Infrastructure-Free AR Headset Tracking System Using Thermal Sensors: Project Plan

CHANDAFA, Abraham Ernest

October 1, 2025

1 Introduction

Augmented Reality (AR) technology is one of the most rapidly expanding fields of the 21st century [1]. With vast applications in many fields such as creating 3D virtual tours for real estate, advanced surgical simulations in medicine, and immersive 3D gaming experiences [2], Augmented Reality creates an immersive experience by enabling real-time superimposition of digital information and virtual objects onto the physical world.

For a great immersive experience, one of the key tasks for Augmented Reality (AR) systems is to accurately estimate the 6-Degree-of-Freedom (6-DoF) pose of the camera. As pointed out by Xu et al. [3], modern image-based camera pose estimation techniques can be categorised into two paradigms: **structure feature-based localisation**, which recovers the camera pose by establishing correspondences between features in a query image and a 3D model of the scene, and **regression-based localisation**, where deep learning models, such as PoseNet, are trained to directly regress the camera pose from an input image in an end-to-end fashion.

To estimate camera pose, these traditional pose estimation techniques rely on visible texture and patterns in the environment, a method that faces serious drawbacks in poorly textured environments [3]. In the absence of good texture and patterns, artificial physical patterns, such as QR codes or remotely projected patterns can be introduced [4]. But introducing physical patterns is often invasive and not ideal for most scenarios, and projecting patterns is inefficient for dynamic vision tasks because the projector moves with the camera, as illustrated in figure 1. As will be shown, a thermal camera and laser projector can achieve what a normal camera and projector cannot: draw a pattern that "sticks" to the surface and track that pattern in the thermal domain.

2 Background

Thermal sensing and imaging have undergone a significant expansion in recent years, driven by their increasing adoption across diverse domains. As highlighted by Wilson et al. [6], machine learning has enabled the advancement of thermal sensing technology beyond traditional military and defense applications, where thermal cameras can easily identify intruders in the dark, as well as medical applications, where thermal imaging helps in medical diagnosis.
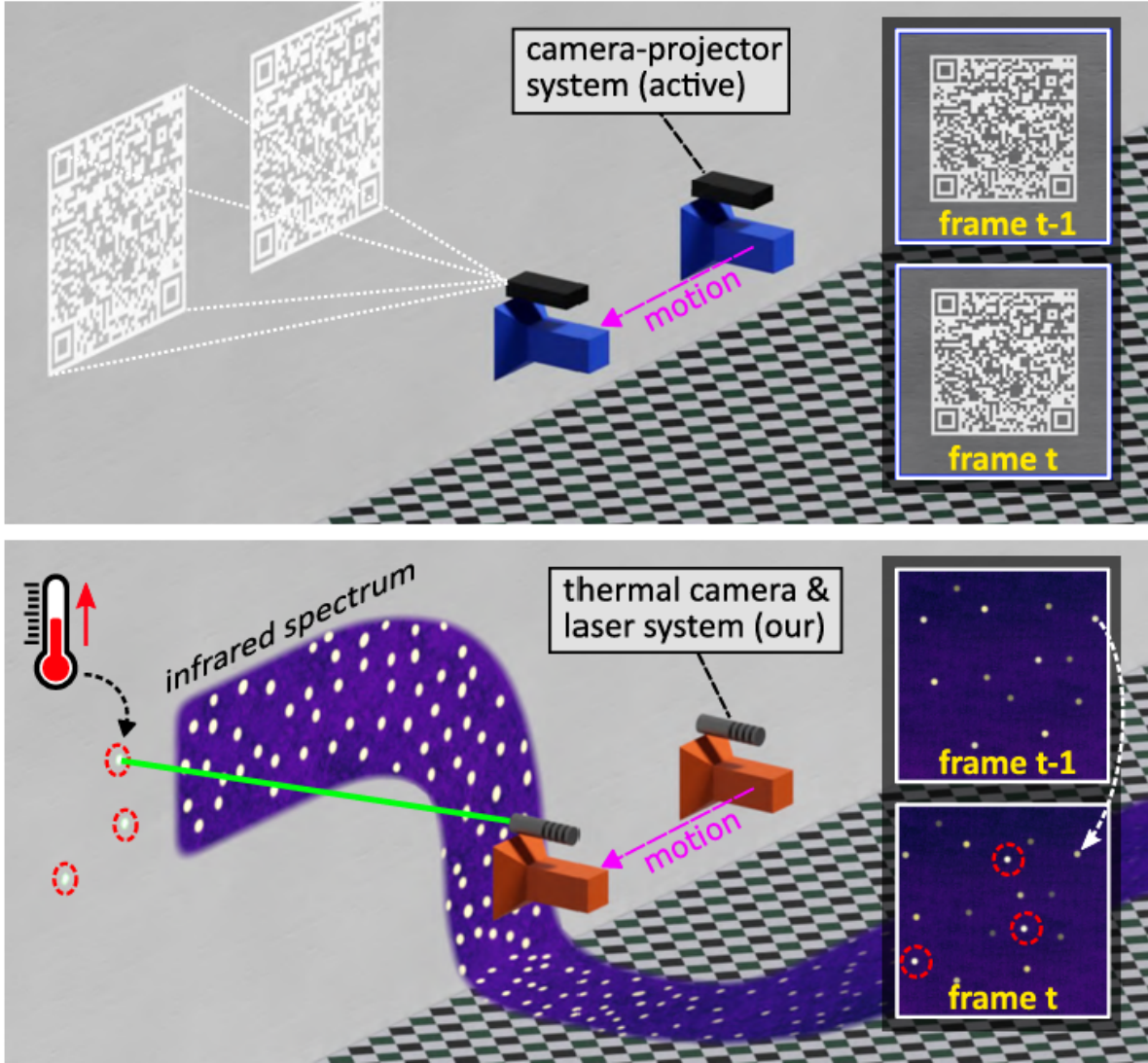
Figure 1: Projecting trackable thermal patterns. Adapted from [5]. (Top) A camera
projector system can introduce distinct features, but these features cannot be used
as a reference between frames because the projected pattern moves along with the
system. (Bottom) The laser can 'paint' a temporary heat pattern that 'sticks' to objects'
surfaces, and can, therefore, be tracked in the thermal domain.

One such pivotal advancement was introduced by Sheinin et al. [5], who proposed a novel framework for projecting trackable thermal patterns using a co-located thermal camera and steerable laser system. Their approach utilises the transient yet persistent nature of laser-induced thermal signature to "paint" features onto otherwise texture-less surfaces. These heat patterns can be tracked by thermal cameras, overcoming a fundamental limitation of traditional projected patterns that move with the projector.

A fundamental challenge with thermal patterns is heat diffusion. In order for a thermal pattern to be trackable, it needs to have a spatial structure with trackable feature points, and these feature points have to persist for at least two consecutive video frames [5]. Sheinin has addressed this heat diffusion through a learning-based approach called Projection-Diffusion Reversal (PDR), where they trained a neural network to remove newly added thermal points and reinforce existing points that have diffused, producing a consistent thermal pattern across frames.

While Sheinin et al. demonstrated the feasibility of thermal pattern projection for general computer vision tasks, its application to the specific, demanding problem of AR headset pose estimation remains unexplored.

3 Project Motivation

Motivated by the critical need for robust, infrastructure-free tracking for Augmented Reality headsets, this project builds upon the foundational work done by Sheinin et al. by adapting active thermal pattern projection for real-time estimation and tracking of the 6-degree-of-freedom (6-DoF) pose of an Augmented Reality (AR) headset.

As comprehensively surveyed by Zhou et al. [7] in their review of the ten years of International Symposium on Mixed and Augmented Reality (ISMAR) research, tracking has consistently been the most challenging and actively searched area in AR, constituting over 20% of all publications. See table 1. Traditional tracking methods still face fundamental limitations, despite decades of development:

- **Marker-based tracking**, while robust, requires maintenance, and often suffers from limited range and intermittent errors, because it provides location information only when markers are in sight, rendering them impractical for dynamic, large-scale environments [7].

- **Natural feature tracking** methods struggle in texture-poor spaces where distinguishable objects that can be used as markers are not always available, which often leads to a tracking failure [3, 7].

- **Model-based approaches** usually require the cumbersome process of modelling, especially when creating detailed models for large cluttered environments [7], thus limiting their scalability.

AR headsets are particularly sensitive to these limitations, since they must operate reliably in real-world environments, including dimly lit rooms, featureless corridors and hallways, and other visually challenging conditions.

By adapting Sheinin's thermal projection framework specifically for AR headset pose estimation, this work aims to overcome the critical limitations identified in the ISMAR

Topics	% Papers	% Citations
Tracking	20.1	32.1
Interaction	14.7	12.5
Calibration	14.1	12.5
AR App.	144	12.5
Display	11.8	5.4
Evaluations	5.8	1.8
Mobile AR	6.1	7.1
Authoring	3.8	8.9
Visualization	4.8	5.4
Multimodal	2.6	0.0
Rendering	1.9	1.8

Table 1: Proportion of Highly Cited Papers. Adapted from [7]. This table considered all IS MAR/ISAR/ISMR/IWAR papers with an average citation rate of 5.0 or more per year.

survey and enable AR systems to work reliably under varying lighting conditions and textureless surfaces.

4 Project Objectives

The goal of this project is to develop a prototype system demonstrating real-time 6-DoF pose estimation for an AR headset using actively projected thermal patterns. To achieve this, the following objectives have been defined:

1. **System Design and Integration:** Design and assemble a hardware prototype comprising a thermal camera, a steerable laser projector, and computing hardware, configured to be head-mountable. The system will adopt Sheinin's coaxial configuration where "each camera pixel can be mapped to a corresponding outgoing projector ray" [5], ensuring predictable pattern projection regardless of scene geometry.

2. **Pattern Projection and Tracking:** Implement software to control the laser projector, creating sparse, trackable thermal dot patterns on surfaces in the environment. Following Sheinin's findings, we will use discrete patterns with "only a few dots per frame as opposed to 'continuous' patterns" [5] to maintain trackability, and develop a tracker to follow these dots in the thermal video stream.

3. **Pose Estimation Pipeline:** Develop and implement a software pipeline that uses the tracked 2D positions of the thermal dots to compute the 6-DoF pose of the headset relative to the room. This extends Sheinin's structure-from-motion approach, which demonstrated that "the system continuously projects new 'heat points' on the object's surface during motion" to generate accurate camera trajectories [5].

4. **Performance Evaluation:** Quantitatively evaluate the system's performance in terms of tracking accuracy (translation and rotation error), latency, and robust-

ness compared to traditional RGB-based tracking in both textured and textureless environments. We aim to achieve performance comparable to Sheinin's indoor localization experiment, which demonstrated "accurate tracks and loop closure" over a 20-meter trajectory [5].

5 Project Scope and Deliverables

Scope

- **Environment:** Testing will be conducted in controlled indoor environments.

-

- **Scale:** The system is designed for room-scale tracking (approximately 5x5 meter areas), not large-scale building or outdoor navigation.

- **Real-time Focus:** The primary metric is real-time performance sufficient for AR applications (target ¿15 Hz), prioritizing latency over ultimate accuracy, while maintaining the "temporal continuity relations across frames" essential for dynamic tracking.

- **Material Assumptions:** The initial work will assume surfaces responsive to laser heating (e.g., wall paint, wood, common office materials).

- **Pattern Density:** Following Sheinin's optimization, we will use sparse dot patterns rather than dense thermal textures, as they found that in order to maintain the patterns' trackability for a reasonable duration, discrete patterns having only a few dots per frame must be used [5].

Deliverables

- **A Functional Hardware Prototype:** A head-mounted rig integrating a thermal camera and laser projector in a coaxial configuration, capable of projecting and tracking thermal patterns in real-time.

- **Open-Source Software Repository:** Code for pattern projection, thermal feature tracking, and the 6-DoF pose estimation pipeline.

- **Final Project Report:** A comprehensive document detailing the system design, methodology, experimental results, and analysis.

- **Demonstration Video:** A video showcasing the system performing real-time pose estimation in both textured and textureless environments.

References

[1] T. A. Syed, M. S. Siddiqui, H. B. Abdullah, S. Jan, A. Namoun, A. Alzahrani, A. Nadeem, and A. B. Alkhodre, "In-depth review of augmented reality: Tracking

technologies, development tools, ar displays, collaborative ar, and security concerns," *Sensors*, vol. 23, p. 146, dec 2023. Academic Editors: Calin Gheorghe Dan Neamtu, Radu Comes, Jing-Jing Fang and Dorin-Mircea Popovici.

[2] M. S. Siddiqui, T. A. Syed, A. Nadeem, W. Nawaz, and A. Alkhodre, "Virtual tourism and digital heritage: an analysis of vr/ar technologies and applications," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 7, 2022.

[3] M. Xu, Y. Wang, B. Xu, J. Zhang, J. Ren, Z. Huang, S. Poslad, and P. Xu, "A critical analysis of image-based camera pose estimation techniques," *Neurocomputing*, vol. 570, p. 127125, 2024.

[4] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado, "A state of the art in structured light patterns for surface profilometry," *Pattern Recognition*, vol. 43, no. 8, pp. 2666–2680, 2010.

[5] M. Sheinin, A. Sankaranarayanan, and S. G. Narasimhan, "Projecting trackable thermal patterns for dynamic computer vision," in *Proc. IEEE/CVF CVPR*, 2024.

[6] A. N. Wilson, K. A. Gupta, B. H. Koduru, A. Kumar, A. Jha, and L. R. Cenkeramaddi, "Recent advances in thermal imaging and its applications using machine learning: A review," *IEEE Sensors Journal*, vol. 23, no. 4, pp. 3395–3407, 2023.

[7] F. Zhou, H. B.-L. Duh, and M. Billinghurst, "Trends in augmented reality tracking, interaction and display: A review of ten years of ismar," in *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 193–202, 2008.