero-order spherical harmonic (DC) coefficients) thus encodes the local temperature as a perceptually uniform colour, which can subsequently be inverted. The reconstruction produced a thermal mesh with approximately 3 million vertices for the full-resolution output.

For AR deployment, the original Gaussian splat reconstruction was cleaned to isolate a single target tree from the surrounding scene. The resulting tree-only model was used in the AR application, while retaining the principal thermal patterns of the canopy, trunk, and major branches.

4.3 AR Visualization

The AR visualization part of our pipeline is a 3DGS viewer for Android. It was developed in Unity to visualize thermal splat models in an AR environment. Rather than rendering a conventional mesh, the system loads and displays a 3DGS .ply file. Within Unity, a custom PLY loader adapted from the *UnityGaussianSplatting* implementation by Pranckevičius (2025) was used to parse the .ply file at runtime and pass the Gaussian attributes to a GPU-accelerated rendering pipeline. The application integrates two complementary spatial anchoring modes, marker-based and GPS-based, to support both precise inspection and flexible field deployment. In the primary marker-based mode, Vuforia Engine (PTC Inc., 2025) was used for image-target tracking, where a printed fiducial marker served as the spatial anchor and the thermal splat model was registered to this target so that it appeared when the marker was detected by the device camera. Within this mode, the AR scene was configured using Vuforia’s ARCamera and Image Target, and the resulting *SplatRenderer* object was attached to the image target to maintain spatial alignment during detection and tracking. In the secondary GPS-based mode, the splat model was positioned using coordinates obtained from Google Maps and rendered at the tree location, enabling in-situ visualization without requiring a physical marker. This location-based placement was implemented in Unity using ARCore Extensions for AR

Foundation to support the mobile AR session, while Cesium for Unity was used to handle geospatial placement within the scene.

Real-time visualization on mobile hardware was supported through the GPU-accelerated rendering pipeline provided by the *UnityGaussianSplatting* framework. For AR deployment, the original thermal Gaussian splat was cleaned to isolate a single target tree from the surrounding scene, yielding a tree-only model with 31,719 splats. The thermal colour information already encoded in the splat representation was preserved in the Unity visualization, and an interactive temperature scale overlay was included to help users interpret approximate temperature values directly in the field. The implementation was developed using Unity 6000.3.5f2, ARCore Extensions for AR Foundation v1.52.0, Vuforia Engine 11.4.4, Cesium for Unity 1.16.0, and *UnityGaussianSplatting* v1.1.1.

as unsigned 8-bit RGB values (range 0-255) encoding temperature indirectly through the FLIR RHC palette.

To recover per-vertex temperature estimates, a colourmap inversion procedure was applied. This produced a lookup table (LUT) of N = 2,048 entries mapping normalized positions t ∈ [0, 1] to RGB triplets. For each mesh vertex with colour (R, G, B), the nearest LUT entry is identified by minimizing Euclidean distance in RGB space as described in equation. 1:

$$i* = \arg\min_i \; \|(Rv - Ri, Gv - Gi, Bv - Bi)\|2 \qquad (1)$$

The temperature for vertex v is then recovered using equation 2:

$$Tv = Tmin + (i*/N) \times (Tmax - Tmin) \qquad (2)$$

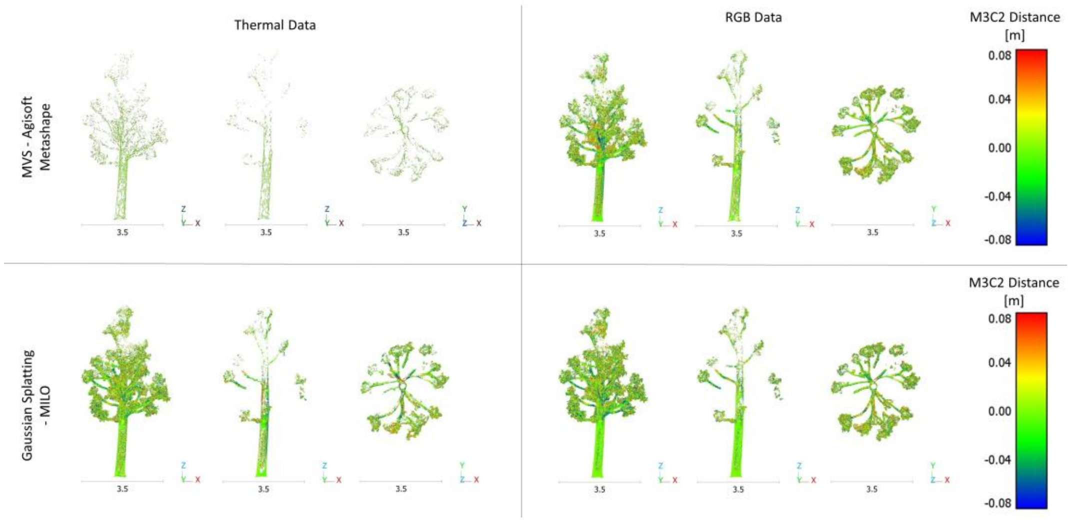


Figure 2. M3C2 signed distance maps comparing Metashape (top) and MILo (bottom) reconstructions of the tree against the TLS reference, for thermal (left column) and RGB (right column) datasets. Colour scale: ±8 cm.

4.4 Evaluation Framework

The reconstructed thermal mesh vertices were evaluated geometrically against the Z+F T-Cam reference point cloud using the Multiscale Model-to-Model Cloud Comparison (M3C2) algorithm in CloudCompare. M3C2 computes signed distances between two point sets along locally estimated surface normals, providing both a measure of mean geometric deviation and a confidence interval accounting for point cloud roughness. An analysis range of ±8 cm was adopted to facilitate interpretation of M3C2 deviation, accounting for natural tree movement (e.g., wind-induced displacement of branches and foliage) and the inherent challenges of comparing dynamic organic structures with TLS ground truth. Completeness (proportion of the tree geometry successfully reconstructed) and correctness (absence of hallucinated geometry) were assessed qualitatively from deviation maps in lateral and planimetric cross-section views (Figure 2), and quantitatively via mean error, standard deviation, and outlier rate (Table 1). For reference, results from Agisoft Metashape (traditional MVS) are reported alongside MILo to contextualise performance on both RGB and TIR imagery.

Unlike the geometric evaluation, the thermal evaluation cannot rely on direct temperature comparisons extracted from the MILo output. The MILo pipeline stores per-vertex colour information

5. Results

5.1 Qualitative Geometric Evaluation

Visual inspection of the M3C2 deviation maps reveals clear differences in reconstruction completeness between MILo and Metashape across both imaging modalities (Figure 2).

For the RGB dataset (Figure 2, right column), both methods produce predominantly green coverage in lateral and sectional views, confirming that geometric deviations are largely confined to the central portion of the ±8 cm analysis range. Metashape's reconstruction is nonetheless visibly sparser in interior canopy regions, where branch overlap creates ambiguities for feature matching, resulting in white patches and gaps in the sectional slices.

The contrast is far more pronounced for the thermal dataset (Figure 2, left column). MILo maintains continuous volumetric coverage across all views, while Metashape's thermal reconstruction is dominated by blank regions indicating failed reconstruction or out-of-range deviations. In terms of completeness, MILo achieves near-full canopy coverage across both imaging modalities: lateral views show continuous coloured surfaces from the lower trunk through the upper canopy, and planimetric slices confirm that interior branching structure is represented throughout the tree volume. Metashape's RGB reconstruction is largely complete at the canopy periphery but exhibits visible gaps in interior regions where branch overlap creates feature-matching ambiguities. For the thermal dataset, Metashape's completeness collapses almost entirely: planimetric cross-sections show only sparse peripheral clusters, with the interior canopy structure nearly absent, while MILo maintains coherent volumetric coverage comparable to its RGB reconstruction. In terms of correctness, neither method produces significant hallucinated geometry beyond the tree envelope. Out-of-range deviations in the deviation maps are spatially consistent with known occlusion zones and wind-induced branch displacement between acquisitions rather than systematic reconstruction artefacts.

5.2 Quantitative Geometric Evaluation

The quantitative metrics (Table 1) confirm the trends observed qualitatively and reveal important nuances that vary by imaging modality.

For the RGB dataset, MILo and Metashape achieve comparable geometric accuracy, with standard deviations of 3.8 cm and 4.1 cm, respectively, and negligible mean errors for both, indicating no systematic bias. Outlier rates are also similar (55% vs. 59%), reflecting the shared challenge of reconstructing complex branching and dense foliage. The critical distinction lies in coverage: MILo generates approximately 2.8 times more points (1,415,283 vs. 506,005), at the cost of a roughly 57-fold increase in processing time.

For the thermal dataset, this trade-off decisively favours MILo. While Metashape achieves a marginally lower standard deviation (2.7 cm vs. 4.0 cm) in regions where reconstruction succeeds, this apparent advantage is undermined by an outlier rate of 80 % (nearly 24 percentage points above MILo's 56 %) and a point count of just 52,883 compared to MILo's 714,724. Metashape thus captures less than 8 % of the geometric coverage produced by MILo, rendering its speed advantage (0.13 min vs. 17.05 min) practically irrelevant. The root cause lies in the texture-poor nature of thermal imagery: unlike RGB scenes, tree canopies in the thermal spectrum exhibit minimal differentiation, undermining the feature correspondence assumptions on which traditional MVS relies. MILo's GS optimization, which does not depend exclusively on discrete feature matches, is better suited to interpolate geometric information across views under these conditions. Taken together, these results confirm that for thermal imagery, MILo's extended computational cost represents a necessary investment rather than a simple accuracy-speed trade-off.

		N Points	Mean Error (cm)	Std (cm)	Outliers (%)	Time (min)
TIR	MILo	714,724	0.0	4.0	56	17.05
	MVS	52,883	-0.1	2.7	80	0.13
RGB	MILo	1,415,283	-0.2	3.8	55	40.17
	MVS	506,005	-0.1	4.1	59	0.70

Table 1. Quantitative metrics extracted using M3C2 for both datasets.

5.3 Thermal Evaluation

Per-vertex temperature estimates were recovered via the colourmap inversion described in Section 4.4. Prior to analysis, 6,246 vertices (0.9% of the total 714,724) were excluded as sky-mixel artefacts, pixels at the leaf-sky boundary that integrate thermal radiance from both the leaf surface and the cold sky background. The recovered distributions were then compared against direct T-Cam measurements from the Z+F reference scanner. Thermal distributions were compared using descriptive statistics (Table 2) and the two-sample Kolmogorov–Smirnov (KS) test. The KS statistic $D = \max|F1(x) - F2(x)|$ measures the maximum absolute difference between the empirical cumulative distribution functions (CDFs) of the two datasets, with values close to 0 indicating agreement between the distributions.

Dataset	Mean (°C)	Std (°C)	Median (°C)
Z+F T-Cam	29.5	4.5	29.2
MILo	29.0	3.6	27.5

Table 2. MILo vs. Z+F T-Cam reference - Thermal statistics.

The MILo mesh yields a mean temperature of 29.0 °C, compared with 29.5 °C for the TLS reference, a mean absolute difference of 0.7°C, well within the ±2 °C manufacturer's accuracy for both instruments. The KS statistic D = 0.227 indicates moderate agreement between the distributions. To investigate whether this distributional divergence reflects spatially structured rather than uniform bias, we performed a spatial analysis of temperature profiles. As illustrated in Figure 3, the spatial distribution of temperature differences ($\Delta T = T\_MILo - T\_ZF$) was mapped across the 3D tree canopy. The unified colour scale, ranging from -2.0 °C to +4.5 °C, highlights these localized thermal discrepancies. Specifically, warmer colours (yellow to red) indicate regions where the MILo-derived temperatures are overestimated compared to the reference, which are predominantly located at the uppermost part of the crown and along the outermost radial foliage. Conversely, cooler colours (light to dark blue) represent areas of underestimation, often found in the lower to middle sections of the canopy and deep within the interior near the central trunk axis.

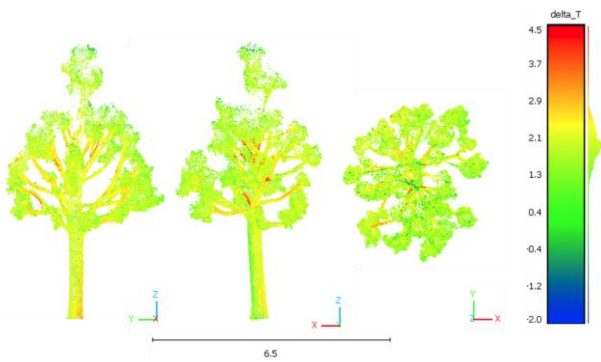


Figure 3. Spatial distribution of temperature differences (T_MILo−T_ZF) in °C, between the MILo and Z+F T-Cam reference. (Left) South view, (Center) East view, (Right) Top view.

To characterise the spatial structure of the thermal bias between MILo and the ZF reference, temperature was analysed as a function of two geometric descriptors: normalised height along the tree's vertical axis (0 = trunk base, 1 = apex) and radial distance from the estimated trunk axis in the horizontal plane. Results are presented in Figure 4.

For the vertical profile (Figures 4a, 4c), both datasets reproduce the expected physical vertical temperature gradient: the trunk base is the warmest zone, with temperatures decreasing progressively toward the upper canopy. The profiles are well-aligned in the mid-canopy (normalised height 0.3–0.7), where both instruments have good mutual visibility. However, a positive bias of MILo relative to ZF is observed across the full vertical range, with the strongest overestimation at the trunk base (+3.3 °C for h < 0.2). This is consistent with the close-range camera geometry of the FLIR T560 acquisition: near-ground viewpoints capture the sun-heated bark surface at high spatial resolution, whereas the T-Cam integrates a wider field of view at the same height, including cooler surrounding ground and lower canopy elements. The overall mean vertical bias is +1.6 °C.

The radial profile analysis (Figure 4b, 4d) reveals that MILo and ZF agree well in the inner canopy (r < 1 m, mean bias −0.1 °C), where both instruments have unobstructed views of the trunk and main branches. In the outer canopy (r > 1.5 m), however, MILo is systematically cooler than ZF by a mean of −3.0 °C. This outer-canopy cold bias directly explains the narrower standard deviation in MILo's global temperature distribution (3.6 °C vs. 4.5 °C): the camera-based acquisition captures the outer leaf surface predominantly from a small number of viewing directions, where shaded (cooler) leaf faces dominate the sampled signal. The ZF scanner, deployed from five stations spaced 10 m apart around the tree, achieves far better angular penetration of the canopy perimeter, sampling a more complete mix of illuminated and shaded leaf surfaces. Residual sky-mixel contamination at the outer canopy boundary likely contributes an additional negative bias at large radii. Taken together, the radial profile confirms that the global KS statistic of D = 0.227 reflects a genuine geometric acquisition effect rather than a calibration artefact and identifies canopy angular coverage as the primary limitation of the camera-based GS thermal reconstruction approach for trees.

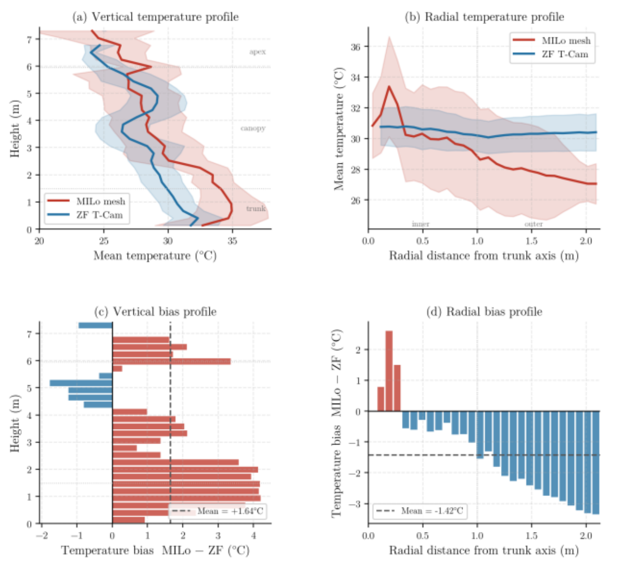
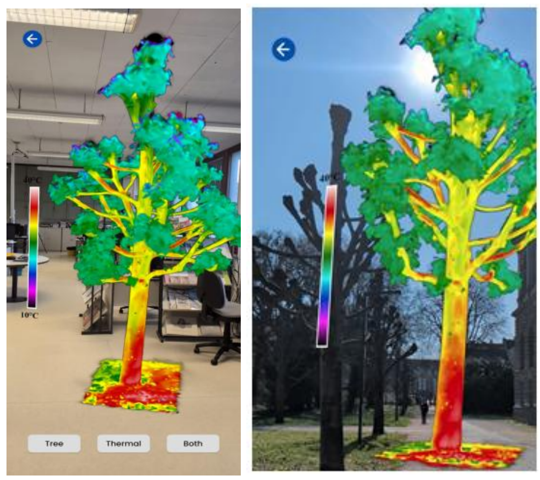


Figure 4. Spatial temperature profile analysis comparing MILo mesh (red) and Z+F T-Cam reference (blue). (a) Mean temperature vs. normalised height; (b) mean temperature vs. radial distance from the trunk axis; (c) temperature bias (MILo − ZF) as a function of height; (d) temperature bias as a function of radial distance. Shaded bands indicate ±1 standard deviation.

5.4 AR Visualization

In the marker-based mode, the thermal splat visualization was successfully anchored to the printed image target and viewed in situ on an Android device (Figure 5a). The overlay generally remained spatially aligned while the target was clearly visible to the camera, enabling close-range inspection of the thermal structure. The interface includes three display options: *“Tree”*, which presents the 3DGS model of the tree; *“Thermal”*, which displays the thermal splat visualization; and *“Both”*, which combines the tree and thermal views in a single overlay. Informal observations during deployment confirm that the visualization is easy to activate and interpret once the marker is detected, although maintaining alignment requires continued target visibility and suitable viewing conditions. These observations are reported as development-stage usability impressions rather than findings from a formal usability evaluation.

Figure 5b illustrates the GPS-based georeferencing mode, in which the thermal splat is rendered at the tree location using coordinates obtained from Google Maps, without requiring a physical marker. A visible spatial offset between the rendered splat and the physical tree can be observed. Such misalignment is expected in location-based AR, as coordinate-based placement provides only coarse spatial registration. In contrast, the marker-based configuration uses Vuforia image-target tracking to estimate pose from visual features on a known target, enabling tighter local alignment. Accordingly, the GPS-based mode should be understood as an approximate positioning strategy that helps orient users toward the correct tree at site scale and conveys a general impression of the canopy’s thermal structure.



(a) (b)

Figure 5. In-situ AR visualization of the MILo-derived thermal Gaussian splat of the tree on an Android device: (a) marker-based tracking with display mode options for *“Tree”* (RGB), *“Thermal”* (TIR), and *“Both”*; (b) GPS-based georeferencing.

6. Discussion

This study demonstrates a complete pipeline from in-situ TIR image acquisition through GS-based thermal mesh reconstruction and validation to mobile AR visualization of urban tree canopy temperature. The results confirm that MILo-based GS is both technically feasible and quantitatively meaningful for thermal tree modelling and visualization.

The collapse of traditional MVS under thermal imagery is the most consequential finding of the geometric evaluation. The root cause is structural: tree canopies in the 8–14 µm band present low chromatic and spatial contrast, undermining the feature correspondence assumptions on which photogrammetric pipelines rely. GS photometric optimization, which minimizes a dense image-level loss rather than relying on discrete keypoint matches, is intrinsically better suited to this task. This distinction is not a marginal accuracy advantage but a qualitative difference in reconstruction viability: at less than 8% of MILo's point count and an 80% outlier rate, Metashape's thermal output does not constitute a usable canopy model, regardless of its marginally lower standard deviation in the regions where reconstruction does occur. The 4.0 cm standard deviation achieved by MILo on thermal data, only 0.2 cm above its RGB result, confirms that the GS optimization framework absorbs the added difficulty of thermal imagery with a modest and practically acceptable penalty.

The high outlier rates observed across modalities with both methods (55–80%) warrant separate interpretation. The FLIR T560 and Z+F acquisitions introduce minor geometric discrepancies that the M3C2 algorithm might register as out-of-range deviations.

The colourmap inversion approach for thermal recovery is a pragmatic solution to a fundamental constraint of current GS frameworks. Recovering temperature through LUT inversion of the FLIR RHC palette is therefore not an approximation of a more principled approach but the only viable route given the current state of the framework. The 0.7°C mean absolute difference relative to the Z+F T-Cam reference, within the ±2°C accuracy of both instruments, confirms that semi-quantitative thermal analysis is achievable through this indirect route without modifying the GS optimization itself. The spatially structured nature of the remaining bias, concentrated at the trunk base and outer canopy periphery, is directly attributable to geometric acquisition differences between the close-range tripod-mounted FLIR T560 and the five-station TLS survey. This is a meaningful distinction: it implies that the colourmap inversion is not the primary source of thermal error, and that improvements in acquisition geometry (additional viewpoints, greater angular coverage of the outer canopy) would reduce distributional divergence more effectively. The most direct remediation remains the integration of temperature as a learnable per-Gaussian scalar within the MILo optimization objective, which would simultaneously eliminate palette dependency, reduce quantization error, and enable physically grounded uncertainty estimates over the reconstructed mesh.

The hybrid architecture combining marker-based and GPS-based AR addresses a practical trade-off in field deployment. Marker-based tracking through Vuforia offers the spatial stability required for close-range inspection and more structured educational use but depends on the placement and visibility of a physical target. In contrast, GPS-based placement relaxes this requirement and supports initial orientation toward the tree from a distance, although with substantially lower positional precision. These two modes therefore function as complementary strategies: GPS-based placement supports approximate site-level navigation, while marker-based tracking enables more precise local inspection once the user reaches the tree. In addition, preserving the thermal colour encoding in the AR rendering, together with the interactive temperature scale overlay, helps maintain the interpretability of the thermal data for non-specialist users. From a technical perspective, the prototype indicates the practical feasibility of deploying thermal Gaussian splats on consumer Android devices for in-situ AR visualization. However, this should be interpreted as a proof of deployment rather than a full performance evaluation, as rendering behaviour across device classes and model scales was not systematically assessed.

Three limitations bound the generalizability of these findings. First, the study is restricted to a single *Tilia tomentosa* specimen at one acquisition date; while this species may offer preferential reconstruction conditions due to its relatively thick branches and grouped foliage packets, species with higher thermal homogeneity or acquired under high-contrast conditions with strong shadow boundaries may exhibit different GS convergence behaviour and reconstruction completeness. Second, the temporal gap between the FLIR T560 and Z+F T-Cam acquisitions, though minimized by conducting both around solar noon on the same day, introduces an inevitable source of thermal discrepancy from shifting atmospheric conditions and canopy microclimate dynamics. Third, MILo reconstruction is computationally intensive and must be performed off-site, requiring either on-device GS optimization, which is prohibitive on current mobile hardware, or a cloud-based reconstruction service with low-latency mesh delivery. Closing this gap is a prerequisite for true operational deployment in urban forestry, where data acquisition and visualization must occur within a continuous field session.

7. Conclusion

This paper presents and validates a complete pipeline for 3D thermal mesh reconstruction and mobile AR visualization of urban trees from ground-based TIR imagery, establishing the technical feasibility of GS as a reconstruction framework for field-deployable canopy thermal monitoring.

Three principal contributions were demonstrated. First, the MILo GS framework was applied to TIR imagery of natural vegetation, demonstrating that GS photometric optimization can recover coherent volumetric canopy geometry from texture-poor thermal data where traditional MVS reconstruction fails. Second, a colourmap inversion methodology was developed to recover per-vertex temperature estimates from GS-derived mesh colours, confirming that semi-quantitative thermal fidelity within instrument-accuracy bounds is achievable through an indirect palette-based route without modifying the underlying GS optimization. Third, the resulting thermal Gaussian splat was deployed in a hybrid marker-based and GPS-based Android AR application, demonstrating the practical feasibility of real-time immersive visualization of canopy thermal patterns on consumer mobile hardware in field settings.

The results establish a proof of concept, laying the ground for future work. Direct integration of temperature as a learnable per-Gaussian attribute within the MILo optimization would replace the indirect colour-encoding route with a physically grounded thermal representation. On the acquisition side, expanding angular coverage of the outer canopy through additional viewpoints or UAV-based capture would reduce the geometric bias identified in the radial profile analysis. On the monitoring side, multi-temporal acquisitions across phenological stages and species would transform the current snapshot capability into a longitudinal tool for tracking canopy water stress and cooling performance across urban forest inventories. On the deployment side, cloud-based GS reconstruction with low-latency mesh delivery to the AR client represents the most tractable path toward a fully in-situ workflow.

Acknowledgements

This article is based upon work from COST Action ARiF (CA23135), supported by COST (European Cooperation in Science and Technology - www.cost.eu). We extend our thanks to the ANR TIR4sTREEt project (ANR-21-CE22-0021), through which the instruments used in this study were acquired, and the ANR-funded Digital Twins for Cultural Heritage (Twin4CH – https://twin4ch.github.io/) Junior Professorship Chair. The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

References

Bressa, N., Korsgaard, H., Tabard, A., Houben, S., Vermeulen, J., 2021. What’s the situation with situated visualization? A survey and perspectives on situatedness. IEEE Trans. Vis. Comput. Graph. 28, 107–117.

Chen, Q., Shu, S., Sun, H., Chen, J., Bai, X., 2026. Thermal3D-GS: Physics-induced 3D Gaussians for Thermal Infrared Novel-view Synthesis with a Large-Scale Dataset. IEEE Trans. Pattern Anal. Mach. Intell. https://doi.org/10.1109/TPAMI.2026.3663966. Epub ahead of print.

Garg, N., Kaur, A., Dutta, R., 2025. Marker-Based Augmented Reality in Unity: Enhancing Education Through Interactive Visualization, in: 2025 3rd International Conference on Communication, Security, and Artificial Intelligence (ICCSAI). IEEE, pp. 374–378.

Guédon, A., Gomez, D., Maruani, N., Gong, B., Drettakis, G., Ovsjanikov, M., 2025. MILo: Mesh-In-the-Loop Gaussian Splatting for Detailed and Efficient Surface Reconstruction. https://doi.org/10.48550/ARXIV.2506.24096.

Guédon, A., Lepetit, V., 2024. SuGaR: Surface-aligned gaussian splatting for efficient 3d mesh reconstruction and high-quality mesh rendering, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5354–5363.

Huang, B., Yu, Z., Chen, A., Geiger, A., Gao, S., 2024. 2D Gaussian Splatting for Geometrically Accurate Radiance Fields, in: Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers, ACM, Denver, CO, USA, pp. 1–11. https://doi.org/10.1145/3641519.3657428.

Kerbl, B., Kopanas, G., Leimkuehler, T., Drettakis, G., 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. ACM Trans. Graph. 42, 1–14. https://doi.org/10.1145/3592433.

Lecomte, V., Macher, H., Landes, T., 2022. Combination of Thermal Infrared Images and Laserscanning Data for 3d Thermal Point Cloud Generation on Buildings and Trees. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLVIII-2/W1-2022, 129–136. https://doi.org/10.5194/isprs-archives-XLVIII-2-W1-2022-129-2022.

Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R., 2022. NeRF: representing scenes as neural radiance fields for view synthesis. Commun. ACM 65, 99–106. https://doi.org/10.1145/3503250.

Pranckevičius, A. 2025. UnityGaussianSplatting (Version 1.1.1) [Computer software]. GitHub. https://github.com/aras-p/UnityGaussianSplatting/releases.

PTC Inc. 2025. Vuforia Engine for Unity (Version 11.4.4) [Computer software]. Vuforia Engine Developer https://developer.vuforia.com/references/unity/index.html

Reitmayr, G., Drummond, T.W., 2006. Going out: robust model-based tracking for outdoor augmented reality, in: 2006 IEEE/ACM International Symposium on Mixed and Augmented Reality. IEEE, pp. 109–118.

Santesteban, L.G., Di Gennaro, S.F., Herrero-Langreo, A., Miranda, C., Royo, J.B., Matese, A., 2017. High-resolution UAV-based thermal imaging to estimate the instantaneous and seasonal variability of plant water status within a vineyard. Agric. Water Manag. 183, 49–59.

Satriadi, K.A., Cunningham, A., Smith, R.T., Dwyer, T., Drogemuller, A., Thomas, B.H., 2023. ProxSituated Visualization: An Extended Model of Situated Visualization using Proxies for Physical Referents, in: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, ACM, Hamburg, Germany, pp. 1–20. https://doi.org/10.1145/3544548.3580952

Schall, G., Mendez, E., Kruijff, E., Veas, E., Junghanns, S., Reitinger, B., Schmalstieg, D., 2009. Handheld Augmented Reality for underground infrastructure visualization. Pers. Ubiquitous Comput. 13, 281–291. https://doi.org/10.1007/s00779-008-0204-5

Smigaj, M., Agarwal, A., Bartholomeus, H., Decuyper, M., Elsherif, A., De Jonge, A., Kooistra, L., 2023. Thermal Infrared Remote Sensing of Stress Responses in Forest Environments: a Review of Developments, Challenges, and Opportunities. Curr. For. Rep. 10, 56–76. https://doi.org/10.1007/s40725-023-00207-z.

White, S., Feiner, S., 2009. SiteLens: situated visualization techniques for urban site visits, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, Boston, MA, USA, pp. 1117–1120. https://doi.org/10.1145/1518701.1518871

Willett, W., Jansen, Y., Dragicevic, P., 2016. Embedded data representations. IEEE Trans. Vis. Comput. Graph. 23, 461–470.

Zürcher, R., Zhao, J., Lau Sarmiento, A., Brede, B., Klippel, A., 2023. Advancing Forest Monitoring and Assessment Through Immersive Virtual Reality. AGILE GIScience Ser. 4, 1–12. https://doi.org/10.5194/agile-giss-4-15-2023.