3D Thermal Mapping of Building Interiors using an RGB-D and Thermal Camera

Stephen Vidas[1,2], Peyman Moghadam[2] and Michael Bosse[2]

***Abstract*— The building sector is the dominant consumer of energy and therefore a major contributor to anthropomorphic climate change. The rapid generation of photorealistic, 3D environment models with incorporated surface temperature data has the potential to improve thermographic monitoring of building energy efficiency. In pursuit of this goal, we propose a system which combines a range sensor with a thermal-infrared camera. Our proposed system can generate dense 3D models of environments with both appearance and temperature information, and is the first such system to be developed using a low-cost RGB-D camera. The proposed pipeline processes depth maps successively, forming an ongoing pose estimate of the depth camera and optimizing a voxel occupancy map. Voxels are assigned 4 channels representing estimates of their true RGB and thermal-infrared intensity values. Poses corresponding to each RGB and thermal-infrared image are estimated through a combination of timestamp-based interpolation and a pre-determined knowledge of the extrinsic calibration of the system. Raycasting is then used to color the voxels to represent both visual appearance using RGB, and an estimate of the surface temperature. The output of the system is a dense 3D model which can simultaneously represent both RGB and thermal-infrared data using one of two alternative representation schemes. Experimental results demonstrate that the system is capable of accurately mapping difficult environments, even in complete darkness.**

I. Introduction

The building sector has been shown to dominate energy consumption in developed nations such as the U.S. [1]. Improving energy efficiency in buildings is critical to achieving a sustainable future in the face of anthropomorphic climate change. Conventional thermography for building energy auditing and non-destructive assessments relies on 2D thermal images, which have significant limitations. For example, a lack information on the geometry and location of objects and areas of interest in the scene. As a result, there is growing interest in moving from 2D towards photo-realistic 3D representations of environments with integrated temperature data [2]–[4]. Accurate geometric measurements are able to be made using these 3D representations of heat distribution. The combined knowledge of temperature and geometry then enables quantitative estimates of surface energy loss through heat radiation. Furthermore, the precise detection and location of heat sources, thermal bridges and other phenomena becomes possible [5].

[1]S. Vidas is with the Science and Engineering Faculty, Queensland University of Technology, Brisbane, Australia stephen.vidas@qut.edu.au

[2]All authors are with the Autonomous Systems Laboratory, CSIRO ICT Centre, Brisbane, Australia {steve.vidas, peyman.moghadam, mike.bosse}@csiro.au

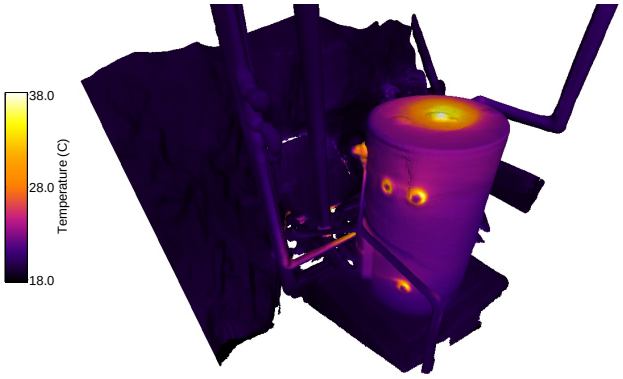


Fig. 1: Result of a 3D thermal mapping of an underground hot water system. Data was collected in complete darkness which is often needed to reduce environmental impacts on temperature readings, such as from sunlight.

This paper proposes a mobile 3D thermal mapping system which incorporates color, thermal-infrared and depth information simultaneously. The system is targeted towards the application of continuous and non-destructive monitoring of building interiors for energy efficiency assessment. The proposed approach uses only an RGB-D camera and a single additional thermal-infrared camera, and is capable of forming dense and high-fidelity 3D models for the purpose of further analysis, as demonstrated in Figure 1. While previous studies have explored the area of 3D thermal mapping, each of these were found to have limitations. These included the inability to operate at night (often important for energy auditing), or in confined spaces with obstacles such as stairs.

The method proposed in this paper is the first that is hand-held, uses readily available hardware, and is capable of operating at night. The mobility associated with being hand-held allows seamless mapping between floors and interiors of a building, however, there is no technical reason that the device could not be mounted on a robot or wheeled platform. The capacity to operate in darkness is also important, because often explicit night-time analysis is required by applications such as building energy auditing. This is due to the need to limit the effects of environmental phenomena such as sunlight on surface temperatures. Data can be captured simply by moving around an environment with the proposed system and streaming to a computer in a wearable unit such as a backpack. Results take the form of a colormapped 3D model which transforms the 4 dimensions of data (3 color channels + thermal-IR) into a single multi-modal representation. This allows human viewers to easily and simultaneously interpret

the available RGB and thermal data under diverse lighting conditions.

Our major contributions are:

- Geometric and temporal calibration of the dual camera (thermal and RGB-D) setup
- Accurate ray-casting through an interpolated pose-estimation scheme
- Simultaneous representation of multiple modalities (RGB and thermal-IR) through a dense 3D model

Our current hand-held 3D thermal mapping prototype (shown in Figure 2) comprises of a low-cost RGB-D camera (Microsoft Kinect) rigidly attached to a thermal-infrared camera (Thermoteknix Miricle 307K). Like any other multi-modal sensor configuration, the first key step is to geometrically and temporally calibrate the sensors with respect to each other, a process which is discussed in detail in Section III. Our past work formed the basis for the problems of both the intrinsic calibraton of the thermal-infrared camera [6], and the extrinsic calibration of the multiple-modality, multi-sensor configuration [7]. We were also required to analyze timing irregularities associated with both the cameras, and a novel technique was developed to smooth capture times and temporally calibrate the cameras.

The KinectFusion algorithm [8] was used to obtain an estimate of the pose of the RGB-D camera for each captured frame, and to optimize a voxel occupancy map. In Section IV we outline how we extended this to obtain estimates of the pose of the thermal camera. This required a good understanding of the temporal differences between cameras and an effective temporal calibration. A combination of cubic spline interpolation and a knowledge of the extrinsic calibration between cameras was used. This enabled accurate raycasting using all of the available image data.

In Section V we describe our raycasting technique to assign values from estimated camera poses. We also show how to get more accurate temperature estimates by considering the mechanical properties of the thermal-infrared sensor.

Finally, two novel methods for the simultaneous representation of the thermal and visible-spectrum data on the dense 3D model are demonstrated. To the best of our knowledge, this is the first time that the visible and thermal-infrared modalities have been fused for the purpose of 3D mapping. Our proposed representations to allow a human viewer to visualize information from both modalities on the same model, are introduced in Section VI.

Our evaluation in Section VII demonstrates that the system is highly capable of generating precise models with incorporated temperature data, even in difficult circumstances such as complete darkness or cluttered environments.

The data captured for this research will be made publicly available and can be found along with demonstration videos online.[3]

II. State of the art

Several previous studies have explored the potential of 3D thermal mapping for applications like building inspection and auditing [2]–[4]. These studies were motivated by many of the limitations of performing such tasks with conventional thermographic images. A key limitation is the lack of a holistic and geometrically accurate model of the environment which could be used for automatic offline analysis. Also, an immersive 3D environment provides an opportunity for a user to investigate and detect anomalies in context, rather than simply within a sequence of 2D images.

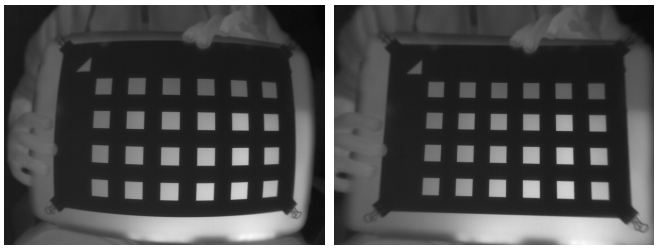
Fig. 2: The sensor configuration prototype: comprised of a Microsoft Kinect and Thermoteknix Miricle 307K thermal camera.

The ThermalMapper project [2] involves a mobile, wheeled robot with a terrestrial laser scanner and thermal-infrared camera, and is capable of projecting thermal data onto a 3D model as it explores an environment. The system implements a “stop and go” approach for performing scans, as opposed to a continuous scanning approach as proposed in our paper. Results from the automated ThermalMapper system are in the form of dense 3D point clouds that can be visualized in either thermal-infrared or RGB. Because the system is mounted on a wheeled robot, it is not capable of traversing up stairs or exploring difficult terrain or confined spaces. There is also a significant cost difference with our approach, as the 3D LiDAR and robotic platform is considerably more expensive than a single RGB-D camera.

An approach using only a single RGB camera together with the thermal-infrared camera was explored by [3]. Their technique uses a Structure from Motion (SfM) and Multi-View Stereo (MVS) pipeline to solve for the geometry of the cameras and form a dense, colored point cloud with thermal data optionally overlayed. However, it is vulnerable to many of the typical failure conditions of SfM to which our approach is immune, such as lack of visual texture or low levels of lighting. This prevents the effectiveness of the approach for many inspection tasks, which are often best conducted in darkness to explore differences in energy usage between night and day. In addition, it is not capable of generating dense 3D models which can enable detailed quantitative analysis. Single-camera SfM systems also have no global scale, so the true dimensions of the model cannot be determined without a reference of known size.

The use of a combined LiDAR and thermal imaging

[3] http://www.youtube.com/user/AutonomousSystemsLab



(a) Original image (b) Undistorted image

Fig. 3: Comparison of a thermal-infrared image before and after the correction of lens distortion using the planar mask.

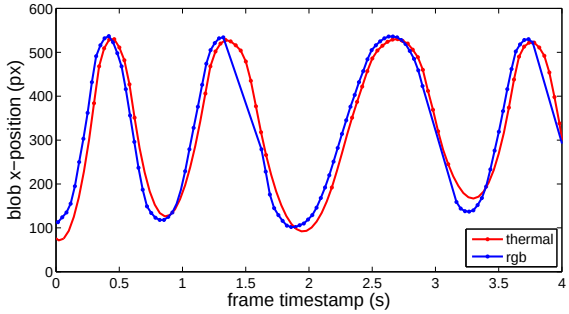


Fig. 4: Interpolated x-position of blob centroids for each camera for a small section of the 5 minute calibration sequence. A slight but consistent lag in the thermal camera timestamps can be observed.

system for 3D thermal mapping has also been proposed [4]. Thermal stereo solutions have been explored [9], but have so far been limited to small-scale reconstructions and suffer from the additional cost of a second thermal camera rather than a much cheaper RGB-D camera. Other approaches for performing 3D thermal mapping have largely been limited to combing thermal imagery with existing 3D models [10], [11].

III. Calibration

A. Intrinsic Calibration

Raw images captured from the thermal-infrared camera have a significant amount of lens distortion. A common method for geometric camera calibration uses a checkerboard to estimate the intrinsic parameters of the camera. However, checkerboard patterns are not clearly visible in the thermal-infrared modality, and thus the approach requires modification. Our previous work adapted Zhang's calibration method [12] for the thermal-infrared modality [6].[4] Our method involves using a 2-dimensional thermal mask held in front of a backdrop with a significantly different level of thermal radiation. An MSER-based [13] detection algorithm is then used to find the mask in the image, and subsequently locate the corners to subpixel accuracy and optimize a distortion model. It should be noted that spatially remapping the image to correct for lens distortion has the effect of reducing the amount of noise, since non-uniformities (a unique problem of thermal-infrared sensors) are more intense at greater distances from the center of the image. Figure 3 demonstrates the effectiveness of this approach for correcting the thermal-infrared images.

B. Temporal Calibration

The Kinect does not support hardware synchronization, so there is no guarantee that pairs of depth and RGB frames are taken at identical times. Experiments demonstrated that the RGB and depth streams are captured at slightly different frame rates, resulting in a continuous drift between the two streams. The delay between two RGB and depth frames can differ by up to 17ms, at an approximate framerate of 30Hz.

The Kinect has internal clocks which run at 60MHz and generate frame timestamps. We use these timestamps to synchronize the RGB and Depth frames. The difference between the RGB internal clock C_c and the depth internal clock C_d frames can be used to recover the offset of clock $C_d(t)$ relative to clock $C_c(t)$ at each time stamp t.

In order to temporally calibrate the thermal camera and Kinect (i.e. to determine the offset between the sensor capture times), we designed an experiment by capturing a series of random cyclic motions while the platform is pointing at a hot, bright object in front of a relatively uniform background. In our case we used a computer monitor that had been powered for at least 10 minutes (in order to warm up), displaying a full-screen red image. Then, we use blob detection to detect and track the region of interest in both the RGB and thermal images. The x-position of the centroid of the blob in each modality generates two signals. Since the two cameras capture images at different frequencies, and the blob may not have been tracked successfully in all frames, the signals are resampled to the same rate using cubic spline interpolation. The two reprocessed and interpolated signals are shown in Figure 4.

The shift between the two signals can then be easily found using cross-correlation. However, such temporal alignment is bounded by the sampling frequency of the two signals. To get higher accuracy, we fit a parabola to the cross-correlation result and the maximum is used to further optimize the estimate of the latency between the thermal-infrared camera and the Kinect.

The temporal calibration method is based on the assumption that the thermal image timestamps are noise free. However, under realistic conditions the delay between image capture and arrival at the host computer is unpredictable due to non-deterministic buffering which may occur on the sensor, the bus or the host interface. As a result, the timestamps assigned to the image at the host computer are not linearly related to the true capture times. We use a modified implementation of the convex hull algorithm [14] to correct the timestamps.

First, we find the lower boundary of the convex hull of the measurement timestamps. Then, the algorithm fits a straight line to each consecutive point of the convex hull

[4] http://code.google.com/p/thermalvis-ros-pkg/

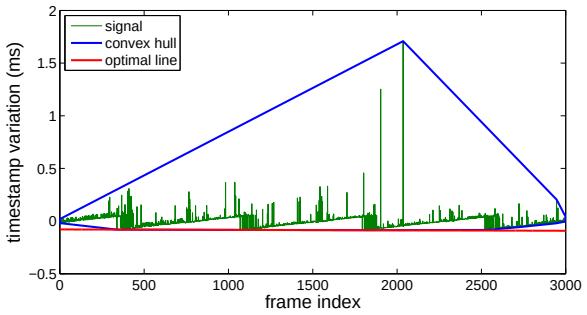


Fig. 5: The convex hull fitted to the thermal camera timestamp signal, with the linear trend subtracted for visualization purposes. The blue line represents the convex hull of the signal minus the linear trend (in green), and the red line represents the smoothed signal (again, minus the linear trend).

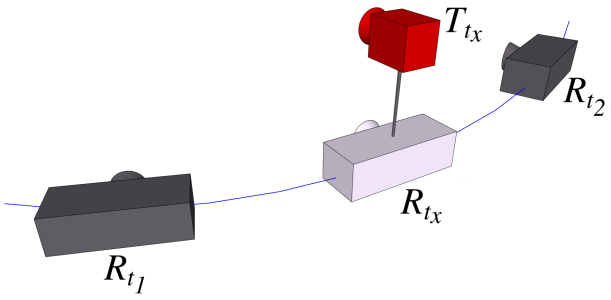


Fig. 6: The relationship of the poses of the two cameras at unsynchronized time instances.

and calculates the residual between the hypothesis line and the points on the lower boundary of the convex hull. The straight line with the minimum residual error will be the optimal straight line representing the smoothed timestamp signal. This is illustrated in Figure 5.

C. Extrinsic Calibration

In order to be able to integrate information from alternative sensing modalities, it is required to represent them in a common geometric reference frame. The process of determining the relative translation and rotation between different sensors' coordinate systems is called extrinsic calibration. Conventional extrinsic calibration techniques are based on observing a planar checkerboard pattern from all sensing modalities [15], [16]. However, accurately detecting a checkerboard pattern accurately with a low-resolution thermal-infrared imaging system is difficult. In addition, the procedure of extrinsic calibration using checkerboards is a labour intensive process, requiring the placement of an artificial target in the scene at several poses with respect to the sensors.

We present a new technique for estimating the extrinsic parameters of a thermal camera rigidly attached to the Kinect sensor with no requirement of artificial targets [7]. The only assumptions are that both the Kinect and the thermal-infrared cameras share similiar fields of view, and that there are non-colinear straight lines detectable in both the thermal image and the range data. This is so that the 3D linear features extracted from the range data are associated with the 2D linear features in the thermal image, and constrain the geometric alginment. The backgrounds of the images are unimportant.

For our experiment, we used a powered computer monitor which produced distinctive linear features in both modalities because of its shape and temperature. First, a set of natural 2D and 3D lines (plane intersections and boundary lines) are extracted from the thermal-infrared image and Kinect range data respectively. Next, given a set of matched 2D-3D lines, we use an iterative non-linear optimization solution to minimize the distance between the 2D-3D line pairs. A minimum number of three distinct lines are required for the optimization to converge. Finally, the estimated six degrees-of-freedom (6DoF) transformation is used to register the thermal-infrared camera coordinate system with respect to the Kinect coordinate system.

IV. Geometry Estimation

The KinectFusion algorithm [8] can retrieve poses for each RGB-D frame. We use these poses as a basis for estimating the 3D pose of the thermal-infrared camera for each thermal frame. The following process is used which is also illustrated in Figure 6:

1) Four neighbouring RGB-D frames are selected: the two immediately preceding the thermal frame and the two immediately following (times t_0, t_1, t_2 and t_3).
2) A cubic spline is fitted to the four poses corresponding to these RGB-D frames, representing the 3D path of the camera throughout this interval.
3) The pose of the range sensor (R) for the exact time of the thermal frame (t_x) is estimated using the spline.
4) The pose of the thermal camera (T) for the time instance t_x is calculated using the 6DoF transform relating the two cameras (determined through extrinsic calibration).

V. Temperature Assignment

The output of the KinectFusion algorithm includes an uncolored 3D voxel occupancy map, which is converted to a dense 3D model using the marching cubes algorithm, and then simplified using a quadric-based edge collapse strategy. Raycasting is performed for each valid frame from both the thermal and RGB camera, in order to assign RGB and thermal values to each vertex in the model.

Assignment of thermal values requires greater care, because of the nature of the sensor. Many thermal-infrared cameras require regular intensity calibrations to correct for a gradual decrease in measured signal accuracy during operation. These operations are known as NUCs (Non-Uniformity Corrections) and involve a mechanical shutter operation which temporarily blocks the imaging sensor with a sheet of material of uniform temperature. This causes an interruption to data streaming of up to one second, but is necessary to

ensure that the pixel values continue to represent an accurate estimate of thermal radiance (and therefore temperature) of the scene. NUCs are generally triggered by either an increase in equivalent thermal noise, or the passage of a specified period of time - whichever comes first. The more frequently NUCs can be performed, the more accurate the thermal-infrared image will be. An additional factor that can affect the accuracy of the recorded thermal value is the incident angle between the sensor and the radiation emitted from the surface. It is best practice to minimize this angle if at all possible.

Therefore several strategies are employed in order to more accurately assign thermal values to the model:

- The frequency of NUCs is increased to occur approximately each minute.
- Frames retrieved from the camera during a NUC are ignored.
- Only rays which have an incident angle of less than 30 degrees are considered for temperature assignment.

VI. Multi-modal Representation

Thermal-infrared image data is typically visualized as a colormapped 2D image with a colorbar to help the viewer associate pixel RGB values with approximate temperature estimates. Using a colormap to represent the range of intensities allows a human viewer to discern more discrete intensity levels than if the image were rendered in grayscale [17].

While these schemes can be extended to the 3D voxel case, a problem remains when attempting to visualize RGB and thermal-IR data simultaneously. This is because there are now 4 dimensions (red, green, blue and thermal-IR) but only 3 channels available (red, green and blue) for visualization. The significant difference in wavelengths between visible light and thermal-infared radiation results in a large amount of complementary information between modalities that would be useful if able to be viewed simultaneously. For example, in an electrical maintainence context, visualizing a model in the thermal-infrared modality would enable the viewer to determine which of many electrical connections were heating up abnormally due to faulty connections. However, without seeing the visible data, there may be no way to retrieve information associated with a text or color-based label, which may be critical for energy auditing applications where particular appliances or connections must be easily identified by a user.

In this paper, two alternative 4D-3D representation schemes are proposed that attempt to incorporate both visible-spectrum and thermal-infrared data in the same model. These are Intensity-Hue (I-H) Mapping and Thermal Highlighting. Each of these schemes is designed to have specific advantages for specific circumstances. Both schemes are used in conjunction with a colormap which is used to assign RGB values for each level of thermal-infrared intensity.

I-H Mapping first converts both the colormapped thermal value and the visible-spectrum RGB value into HSL space. For the fused value, the hue is from the colormapped thermal representation, while the luminance is from the visible spectrum. In the experiments, luminance is remapped from $(0,1)$ to $(0.2,0.8)$ so that very dark or light regions do not obscure the hue representing the temperature. Saturation is set to its maximum value for all voxels. The effect of this combination is that the relative temperatures can be interpreted almost as easily as with a thermal-only colorization, but the texture and lighting conditions from the visibile spectrum can also be seen. This is useful for reading text and identifying patterns that are not able to be discerned from thermal-infrared alone.

Thermal Highlighting uses a weighting scheme to vary the bias of fused voxels towards either the thermal or visible modalities. For average thermal-infrared intensities, representing the average surface temperature of the scene, the RGB value from the visible-spectrum is used. As the thermal-infrared intensity deviates increasingly from the mean, the bias is shifted towards the colormapped thermal value. This serves the purpose of highlighting the "hottest" and "coldest" surfaces with strong colors that make them highly salient to a human viewer, whilst still retaining visible-spectrum color and texture data for the majority of the scene. This color scheme provides a solution for when color information is useful, but the detection of extreme temperatures in the environment is most critical. The implementation of the weighting scheme as applied to each voxel is shown in Equation 1:

$$f_{rgb} = \quad v_{rgb} - w(t_{raw}, T_{raw}) \left(t_{rgb} - v_{rgb}\right) \tag{1}$$

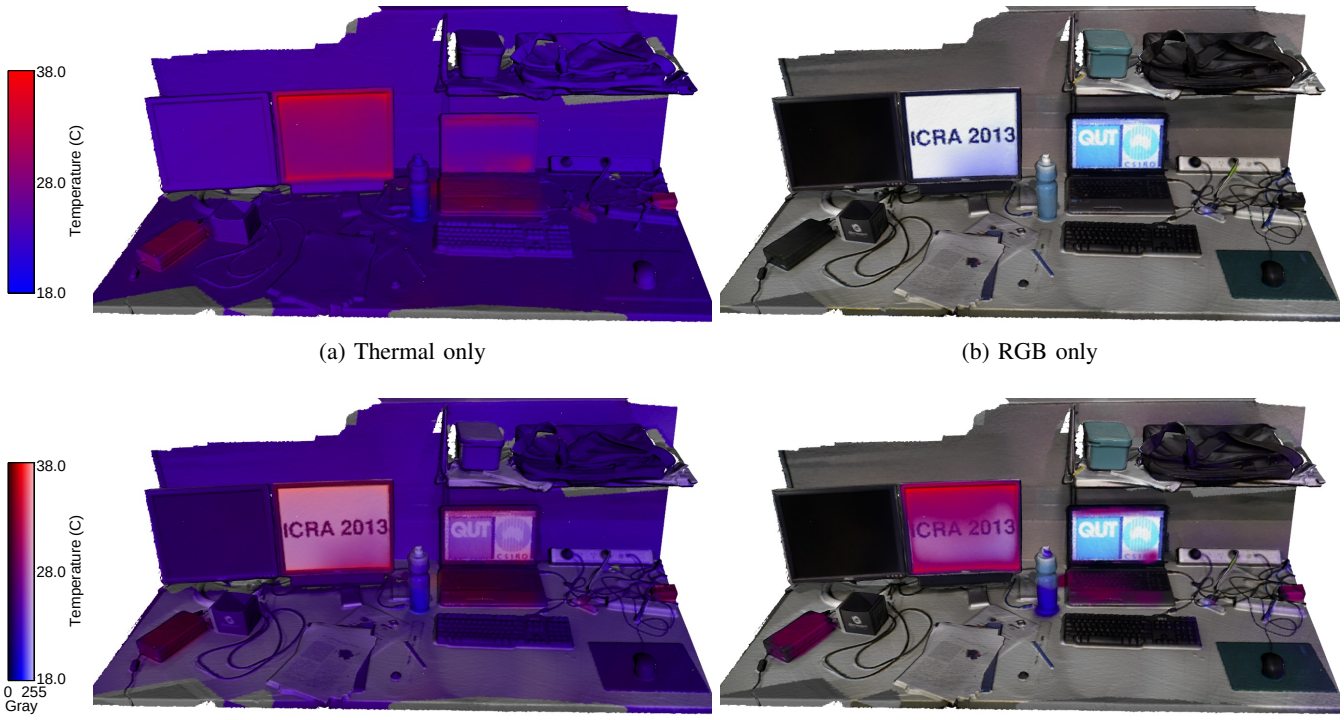
where, t_{raw} is the original monochromatic thermal value, and T_{raw} is the distribution of all values for all voxels. The weighting function w() is then used to bias the final color f_{rgb} towards either the RGB value v_{rgb} or the colormapped thermal value t_{rgb}. For our experiments, the weighting function smoothly transitions from 0 to 1 as a thermal value departs from the mean thermal value to the minimum or maximum.

Figure 7 shows these two alternative schemes in comparison with a thermal-infrared only model and an RGB only model. The effectiveness of these schemes is discussed further in the next section.

VII. Evaluation

From Figure 7 it is seen that the thermal-only visualization (Figure 7a) clearly shows temperature variation over the scene, including hot spots such as computers, monitors and power supplies. However, the RGB only visualization (Figure 7b) reveals many details that are hidden to thermal-infrared, such as text displayed on the computer monitors, a logo on the small black box to the left, and the details of this very paper which can be seen on the table. In practical inspection applications, these potentially identifying details (as well as colors) may be needed in the visual representation simultaneously along with the temperature information.

The proposed I-H Fusion scheme in Figure 7c can be seen to effectively preserve the representation of the surface temperatures using the same false colormapping as the thermal only scheme, whilst simultaneously allowing written and textural information to be perceived clearly. If color



(a) Thermal only

(b) RGB only

(c) I-H Mapping

(d) Thermal Highlighting

Fig. 7: Demonstration of alternative fusion schemes for texturing a 3D model with RGB and thermal data simultaneously.

information must be preserved, the Thermal Highlighting scheme in Figure 7d is able to do so, whilst still enabling the user to identify hot and cold spots easily.

For the field experiments, data were captured using ROS Electric on an Ubuntu 11.10 installation. Camera specifications are summarized in Table I.

TABLE I: Sensor specifications

Sensor	Thermal Camera	RGB-D Camera
Make	Thermoteknix	Microsoft
Model	Miricle 307K	Kinect
FOV	60°	57°
Resolution	640×480	640×480
Framerate	25 FPS	30 FPS
Wavelength	6 – 14μm	-
Range	-20 – 150 °K	-
NEDT	85mK	-

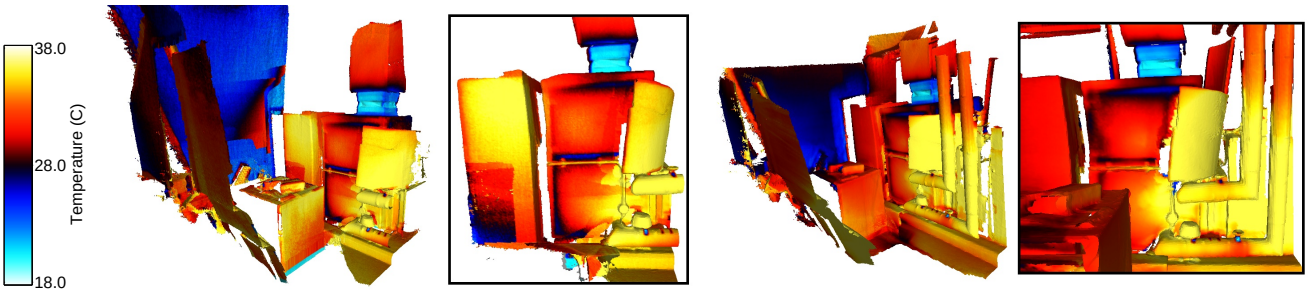
We acquired 7 data sequences from 3 different environments for the evaluation. For all environments, one illuminated and one dark sequence were captured. For environment 3, an additional sequence was captured with the maximum duration between NUCs reduced from 45 seconds to 10 seconds. The following is a description of the 3 environments:

1) A large room containing various HVAC-related equipment (Figure 8)
2) An underground hot water system (Figure 9)
3) A reverse cycle air-conditioning unit (Figure 10)

For environment 1, we focussed on investigating the performance of the system under different lighting conditions. From the results in Figure 8, it can be seen that the appearance of the thermal-only model is virtually identical regardless of the lighting conditions. Temperature readings are also consistent, being unaffected by the presence or lack of visible-spectrum illumination. However, it should be noted that visualizing the model using I-H Mapping without lighting offers no benefit over the thermal-only representation. The utility of the thermal-highlighting representation scheme is demonstrated by its ability to draw attention to hot and cold patches on the surfaces in the scene.

For environment 2, we investigated the geometric characteristics of the model. The camera path shown in Figure 9 includes a pose-axis for a set of arbitrary keyframes, with red, green and blue representing the x, y and z axes respectively for the corresponding camera coordinate frame. The geometric measurements extracted from this model were found to correspond well with physical measurements. The ability to accurately measure distances between points in 3D space using the 3D thermal model has many applications. For example, in the context of industrial plant monitoring, the distance between a faulty component with an abnormally high temperature and a vulnerable component nearby can be closely tracked. In addition, estimates of energy loss can be made by calculating the surface area of a hot patch.

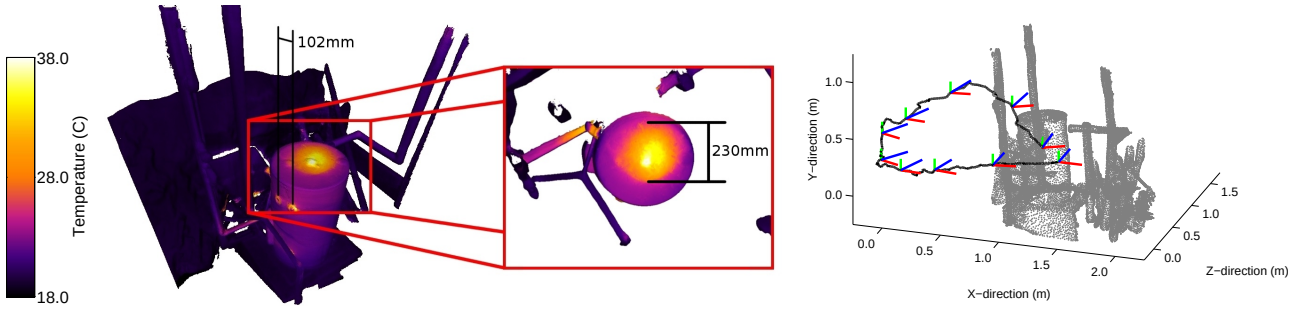
For environment 3, we explored the usefulness of the I-H Mapping method for visualizing 4 dimensions (red, green, blue and thermal-infrared) in a colorized 3D model. A major advantage of our approach is the ability to read text-based labels whilst still simultaneously visualizing the thermal data.



(a) Illuminated (b) Dark

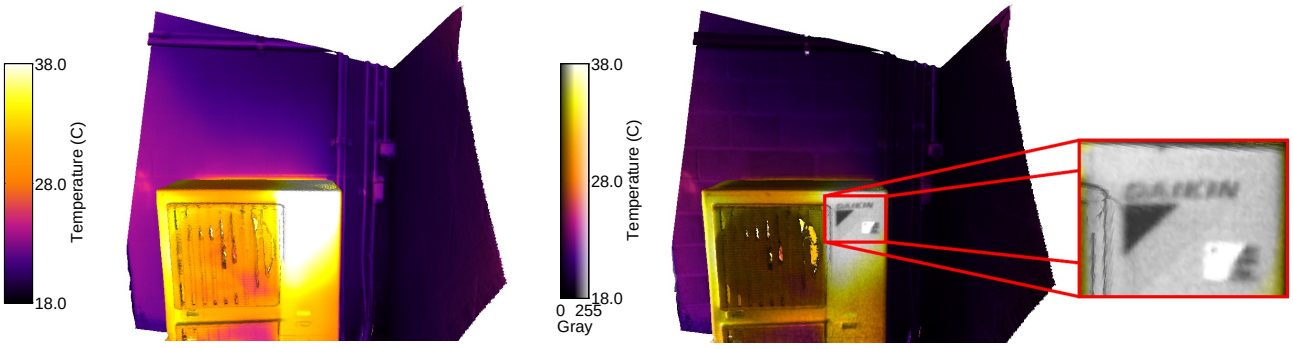
(c) Thermal only (d) RGB Only (e) Thermal Highlighting

Fig. 8: Experiment 1: Demonstration of illumination invariance and incorporation of visible-spectrum RGB.



(a) Thermal-only model with geometry overlay (b) Camera path

Fig. 9: Experiment 2: Demonstration of the geometric integrity of the model.



(a) Thermal-only representation (b) Fused Representation with highlighted label

Fig. 10: Experiment 3: Demonstration of the effectiveness of fused representation for scene investigation.

This is shown in Figure 10.

Many of the major limitations of our system are those common to RGB-D based camera systems, such as a difficulty in functioning outdoors under strong sunlight, and the requirement that all surfaces must be within several meters range of the sensor. In addition, accurate thermal-infrared values cannot be obtained from surfaces which have atypical thermal emissivity, such as many polished metals, and glass.

VIII. Conclusion

We have presented a highly mobile system capable of producing high-fidelity 3D models incorporating both color and surface temperature data. The system is the first of its kind in that it is hand-held, comprised of readily available hardware and can perform 3D thermal mapping in total darkness.

The cost of the system is limited simply by the cost of a single thermal camera and a single RGB-D camera. Data acquisition can be done quickly as the system functions in continuous operation, with a computer stored in a wearable unit such as a backpack. Being lightweight and hand-held, the system can be used to map environments not easily accessed by a wheeled platform such as staircases and cluttered areas. The operation of the system and interpretation of results can be done by an inexperienced operator with only brief, informal instruction.

The output of our proposed system could see immediate applications in areas which would benefit from rapid 3D visualization or analysis of surface temperature information. These include fire management and response, as well as within electrical and structural building inspection. We predict a particularly strong growth in demand for this technology in the area of building energy audits for improving energy efficiency, based on the recent increase in literature regarding these techniques.

Planned future work includes conducting a user study for an objective understanding of the effectiveness of different 4D-3D representations, and exploring the potential of adaptive visualization schemes. In addition, the system will be improved to operate in more extended environments in real-time.

Acknowledgment

The authors would like to thank undergraduate student intern Obadiah Lam for the development of a multispectral visualization software package which was used for experimentation and to generate several of the images used in the paper.

References

[1] U.S. Energy Information Administration, "Annual energy review," Free Pr, USA, Tech. Rep., 2010. 1

[2] D. Borrmann, A. Nuchter, M. Dakulovic, I. Maurovic, I. Petrovic, D. Osmankovic, and J. Velagic, "The project ThermalMapper thermal 3D mapping of indoor environments for saving energy," in *Proceedings of the 10th International IFAC Symposium on Robot Control (SYROCO)*, vol. 10, 2012. 1, 2

[3] Y. Ham and M. Golparvar-Fard, "Rapid 3D energy performance modeling of existing buildings using thermal and digital imagery," in *Construction Research Congress*. ASCE, 2012. 1, 2

[4] Y. Cho and C. Wang, "3D thermal modeling for existing buildings using hybrid LIDAR system," in *Computing in Civil Engineering*. ASCE, 2011. 1, 2, 3

[5] J. Wardlaw, M. Gryka, F. Wanner, G. Brostow, and J. Kautz, "A new approach to thermal imaging visualisation, EngD Group Project," 2010. 1

[6] S. Vidas, R. Lakemond, S. Denman, C. Fookes, S. Sridharan, and T. Wark, "A mask-based approach for the geometric calibration of thermal-infrared cameras," *IEEE Transactions on Instrumentation and Measurement*, 2012. 2, 3

[7] P. Moghadam, M. Bosse, and R. Zlot, "Line-based extrinsic calibration of range and image sensors," in *The 2013 IEEE International Conference on Robotics and Automation*, 2013. 2, 4

[8] S. Izadi, R. A. Newcombe, D. Kim, O. Hilliges, D. Molyneaux, S. Hodges, P. Kohli, J. Shotton, A. J. Davison, and A. Fitzgibbon, "KinectFusion: real-time dynamic 3D surface reconstruction and interaction," in *ACM SIGGRAPH 2011 Talks*, 2011, p. 23. 2, 4

[9] S. Prakash, P. Y. Lee, and T. Caelli, "3D mapping of surface temperature using thermal stereo," in *Control, Automation, Robotics and Vision, 2006. ICARCV'06. 9th International Conference on*, 2006, p. 14. 3

[10] L. Hoegner and U. Stilla, "Texture extraction for building models from IR sequences of urban areas," in *Urban Remote Sensing Joint Event, 2007*, 2007, p. 16. 3

[11] D. Iwaszczuk, L. Hoegner, and U. Stilla, "Matching of 3D building models with IR images for texture extraction," in *Urban Remote Sensing Event (JURSE), 2011 Joint*, 2011, p. 2528. 3

[12] Z. Zhang, "A flexible new technique for camera calibration," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 11, p. 13301334, 2000. 3

[13] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and Vision Computing*, pp. 761–767, 2004. 3

[14] L. Zhang, Z. Liu, and C. Honghui Xia, "Clock synchronization algorithms for network measurements," in *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 1, 2002, p. 160169. 3

[15] Q. Zhang and R. Pless, "Extrinsic calibration of a camera and laser range finder (improves camera calibration)," in *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, vol. 3, 2004, p. 23012306. 4

[16] R. Unnikrishnan and M. Hebert, "Fast extrinsic calibration of a laser rangefinder to a camera," Carnegie Mellon University Robotics Institute," Research Showcase, 2005. 4

[17] A. Toet, "Natural colour mapping for multiband nightvision imagery," *Information Fusion*, vol. 4, no. 3, p. 155166, 2003. 5