Multisensory Fusion: Integrating 3D Modelling with Thermal Imaging for Enhanced Scene Perception

Abstract- **This paper discussess a comprehensive approach to scene perception and modelling through the integration of cutting-edge imaging technologies. Leveraging the OpenMV-based Lepton 3.5 camera H7 plus, we employed the OpenMV IDE tool to capture thermal images in .jpg format, implementing MicroPython code for streamlined image capture. Additionally, we experimented with Python code in PyCharm, capturing thermal images. In parallel, the Lucid Vision Helios2 Time of Flight Camera facilitated the collection and conversion of point cloud data into a CAD model in .stl format using CloudCompare. Furthermore, our study delved into the realm of 3D scanning with the Intel RealSense Camera D435, utilising the LipScan3D software to capture a detailed 360-degree scan of a water bottle, subsequently saved in .stl format. A novel aspect of this work involves the superimposition of thermal imagery onto 3D model. Through Python code, we seamlessly integrated thermal information onto the 3D model, providing a comprehensive visualisation that enhances the understanding of the temperature dynamics within the scanned scene.**

Keywords- OpenMV IDE tool, 3D model, CloudCompare, LipScan3D, PyCharm, CAD, D435

I. INTRODUCTION

1.1 Motivation

The motivation behind this study lies in addressing the increasing demand for advanced imaging solutions capable of capturing multidimensional data accurately and efficiently. Traditional imaging techniques often present limitations in comprehensive data acquisition, hindering the detailed analysis required in various fields such as engineering, medicine, and environmental monitoring. By leveraging innovative imaging technologies like the OpenMV-based Lepton 3.5 Camera H7 plus, Lucid Vision Helios2 Time of Flight (ToF) Camera, and Intel RealSense Camera D435, we aim to overcome these limitations and facilitate robust data collection processes. Through the seamless integration of these technologies and the utilization of specialized software tools like the OpenMV IDE, CloudCompare, and LipScan3D, our project seeks to enable the capture of thermal images, point cloud data conversion to CAD models, and detailed 360-degree scans, all essential for comprehensive analysis and modeling. By exploring and harnessing the capabilities of these cutting-edge imaging tools, we strive to contribute to the advancement of imaging technologies and their practical applications in real-world scenarios, aligning with the evolving needs of modern research and industry standards.

1.2 Related Works

Yasuda et al. [1] consider the situation where the acquisition device is composed of at least an RGB and a thermal infrared camera. In a standard situation, the two (or more) cameras cannot share the same position. Thus, the matching of these data requires some geometric computation. A feasible way to proceed consists of adding a 3D sensor to the thermal-infrared and RGB cameras to handle these geometric constraints.

Borrmann et al. [2] use a laser scanner and an RGB camera. Due to its heavy weight, the laser scanner is attached to the RGB camera and mounted on a wheeled robot. The robot moves around the scene and performs laser scans repeatedly to generate a 3D RGB map. For each scan of the scene, the robot also acquires the corresponding 3D thermal map.

Ham et al. [3] compute the geometry of the scene using the Structure from Motion (SfM) technique [4] and superimpose the thermal information on the resulting 3D structure. This system uses only an RGB camera and a thermal-infrared camera, and thus handles both indoor scenes and outdoor scenes.

Faccio et al. [4] depth camera technology has been commercially available for more than a decade and its use in various fields, such as manufacturing, mobile robotics, object detection, and scene understanding, has become an important part of many applications. The 3-D geometry acquisition within these fields usually refers to measurement ranges of a few meters with typically required accuracies of a few millimeters up to several centimeters. Various depth camera measurement principles, such as stereo triangulation, structured light, or time-of-flight (ToF) approaches, can be used for capturing the scene [5].

1.3 Contribution of this article

The contributions of this research lie in several key areas. Firstly, it provides a comprehensive evaluation and integration of state-of-the-art imaging technologies, offering insights into their capabilities and potential applications.
Secondly, novel methodologies are developed and refined to maximize the efficiency and effectiveness of each imaging tool, addressing specific challenges in data acquisition and processing.
Thirdly, the documentation and dissemination of these methodologies serve to enhance the broader understanding and utilization of advanced imaging techniques within the scientific community. Overall, the research contributes to advancing the field of imaging technology by bridging theoretical knowledge with practical implementation, thereby fostering innovation and progress in multidimensional data acquisition and analysis.

II. THEORETICAL BACKGROUND

- **Methodology and Theory**

2.1 Data Acquisition for 360-degree 3D models

The methodology seamlessly transitions into the data acquisition phase, highlighting the utilization of Lipscan3D software in conjunction with the Intel RealSense D435 camera and rotary table. A parallel approach involves python code implementation, utilising pyrealsense2 and open3d libraries to capture pointcloud and CAD model data. Notably, the inclusion of thermal imaging is detailed, achieved through OpenMV IDE with MicroPython code for a Lepton 3.5 camera. This thermal data is saved as '.jpg,' (shown in Fig. 1) and an alternative execution using Python in PyCharm is explored, providing flexibility in the workflow.

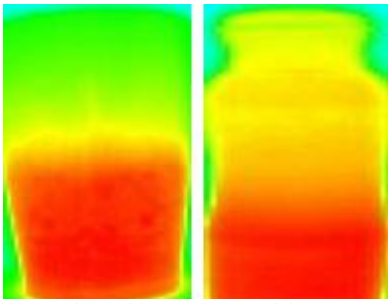

Fig. 1 : Different thermal images captured by OpenMV based Lepton 3.5 camera

2.2 Processing of 3D models

The methodology delves into the meticulous processing of 3D models, utilizing tools such as CloudCompare, Meshmixer, and FreeCAD. This stage is crucial for converting models from the .ply to .stl file format, ensuring compatibility and readiness for subsequent stages of the workflow. For thermal 3D mapping involving a range camera, respective 3D coordinates are available for each pixel of the intensity image. Accordingly, feature correspondences have to be derived either between standard intensity images (i.e. color or gray-valued images) and thermal infrared images, or between 3D point clouds and thermal infrared images.

2.3 Superimposing Thermal Images onto 3D models

A focal point of the methodology involves the superimposition of thermal images onto 3D models, accomplished through Python code leveraging Mayavi, VTK, and PIL libraries. A step-by-step breakdown is provided, encompassing the loading of the 3D model in STL format, display of thermal images, creation of a vtkImageData object, mapping thermal data onto the 3D model, and meticulous adjustments to camera position and orientation for optimal viewing. An illustrative alternative case demonstrates superimposition onto a water bottle model, elucidating the manual adjustments made to camera parameters

2.4 Integration of Hardware and Software Components

The methodology concludes by emphasizing the holistic integration of hardware and software components throughout the workflow. This inclusive approach not only captures the intricacies of 3D modelling and thermal imaging but also showcases the synergy between diverse tools, programming languages, and platforms. The comprehensive nature of the methodology ensures a versatile and effective workflow for 3D modeling and thermal imaging applications

III. PROPOSED FRAMEWORK

The method being proposed has the capability to overlay or project two-dimensional thermal images onto the corresponding three-dimensional mesh of an object. Importantly, this process does not require capturing the thermal images and the 3D mesh simultaneously. In other words, the thermal images and the 3D mesh can be obtained separately and then aligned or matched in the projection process. This approach offers flexibility in data acquisition, allowing for thermal imaging and 3D mesh creation to be conducted at different times or using different methods while still enabling their integration for analysis or visualization purposes. Additionally, we explored the simultaneous capture of thermal images and 3D mesh data, providing insights into the feasibility and effectiveness of different capture methods. Fig. 2 describes the main steps of the proposed method.

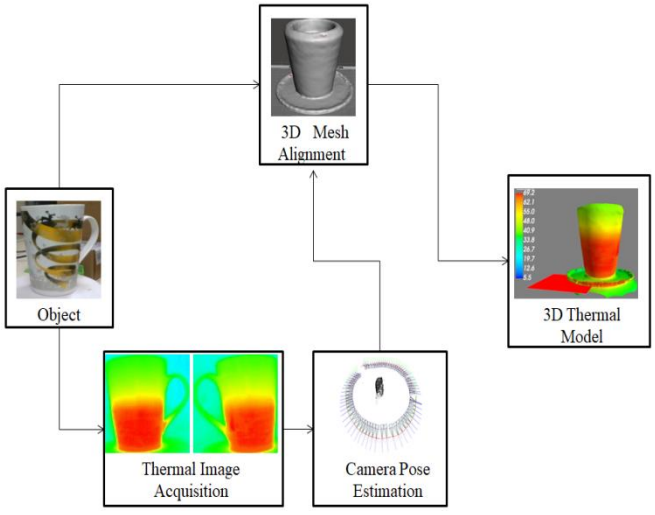

Fig. 2: Overview of Proposed Model

3.1 Imaging tools utilised

- **OpenMV based Lepton 3.5 Camera H7 plus**

Leveraging the OpenMV IDE tool, thermal images were captured in .jpg format using MicroPython code . Additionally, experimentation with Python code in PyCharm enhanced the image capture process using thermal camera as shown in Fig. 3.

Fig. 3: OpenMV based Lepton 3.5 cam H7 plus

- **Lucid Vision Helios2 Time of Flight Camera:** Lucid Vision Helios2 Time of Flight Camera shown in Fig. 4 incorporates the Sony DepthSense IMX556 CMOS sensor, featuring a resolution of 0.3 megapixels (640 x 480 pixels). Operating at a frame rate of 30 frames per second, the camera provides reliable performance across a working distance range of 0.3 meters to 8.3 meters.
  This camera facilitated the collection and conversion of pointcloud data into a CAD model in .STL format. CloudCompare played a crucial role in this process, providing a foundation for 3D modelling.

Fig. 4: Helios2 3D Camera Time of Flight

- **Intel RealSense Camera D435:** Delving into 3D scanning, the Intel RealSense Camera shown in Fig. 5, coupled with LipScan3D software, captured a detailed 360-degree scan of a water bottle stored in .stl format.

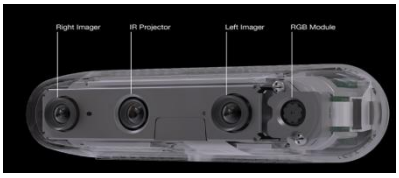

Fig. 5: Intel RealSense Camera D435

3.2 3D Modelling by Scanning

The utilization of advanced imaging tools such as the Intel RealSense camera and Lucid Vision Helios2 ToF camera. These cameras enable precise and detailed 3D scanning of objects, allowing for the acquisition of comprehensive data for further analysis and visualization. The data acquisition process

involves the integration of Lipscan3D software with the Intel RealSense D435 camera and a rotary table, ensuring meticulous capture of object geometries. Additionally, Python code implementation with pyrealsense2 and open3d libraries facilitates the capture of point cloud and CAD model data, providing flexibility and efficiency in the scanning process.

3.3 Preparation of Real Dataset

The depth camera is placed on a tripod tilted down 20 degrees. The objects are placed on a flat surface or ground plane (GP). Camera distance with respect to the objects is adjusted according with the size of the objects. As a result of the lack of a complete object shape measurement, it is required to perform a dataset augmentation by physically rotating the objects along the Z axis, Fig. 6.

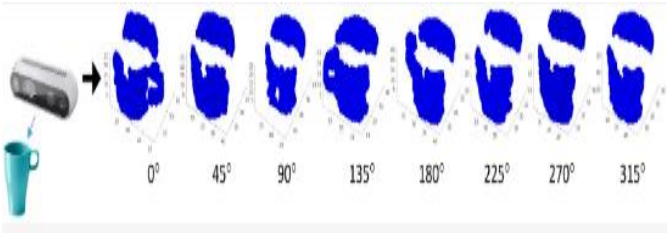

Fig. 6: Real dataset by rotation of each object along the Z axis 0°,45°, 90°, 135°, 180°, 225°, 270°, 315°

3.4 Camera Pose Estimation

Structure from Motion (SfM) is a process that simultaneously estimates the 3D structure of an object and the camera poses during image acquisition as in [6]. A transformation m(T) is proposed on the intensity levels T of the thermal image (normalized in the [0, 255] range), in order to increase the number of regions with high-contrast texture and, consequently, the number of interest points detected. The transformation m(T) is described in Equation 1:

$$m(T)=\frac{255}{2}(1-\cos(\frac{2\pi fT}{255})) \quad (1)$$

Additionally, it is possible to significantly improve the precision of pose estimation by incorporating the internal parameters of the camera – the intrinsic matrix K, described in Equation 2:

$$K=\begin{bmatrix} f_x & a_x & c_x & 0 \\ a_y & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 720 & 0 & 160 & 0 \\ 0 & 720 & 120 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

The matrix K is previously computed and is the same for all poses, since all thermal images were captured by the same camera. At last, determining the camera poses consists in the computation of the extrinsic matrix G, described by Equation 3:

$$G = R_x^{-1}.R_y^{-1}.R_z^{-1}.T_x^{-1}.T_y^{-1}.T_z^{-1} \quad (3)$$

3.5 Training

In this the distance threshold d_{th} is set to the resolution of the model point after subsampled. The model is uniformly downsampled and the density of the model point is lower than the partial view point clouds. So, the model point which is observed will have at least one point of the partial view point clouds within the distance threshold d_{th} . If the model point isn't observed, the corresponding points within the distance threshold d_{th} in partial view point cloud will not appear. At last, the observability of the model point p is calculated according to Equation 4.

$$Obs\ (p) = \frac{1}{K}\sum_{K=0}^{K-1}\delta(p, o_k) \quad (4)$$

Here, K represents the number of viewpoints. o_k represents the direction of the viewpoint vector, and the function δ returns "1" when the model point p is observed.

Using the same approach as calculating the observability of model points, we can calculate the observability of point pair feature $obs(F(m_r, m_i))$. Because m_r and m_i are observed simultaneously, $obs(F(m_r, m_i))$ is calculated by Equation 5.

$$Obs(F(m_r, m_i))=\frac{1}{K}\sum_{K=0}^{K-1}\delta(m_r, o_K)\ \ \delta(m_i, o_K) \quad (5)$$

Based on the observability of the point pair feature $Obs(F(m_r, m_i))$, we can set an observability threshold τ_{obs} . When $Obs(F(m_r, m_i))> \tau_{obs}$, the point pair feature $F(m_r, m_i)$ is computed according to equation 1. When $Obs(F(m_r, m_i)) \leq \tau_{obs}$, the point pair feature $F(m_r, m_i)$ is set to NaN. Then, the point pair features with observability larger than the given threshold are grouped together based on their similarity and then stored in a hash table.

3.6 Thermography

The detailed inclusion of thermal imaging using the OpenMV IDE with MicroPython code for a Lepton 3.5 camera. Thermal imaging adds an additional dimension to the data acquisition process, capturing infrared radiation emitted by objects with temperatures above absolute zero. This enables the detection

of temperature variations across surfaces, contributing to a more comprehensive understanding of object characteristics.

Tsai and Huang introduced a multiplicative model by coming with shadows, and the segmented results do provide the surface temperature differences though shadows do exist in the thermal infrared images [7]. Hence, an image model introduced shadows was given as follows equation (7):

$$I=BT+n \tag{6}$$

Where, I is the given thermal image, B is the environmental effects, T is the calibrated surface temperatures, and n is the noise.

Vese and Chan introduced the regional constants (average values of the segmented regions) into the image segmentation model such that the image segmentation can be modelled as follows equation (8):

$$E(c_i, C) = \sum_{i=1}^{n} \int_0^{\Omega_i} (I - c_i)^2 dxdy + v|C| \tag{7}$$

where Ω_i is the i^{th} segmented regions, c_i is the corresponding regional constants, v is the positive constant, and $|C|$ indicates the length of the segmented regional boundaries.

The image segmentation model proposed by Li is given as follows eq.[9]

$$EE(\emptyset, b, c)=\int (\sum_{i=1}^{N} \int K(y-x)|I(x) - b(y)c_i{}^2 M_i(\emptyset)dxdy) + v\int |\nabla H(\emptyset)|dx + \mu \int p(|\nabla \emptyset|dx \tag{8}$$

where, ∇ is the gradient operator, $\emptyset$ is the set of the level set functions $\emptyset_1$ and $\emptyset_2$. $M_i(\emptyset)$ is defined as the Heaviside function of the level sets, and is represented as follows.

$$M_1(\phi) = H(\phi_1)\, H(\phi_2)$$

$$M_2(\phi) = H(\phi_1)(1 - H(\phi_2))$$

$$M_3(\phi) = (1- H(\phi_1))\, H(\phi_2)$$

$$M_4(\phi) = (1- H(\phi_1))\, (1 - H(\phi_2)) \tag{9}$$

The regional constants c_i can be obtained by taking the derivative of eq. (8) with respect to c_i, and regional constants can be given as follows

$$c_i=\frac{\int B\otimes K)\, I_0 M_i(\phi(y))dy}{\int (B^2\otimes K) M_i(\phi(y))dy},\ i=1,..,\ N \tag{10}$$

where, $\otimes$ is the convolution operation. The shadow influence B can be derived in the similar way, and can be given as follows

$$S=c_1M_1(\phi) + c_2M_2(\phi)+ c_3M_3(\phi)+ c_4M_4(\phi) \tag{11}$$

As for the level set functions ϕ_1 and ϕ_2, the finite difference approach is introduced in a iteration scheme such that the level set functions ϕ_1 and ϕ_2 can be represented respectively. The image segmentation is simultaneously applied on the thermography collected by thermal infrared camera. The image segmentation employed in this approach can correctly segment the given thermal infrared images with environmental effects. In doing so, the surface temperature differences can be identified from the segmented results and the surface temperature boundaries can be located.

3.7 Superimposing Thermal images onto 3D models

The methodology involves the integration of Python code leveraging Mayavi, VTK, and PIL libraries. This allows for the precise alignment and overlay of thermal images onto corresponding 3D models, enhancing the visualization and analysis capabilities of the data. The step-by-step breakdown provided ensures accurate mapping of thermal data onto the 3D model, with meticulous adjustments made to camera parameters for optimal viewing. An illustrative case study involving the superimposition onto a cup model shown in Fig. 7 demonstrates the effectiveness and applicability of the methodology in real-world scenarios.

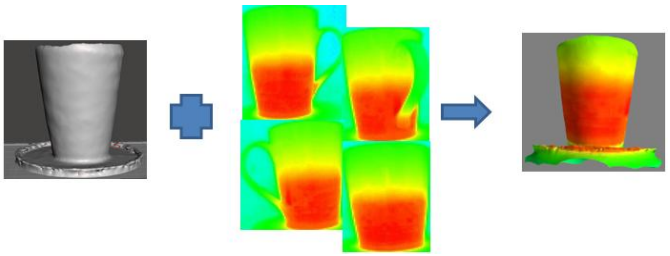

Fig. 7: The 3D model from a 3D scanner (left) combined with the 2D thermal images obtained using thermal camera (middle) to produce the final 3D thermal model(right)

IV. EXPERIMENTATION AND DISCUSSION

The comprehensive analysis of imaging tools highlights the capabilities of the Lucid Vision Helios2 Time of Flight (ToF) camera, which facilitated the collection and processing of point cloud data. This data was then transformed into a detailed CAD model in .STL format, with the intricate details vividly illustrated in the left image (Fig. 8) through

CloudCompare. Furthermore, thermal imaging was seamlessly integrated into our 3D modeling endeavors by leveraging the OpenMV-based Lepton 3.5 Camera H7 plus and the Lucid Vision Helios2 ToF camera. Through Python code implementation, thermal images were superimposed onto the 3D model, as depicted in the right image (Fig. 8), showcasing the fusion of thermal data onto the detailed 3D structure. These results underscore the synergy between different imaging technologies, demonstrating their combined potential in creating a holistic understanding of real-world environments. Moreover, the successful superimposition of thermal images onto the 3D model opens up new avenues for multidimensional imaging applications, providing valuable insights into scene perception and modeling.

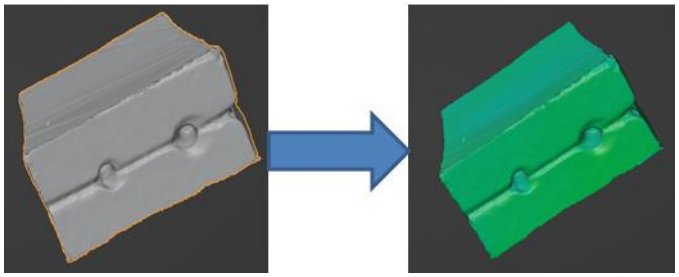

Fig. 8: The left image reveals the CAD model, while the right image displays the superimposed thermal image onto 3D model

In this study, we utilized advanced imaging tools to perform a comprehensive 360-degree 3D scan of a water bottle. Employing the Intel RealSense camera with LipScan 3D software, we captured the bottle's shape in detail, resulting in a 3D model (Fig. 9, left). By integrating thermal imaging into our methodology using Python code and the Lepton 3.5 Camera H7 plus and Helios2 ToF camera, we successfully overlaid thermal images onto the 3D model (Fig. 9, right). This integration not only enhances our understanding of the bottle's shape but also provides insights into its surface temperature variations.

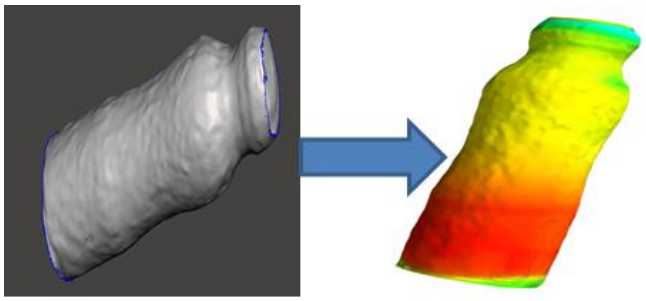

Fig. 9: The left image shows the CAD model, while right image displays the superimposed thermal image onto 360 degree 3D model

We initiated our study by acquiring and integrating the Boest 3D Scanner software using Python code. Through this implementation, we successfully generated a .STL file, as seen in the left image (shown in Fig. 10). This file, obtained through the Boest 3D Scanner software, serves as a fundamental component for our subsequent analyses. Furthermore, we explored the synergy between thermal imaging and 3D modeling by superimposing a thermal image onto the 3D model, as illustrated in the right image(shown in Fig. 10). This integration provides a comprehensive view, combining both the structural information from the 3D model and the thermal characteristics captured by our thermal camera.

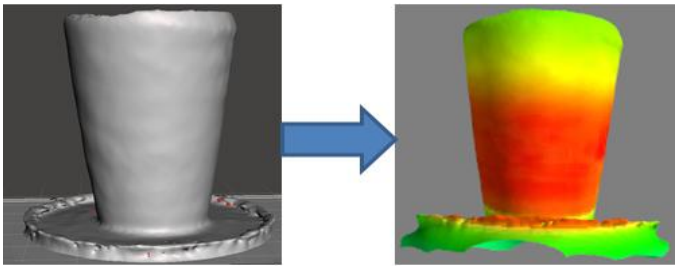

Fig. 10: The left image shows the CAD model, while right image shows the superimposed thermal image onto 3D model

In this phase of study, we introduced hot water into a full cup. We captured a detailed 360-degree 3D scan of cup and saved the resulting models in .STL format, displayed in the left image (shown in Fig. 11). This process not only allows for a thorough examination of the cups physical features but also provides a basis for understanding the impact of the introduced hot water. The right image (Fig. 11) showcase the successful superimposition of thermal images onto the respective 3D models, offering insights into the thermal behavior of the cups when filled with hot water.

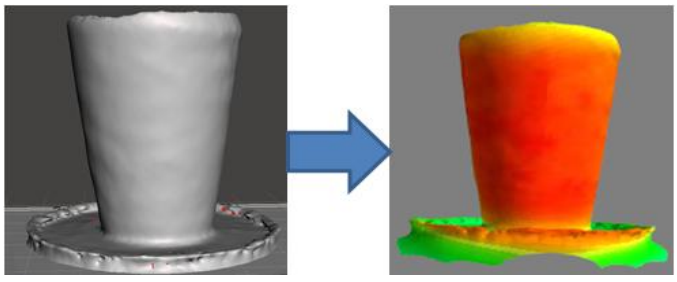

Fig. 11: The left image shows the 360 3D scan of Cup, while right image shows the superimposed thermal image onto 3D model

As part of analysis, utilized Python code to acquire a temperature gradient scale and object temperatures through our thermal camera, as shown in Fig. 12. This step improves our capability to measure and visualize temperature variations

across various surfaces. Integrating temperature information into our study contributes to a nuanced comprehension of the thermal characteristics of scanned objects, offering valuable insights for applications like thermal analysis and quality control.

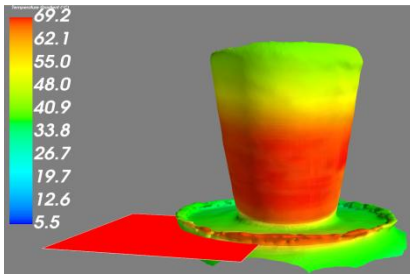

Fig. 12: temperature gradient scale and superimposed thermal image onto 3D model

V. CONCLUSION

This study showcased the synergistic capabilities of advanced imaging tools, particularly the Intel RealSense Camera. The seamless integration of point cloud data and thermal imaging, demonstrated through the detailed 3D scans of a water bottle and a cup filled with hot water, reflects the power of combining diverse imaging technologies. The superimposition of thermal images onto 3D models, illustrated in Figures 8, 9 ,10, and 11 not only provides a comprehensive view of structural details but also offers valuable insights into temperature variations.

The incorporation of the Boest 3D Scanner software and the utilization of Python code for temperature analysis further enriched our methodology. The generated .STL files serve as fundamental components for subsequent analyses, highlighting the adaptability of our approach. Overall, our findings underscore the potential of multidimensional imaging applications, opening new avenues for scene perception, modeling, and thermal analysis in real-world environments.

REFERENCES

[1] K. Yasuda, T. Naemura, and H. Harashima, “Thermo-key: Human region segmentation from video,” IEEE Computer Graphics and Applications, Vol.24,no.1, pp.26-30,2004

[2] D.Borrmann, J. Elseberg, and A. Nuchter, “Thermal 3D mapping of building facades,” *Intelligent Autonomous Systems 12*, vol.193, pp.173-182, 2013.

[3] Y. Ham and M. Golparvar-Fard, “An automated vision-based method for rapid 3D energy performance modeling of existing buildings using thermal and digital imagery,” *Advanced Engineering Informatics*, vol.27, no.3, pp.395-409, 2013.

[4] M**.** Faccio, E. Ferrari, M. Gamberi and F. Pilati, “Human factor analyser for work measurement of manual manufacturing and assembly processes", *Int. J. Adv. Manuf. Technol.*, vol. 103, no. 1, pp. 861-877, Jul. 2019.

[5] B. Langmann, K. Hartmann and O. Loffeld, "Depth camera technology comparison and performance evaluation", *Proc. ICPRAM*, vol. 2, pp. 438-444, 2012.

[6**]** Youngjib Ham and Mani Golparvar-Fard, “ An Automated Vision-based Method for Rapid 3D Energy Performance Modeling of Existing Buildings Using Thermal and Digital Imagery”, *Advanced Engineering Informatics*, 27(3):395 – 409, 2013**.**

[7] P.H. Tsai, Y. Huang and H.S. Tai, “The Feasibility of Identifying Defects Illustrated on Building Facades by Applying Thermal Infrared Images with Shadow”, Proceedings 2019, 27, 15th International Workshop on Advanced Infrared Technology and Applications, Florence, Italy.

[8] L.A. Vese and T. F. Chan, “A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model*”, International Journal of Computer Vision,* Vol.50, No.3, pp. 271-293.

[9] C. Li, R. Huang, Z. Ding, J. C. Gatenby, D. N. Metaxas, and J. C. Gore, “A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI,” *IEEE Trans. Image Process*., vol. 20, no. 7, pp. 2007–2016, Jul. 2011.