POLITECNICO DI TORINO

MASTER's Degree in MECHATRONIC ENGINEERING

Master's Degree Thesis

Tracking People Indoors via Low-Cost Infrared Sensors

Supervisors

Prof. Mihai T. Lazarescu

Candidate

Bassel Mkanna

March 2026

Abstract

This thesis investigates privacy-preserving indoor people monitoring using a low-cost multizone infrared Time-of-Flight (ToF) sensor, with a specific focus on constrained doorway environments where human motion and door dynamics create challenging measurement conditions. The work is deliberately refocused toward a dataset-centric contribution: rather than proposing a single end-to-end people-counting pipeline, the primary objective is to produce a rigorous and well-documented experimental dataset that characterizes the behavior of the STMicroelectronics VL53L7CX in realistic edge cases relevant to residential and hotel-room entrances.

A complete acquisition platform was implemented using a VL53L7CX configured in 8×8 *mode (64 zones) operating at 15 Hz, interfaced through an STM32 NUCLEO-F401RE and logged via UART to a host computer. Each ToF frame was recorded as an* 8×8 *distance matrix (mm) with invalid measurements encoded as a sentinel value. Ground-truth video was captured using an iPhone 13 in ultra-wide mode at 4K resolution and 60 fps, with a millisecond timestamp overlay enabling post-processing temporal alignment with the ToF serial timestamps. The dataset was structured across three mounting configurations (ceiling-mounted, inclined above-door, and side-mounted) and five scenario classes designed to stress sensing robustness, including controlled door angles, entry and exit passages, non-crossing near-door behavior, fast walking, and object carrying.*

Representative synchronized trials demonstrate that mounting geometry strongly influences the measured depth signatures and event timing. The ceiling-mounted configuration provides the most stable coverage within the 2×2 *m region of interest, while the inclined and side-mounted configurations exhibit increased sensitivity to door geometry, occlusion, and delayed detectability due to sensing-volume mismatch. Practical limitations were also observed, including reduced stability for low-reflectance clothing, occasional non-detection of small objects due to coarse zone footprints, and undersampling effects under high-speed motion at 15 Hz. Overall, the thesis delivers a reproducible experimental methodology and a structured dataset intended to support future research on disturbance rejection and data-driven modeling for privacy-first indoor occupancy monitoring.*

Acknowledgments

To my supervisor, Prof. Mihai Lazarescu, for his continuous guidance, encouragement, and valuable feedback throughout the development of this thesis. His expertise and mentorship have been instrumental in shaping both the direction of the research and my own growth as a researcher.

I am also thankful to the faculty and staff of the Politecnico di Torino for providing an excellent academic environment and the resources necessary to carry out this work. Special thanks go to my colleagues and friends in the program, whose discussions and support have made the research process both productive and enjoyable.

I owe heartfelt thanks to my family for their unwavering support, patience, and encouragement during my studies. Their belief in me has been a constant source of motivation and strength.

Finally, I would like to dedicate this work to all those who have inspired me to pursue engineering and research, and who have instilled in me the values of perseverance, curiosity, and continuous learning.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The rapid development of building automation and the widespread deployment of Internet of Things (IoT) devices are transforming how indoor spaces are monitored and managed. In modern buildings, occupancy information is increasingly used to improve energy efficiency, space utilization, and safety. In particular, Heating, Ventilation, and Air-Conditioning (HVAC) systems can be controlled more effectively when real-time evidence of human presence and movement is available, rather than relying on static schedules or conservative design assumptions.

Despite the importance of occupancy data, accurate people monitoring in private indoor environments remains a challenging engineering problem. Doorways and narrow passageways represent a critical monitoring location because they concentrate movement events (entry and exit) into a small region. However, these environments also amplify disturbances: door motion, partial occlusions, varied walking speed, and human-object interactions can generate ambiguous signals that degrade reliability.

This thesis addresses the doorway monitoring problem using a privacy-preserving sensing approach based on low-cost infrared Time-of-Flight (ToF) technology. Rather than focusing primarily on developing a complete end-to-end counting algorithm, the work is deliberately refocused toward creating a rigorous experimental dataset and a scientific characterization of ToF behavior under realistic doorway edge cases. The result is a structured dataset of synchronized multizone ToF measurements and ground-truth video, designed to support reproducible analysis and future algorithmic development.

1.1 Global Context of Building Automation and Energy Optimization

The current practice of architectural design and building operation requires intelligent sensing systems which now function as essential requirements for both government regulations and economic demands. The Energy Performance of Buildings Directive (EPBD), recast in 2024, underscores the urgency of achieving decarbonized and highly energy-efficient building stocks by 2050.

The industry reports show that real-time occupancy signals which control HVAC systems result in energy savings between 10% and 30% through their ability to drive energy usage. The practice of people counting extends beyond its energy benefits to create significant socioeconomic effects.

In retail environments, footfall analysis informs staffing levels and merchandising

strategies, while in healthcare, non-intrusive monitoring provides a safety net for senior living without compromising dignity. The challenge exists because people movement through entranceways which serve as main indoor paths needs to be measured accurately. A system needs to operate at a single-door entrance which has a 36-inch width to identify three types of movement: human entry, human exit, and different types of mechanical or environmental disturbances.

1.2 The Paradigm Shift in Indoor Occupancy Sensing

The field of occupancy detection has historically relied on Passive Infrared (PIR) sensors combined with simple infrared break-beams for occupancy detection. The systems provide affordable operation and reduced energy needs but cannot deliver the detailed information necessary for current analytical procedures.

PIR sensors detect thermal radiation changes to function as binary triggers but they cannot determine human population numbers or movement patterns. The system experiences frequent false positive alerts because of environmental temperature changes while it fails to identify people who remain still.

RGB camera vision-based systems developed deep learning and Convolutional Neural Networks (CNNs) to achieve precise results in crowd counting and tracking through their innovative design. The "camera paradox" presents a major obstacle to camera use because it creates privacy issues through its precise tracking capabilities. The capture of identifiable images in residential environments presents serious ethical and legal problems which arise from regulations like the General Data Protection Regulation (GDPR). The requirement for real-time video processing needs expensive computational power which conflicts with the energy-efficient design of autonomous Internet of Things (IoT) devices.

Time-of-Flight (ToF) technology serves as the essential connector which links PIR systems that work with low accuracy to camera systems that use high levels of privacy intrusion. ToF sensors such as the STMicroelectronics VL53L7CX use invisible infrared pulse timing to measure distance which enables them to create 3D depth maps that maintain complete user privacy while delivering enough spatial detail for advanced people counting.

1.3 Technological Challenges in Constrained Passageways

The implementation of people counting through apartment-sized doors creates two types of difficulties which include physical challenges and algorithmic difficulties. The main technical challenge requires the system to identify and eliminate all non-human movements. The door movement creates continuous sound interference which occurs throughout the doorway area. The opening door creates a large surface area which the basic depth sensor system interprets incorrectly as human movement through the door.

The actual human movement patterns which include walking beside another person while following him and walking at different speeds create the need for high-resolution sensors which can detect human movements. The narrow passageway restricts the sensor's field of view which requires the sensor to detect the entire portal area from a distance which the design specifies. The VL53L7CX system contains a 90° diagonal field of view

which includes an 8x8 multi-zone matrix to provide a solution for these spatial limitations.

1.4 Objectives of the Present Research

The thesis establishes a scientific framework for bidirectional people counting in narrow corridors through the implementation of the VL53L7CX multi-zone ToF sensor. The research investigates how physical sensor placement affects the performance of algorithms used to process data. The researchers defined specific goals to accomplish their study.

- Design and execute a rigorous experimental protocol to acquire a synchronized dataset of multizone ToF depth frames and ground-truth video in a doorway environment.
- Characterize the VL53L7CX sensing behavior under controlled geometric conditions (mounting position and ROI definition) and operational disturbances (door angles, fast motion, non-crossing behavior, and carried objects).
- Provide a structured dataset organization (scenario matrix, logs, synchronized visual evidence, and reproducibility notes) enabling future comparative analysis and algorithm development.
- Identify and document typical failure modes and acquisition limitations of low-cost multizone ToF sensing in realistic edge-case doorway scenarios.

1.5 Thesis Structural Roadmap

The research will be conducted through multiple sequential stages to achieve its research objectives.

- Chapter 1 introduces the research context, motivation, and objectives.
- Chapter 2 presents the literature review on occupancy sensing technologies and Time-of-Flight sensing, with emphasis on doorway constraints and disturbance sources.
- Chapter 3 defines the problem setting and the architectural constraints of doorway monitoring using a multizone ToF sensor.
- Chapter 4 describes the dataset-centric research direction and motivates the experimental characterization approach adopted in this thesis.
- Chapter 5 details the experimental testbed, positioning matrix, scenario design, sensor/camera setup, and the synchronization and logging methodology.
- Chapter 6 presents the experimental results using representative synchronized trials across mounting configurations and scenarios, including typical acquisition limitations and failure cases.
- Chapter 7 concludes the thesis and discusses future work enabled by the released dataset.

Chapter 2

Literature Review

The existing research and technical literature survey develops a complete foundation which supports the design of a multi-zone ToF-based people counting system. The review starts with the examination of fundamental physical principles which enable indoor detection before it studies the VL53L7CX sensor architecture and ends with a detailed investigation of installation methods and signal integrity problems.

2.1 Theoretical Framework of Infrared Detection and Occupancy Monitoring

Occupancy sensing technology has developed through multiple technological generations which aimed to solve the shortcomings of earlier systems. The categorization of these technologies is essential for understanding the unique positioning of multi-zone ToF sensing.

Technology Category	**Primary Mechanism**	**Key Advantages**	**Major Limitations**
Mechanical/ Physical	Turnstiles, pressure mats	Definitive counts, 100% accuracy in controlled flow	Intrusive, high installation cost, hinders traffic
Simple Optical	IR break-beams, photocells	Low cost, reliable for single-file passage	Fails with clusters, triggers only once for groups
Thermal/ Radiometric	PIR sensors, thermopiles	Very low power, non-invasive, cost-effective	Binary only, misses stationary people, thermal noise
Vision-Based	RGB/Stereo cameras, AI analytics	Highest accuracy, multi-human tracking	Privacy invasion, high power, lighting dependent
Radio Frequency (RF)	WiFi radar, UWB, mmWave	Privacy-preserving, works through walls	Multipath interference, signal attenuation
Depth/Time-of-Flight	IR Laser Ranging (VL53 series)	100% privacy, absolute distance, low power	Range limited to $\sim$4 m, sensitive to sunlight

Table 2.1: Comparison of different people counting technologies

2.1.1 The Dominance of Privacy-Preserving Modalities

The people counting market reached its current state in 2025 which demands businesses to maintain exact measurement results while protecting complete user data security. Smart building systems now require privacy-by-design implementation as their fundamental design requirement. Traditional camera systems have lost their ability to monitor private areas because this shift favors non-visual identification sensor technologies. The role requires thermal sensors and ToF depth sensors to function as the main selectable options. ToF sensors are now more popular than thermal sensors because they deliver complete distance information which enables users to distinguish between human beings and stationary thermal sources such as radiators.

2.2 Comparative Analysis of Ranging Modalities for Narrow Passageways

The research needs to assess three different distance measurement methods which include Time-of-Flight (ToF) Doppler frequency shift and reverberation to reach optimum performance at an apartment door. The first rule permits you to identify people while they move through tight spaces, which leads to detecting their walking behavior within restricted environments.

2.2.1 Principles of Optical Time-of-Flight

Time-of-Flight technology operates according to light speed which remains constant. A sensor emits a pulse of infrared light and measures the precise time elapsed until the reflection is detected by the receiver. The term "direct Time-of-Flight" (dToF) describes this system. The time interval Δt for a target at distance d with light speed c is given by the formula:

$$\Delta t = \frac{2d}{c} \tag{2.1}$$

A 940 nm Vertical-Cavity Surface-Emitting Laser (VCSEL) functions as the emitter for the VL53L7CX system. The 940 nm wavelength was chosen because it remains completely undetectable by human vision which guarantees both safety and non-intrusive operation.

The Single Photon Avalanche Diode (SPAD) array serves as an essential part of the ToF system. The SPAD produces a digital timestamp through its avalanche current which activates when a single photon strikes the sensor. The sensor achieves its millimeter-level accuracy at high frame rates through this technology. The integration of these SPADs into a multi-zone matrix allows the VL53L7CX to generate a mini depth-map which is far more descriptive than a single distance point.

2.2.2 Theoretical Foundations of Doppler Frequency Shift

The Doppler effect describes the frequency change of a reflected wave that occurs when an object moves away or toward the transmitter. The Doppler shift in radar and microwave sensing systems shows the relationship between their Doppler shift (Δf) and the person's radial velocity (v) through this equation:

$$\Delta f = \frac{2v f_t \cos\theta}{c} \tag{2.2}$$

The equation shows that f_t represents the transmitted frequency and θ defines the angle that exists between the movement direction and the sensor's line of sight. Recent research has leveraged this principle to extract "micro-Doppler signatures" which create distinct movement patterns that happen when limbs and torso move and chest expands during breathing. Doppler sensors in people counting applications show exceptional motion detection abilities which enable them to track people who approach through walls or heavy clutter environments. The system cannot accurately measure distances which makes it difficult to use for counting people at doorways. The Doppler sensor fails to distinguish between two different walking behaviors because it cannot differentiate between a person walking through a door and a person walking past it in a parallel corridor. The sensor shows complete blindness to a person when they walk at a right angle to its view since their body causes a complete break in sensor visibility ($\theta = 90°$).

2.2.3 Acoustic Reverberation and Environmental Fingerprinting

The process of reverberation-based sensing requires the measurement of a room's acoustic impulse response. The human body serves as an advanced soundproofing system because its main component consists of soft tissue materials. The presence of people in a reverberant environment results in increased absorption coefficient values which lead to shorter reverberation periods according to Sabine's formula:

$$RT_{60} = \frac{0.161\,V}{S\,\alpha} \tag{2.3}$$

The equation defines volume as V and surface area as S. The active sensing system sends out an ultrasonic "chirp" which it uses to determine occupancy levels by measuring how fast the signal decays. The total number of people in large halls or auditoriums can be accurately estimated through reverberation sensing methods, but these techniques cannot reliably track the number of individuals who pass through doorways. Background noise and room furniture arrangements cause the acoustic properties of narrow passages to become highly unstable. The method needs a lengthy training period to create a specific "calibration" for every different environment, which prevents its use as an all-purpose door counting system. The first time optical ToF systems experience their equivalent of reverberation through their "multipath reflection" process when light reflects multiple times before reaching the receiver. The process creates errors which need to be controlled, but people see it as a beneficial aspect.

2.3 State-of-the-Art in Multi-Zone Time-of-Flight Architecture: VL53L7CX

The VL53L7CX represents the pinnacle of STMicroelectronics' FlightSense technology, specifically optimized for wide-field situational awareness. The architectural analysis demonstrates how this hardware system enables the implementation of door counting solutions.

2.3.1 Multizone Matrix and Spatial Resolution

The defining characteristic of the VL53L7CX is its ability to split the SPAD receiving array into separate zones. The sensor offers its default measurement system which includes:

- 8×8 resolution (64 zones) which enables each zone to function as a separate distance sensor that produces 64-point depth measurements for the door area.
- 4×4 resolution (16 zones) which uses zone aggregation to enhance signal-to-noise performance while enabling frame rates that reach 60 Hz.

The system uses its multi-zone feature as its main method to combat the break-beam failure problem which results in multiple people being counted as one person. The algorithm can identify two people who walk together as two distinct targets because they create different zone activation patterns at different horizontal positions within the 8×8 matrix.

2.3.2 Optical Innovation: Metasurface Diffractive Optical Element

The VL53L7CX uses an efficient metasurface lens which functions as a Diffractive Optical Element (DOE) because it needs to produce wide-angle views that extend from door headers. The DOE enables the projection of a consistent $60° \times 60°$ square illumination onto the scene. The system enables 90° diagonal field of view operation which delivers 33% more visual range than the previous generation (VL53L5CX) and provides visual detection at peripheral areas which resembles camera systems. The sensor creates a detection area which extends to both the entrance and most of the approach path area from its 2 m high monitoring point because this FoV needs to support both directional logic operations and tailgating detection functions.

2.3.3 Histogram-Based Distance Computation

The VL53L7CX utilizes a sophisticated histogram-based processing chain. The sensor collects photon arrival data by creating a temporal histogram which tracks all incoming photons to every detection zone. The data structure enables the sensor to identify multiple targets which exist at various distances inside a single detection zone. The Compact Normalized Histogram output serves as a fundamental technological advancement for this sensor family because its designers created it specifically for Artificial Intelligence applications. The CNH data presents unprocessed image data which shows light intensity changes across different distances and this data can be used by edge-based neural networks to identify shapes and detect different types of floor materials. The histogram enables the algorithm to identify door movement through small visual obstructions which include dust particles that cover the lens because these particles would normally stop a single-zone sensor from functioning.

2.4 Physical Dynamics and Constraints of Installation Positions

The main factor which determines system performance happens because of how sensor and doorway space interact. The literature highlights three primary installation topologies

of practical interest for narrow passageways.

2.4.1 Above-Door Inclined Configuration

The sensor system has its sensor installed on the wall which extends from the door header point while it observes the approaching area through its downward facing sensor at an angle of (θ).

Pros:

- Anticipatory Detection: The system uses its wide approach path view to detect entry events which start up to several seconds before actual entry occurs.
- The path of a person walking towards the door results in a clear decrease in distance across a series of zones which provides a robust directional signal.
- Wall mounting is typically simpler than ceiling mounting because it requires less complex wiring and structural modification.

CONS:

- The distance measurements will show perspective distortion because they do not measure distance from the floor. Actual object height (H) can be derived through the equation

$$H = H_{\text{sensor}} - (D \cdot \cos\theta) \tag{2.4}$$

  which requires trigonometric calculation.
- The sensor's FoV directly intersects the door's arc of motion. The moving door creates a large surface area which activates various detection zones when it opens.
- The depth map shows artifacts because the person blocks their lower body with their head and shoulders while their limbs move through the light.

2.4.2 Vertical Ceiling-Mount Orthogonal Perspective

The sensor is fixed on the ceiling to observe the passage center.

Pros:

- Biometric Reliability: The system observes head-and-shoulders facial patterns which scientists believe to be the most stable facial recognition element. From this viewpoint, the anthropometric measurement of head circularity becomes most apparent.
- The orthogonal view allows two people to walk together because their bodies create two separate "bumps" which appear in the depth map.
- The sensor detects only the area which directly faces its mounting position because the door sensor detects its door opening range.

Cons:

- The sensor only detects movement within its immediate sensor range. Only a brief moment makes a person visible when they move through space at high speed.
- The refresh rate needs to be sufficient for fast-moving children and pets so that they do not disappear between frames.
- The ceiling should maintain a height range between 2.15 m and 3.5 m to achieve best operational results. High ceilings reduce signal-to-noise ratio because they permit fewer returning photons from the floor.

2.4.3 Side-Wall Lateral Acquisition at Waist Height

The sensor is installed on the doorframe or side wall at a height of 1 m which allows it to monitor the entire passageway.
Pros:

- Profile Density: This method captures the complete horizontal body width which produces a powerful measurement that delivers high photon return results.
- Which Wait-Time Accuracy System Works: The system effectively detects when people remain stationary because their body distance from the reader or door stays the same.

Cons:

- Total Occlusion: Two people who walk through together will have their close-to-sensor person hide all actual sensor data which leads to extreme counting errors.
- Privacy Concerns: Horizontal views capture more identifiable body-shape and posture information compared to top-down views which may be unacceptable for privacy-first applications.
- Limb Noise: The combination of arm movement and umbrella use together with bag carrying activities generates "phantom targets" which cause the counting system to record multiple events when only one person passes through.

2.5 Algorithmic Strategies for Disturbance Rejection and Signal Integrity

The system requires depth data from VL53L7CX to undergo multiple filtering processes and logic block evaluations in order to achieve commercial-grade counting results. The research work identifies door movement rejection as the primary obstacle which people currently face.

2.5.1 Spatial Footprint and Geometric Clustering

The spatial dimensions of human beings exhibit fundamental dissimilarity to those of doorways. A human body presents itself as an irregularly shaped object that exhibits unpredictable movement patterns. Human beings typically occupy a space that extends

from four to nine zones within the 8x8 depth matrix. The body of a human being maintains an irregular shape which causes these zones to exhibit different measurement distances because the head region remains near the sensor while the shoulder area stays farther away. A door maintains its existence as a flat surface because it consists of unbending material. The movement of a door activates all connected zones to open at once. The sensor detects door depth changes through consistent linear movement that occurs across both vertical and horizontal device sensors. The advanced algorithms utilize segmentation to establish groups of nearby zones which maintain identical distance characteristics. The system distinguishes between actual people and disturbances when a cluster exceeds half of the portal width. The multi-zone FoV tracking system represents the most effective method for counting because it establishes a virtual tracking sequence.

2.5.2 State-Machine Driven Sequence Analysis

The system divides the Field of View into two distinct zones which are categorized as Outer zones and Inner zones.

State	Condition	Interpretation
Idle	Depth Baseline = Floor/Wall	No motion in portal
Approach	Motion in Outer zone only	Potential entry event
Active	Motion in both Outer and Inner	Person is currently in the portal
Entry Verified	Inner zone active, Outer zone clear	Successful ingress passage
Exit Verified	Outer zone active, Inner zone clear	Successful egress passage

The state-machine logic of this system automatically rejects stationary objects that include door entryway chairs because these objects do not fulfill the required zone transition sequence. The system uses this logic to reject door movements because when a door opens and remains open the system enters a new state but the system does not "clear" the zone which enables the baseline filter to treat the door as an unchanging part of the environment.

2.5.3 Temporal Baseline Adaptation

The environment in which a sensor operates is not static. The empty room baseline changes because light levels and temperature shifts and furniture movement create disturbances. The continuous change detection system applies exponential smoothing to establish background models which update for each of the 64 zones according to the formula

$$B_t = (1 - \alpha)B_{t-1} + \alpha D_t \tag{2.5}$$

where B_t represents the baseline at time t and D_t denotes the current distance measurement and α serves as a minimal smoothing parameter. The system maintains its accuracy during extended periods when a door remains partially open because the open door establishes a new baseline while people moving through the door gap create detection thresholds for system operation.

2.5.4 Frequency and Variance Stream Analysis

The research project started in 2024 to study the frequency patterns which represent how humans walk compared to how machines move. Human walking displays a regular walking speed which ranges from 1 Hz to 3 Hz while their limbs move in patterns which create depth signal fluctuations between 3 Hz and 11 Hz. The algorithm uses Fast Fourier Transform (FFT) together with variance stream analysis of multiple zones to separate the door's continuous low-speed movement from a person's unpredictable high-speed movement. The door movement detection method through frequency fingerprinting works effectively for side-wall door installations which maintain a narrow proximity between the door arc and the detection sensor.

2.6 Advanced Algorithmic Trends (2024-2025)

Research projects now focus on Machine Learning (ML) implementation for edge computing as the main research direction during the current research period. The transition to this new system occurs because the VL53L7CX provides organizations with access to high-quality multi-zone data.

- 1DCNN-LSTM Models: The system uses one-dimensional Convolutional Neural Network (CNN) techniques to extract spatial characteristics from depth data, and it employs Long Short-Term Memory (LSTM) units to analyze the time-based sequence of reading material. The models achieve more than 95 percent accuracy for testing situations that involve multiple challenges in high-traffic environments.
- Knowledge Distillation: Researchers use highly accurate teacher models which they trained on video data to create student models that operate with ToF data. The ToF sensor system enables itself to acquire human movement visual information by using complex visual cues which it processes without requiring actual photograph recognition.
- Spiking Neural Networks (SNNs): For battery-powered independent door counters SNNs provide an energy-efficient solution which replaces conventional deep learning techniques. SNNs operate when distance changes reach significant levels because they only respond to "event" occurrences which lead to distance changes.

2.7 Strategic Summary and Synthesis of Findings

The current state-of-the-art review which analyzes all available technologies establishes a complete ranking system that assesses both sensor performance and practical implementation success within narrow passageway environments.

Table 2.2: Comparison of Occupancy Sensing Technologies

Counting Strategy	**Typical Accuracy**	**Power Profile**	**Complexity**
Binary PIR	<60% (No headcount)	Ultra-Low	Low
Ultrasonic Reverberation	80–85%	Moderate	High (Per-room calibration)
Microwave Doppler	85–90%	Moderate	Moderate (Direction sensitive)
Multi-Zone ToF (8x8)	93–97%	Low (5.4mW)	Moderate (State-machine logic)
RGB-AI Camera	98–99%	High	Very High (Privacy risk)

The scientific evidence shows that the VL53L7CX multi-zone ToF sensor represents the best hardware solution for this Master's project. The system achieves spatial awareness through its 90° FoV and 8x8 depth mapping system which delivers spatial awareness that was previously accessible only through camera systems inside a compact energy-efficient module that safeguards user privacy. The upcoming experimental phase will focus on these key experimental findings:

- Mounting Trade-offs: The vertical ceiling-mount offers superior accuracy than any other mounting method for basic head-counting, while the inclined wall-mount enables users to see approaching targets better.
- Rejection triggers: The system requires door movements to be filtered through two methods which include spatial "planar coherence" checks and temporal sequence-based state machines.
- Algorithmic Enhancement: The two false counts which arm swings and carried items generate, will undergo reduction through the application of variance-based gait identification which uses CNH histogram data.

The literature review establishes the essential technical and theoretical foundation required to proceed with Chapter 3 which will demonstrate these concepts through experimental testbed design and synchronized ground-truth data collection.

Chapter 3

Problem Definition and Architectural Constraints

This chapter establishes a complete description of the hardware-software co-design problem which enables the implementation of the VL53L7CX sensor within residential entryway environments. The study aims to identify an appropriate distribution of zones with corresponding timing elements which results in reduced counting errors while meeting strict power and bandwidth requirements.

3.1 Target System Description

The research focuses on a real-time occupancy monitoring system which consists of three main hardware components.

- The **Sensing Layer** includes one STMicroelectronics VL53L7CX multizone ToF sensor which performs continuous depth map acquisition.
- The **Processing Layer** uses an ESP32-S3 or STM32-NUCLEO microcontroller MCU to run the counting process and door rejection state machine operations.
- The **Communication Layer** uses a low-power wireless technology such as Bluetooth LE or WiFi to send occupancy data which has been processed to building management systems.

The system timeline is divided into distinct time intervals called frames. The MCU requires approximately 16.6 ms to process an 8×8 depth matrix and update the occupancy state at a frame rate of 60 Hz. The system requires an algorithm which operates with high efficiency while using minimal memory resources.

3.2 Dynamic Characteristics of the Detection Space

The doorway environment requires modeling of its essential physical characteristics which include:

- Portal Dimensions: A standard apartment door header is at a height of 2.0 m to 2.2 m, with a width of approximately 0.9 m. The VL53L7CX's 60° vertical FoV results in a floor-level detection span of approximately 2.3 m at this height, which is more than sufficient to cover the portal and the immediate side-walls.

- Mechanical Cycle of the Door: The door requires between 0.8 and 2.0 s to complete its standard swing. The "planar triggering" of the sensor zones during this cycle must be modeled as a deterministic transition sequence to allow for its successful rejection.
- Human Biometrics: A person passing through the door is characterized by a "depth bump" relative to the floor. The system requires calibration to identify objects reaching heights of 1.0 m for children and 2.0 m for adults while excluding pets and floor-based debris.

3.3 Objectives of the Counting Algorithm

The algorithm's first phase aims to create a two-way counting system that satisfies all specified requirements. The system must establish two fundamental relationships between the detected objects and their movement across the monitored area to correctly determine entry and exit events.

3.4 Operational Constraints and Trade-offs

The development of the people counting system is subject to a "Trilemma of Constraints". The use of depth maps ensures privacy protection, as no visual images are recorded; however, this approach requires additional processing capabilities to correctly identify and manage detected objects that traditional cameras could track more easily. Furthermore, the system relies on active I2C communication links to send and receive data between multiple devices at predefined intervals. While this communication enables synchronization and real-time data exchange, it also increases power consumption and requires careful timing management.

- Accuracy vs. Privacy: The use of depth maps ensures privacy but requires more sophisticated logic than visual cameras to avoid occlusion errors.
- Refresh Rate vs. Signal Quality: Increasing the frame rate to 60 Hz (4x4 mode) reduces "motion blur" but decreases the photon integration time, making the sensor more sensitive to ambient IR noise.
- Power Consumption vs. Real-time Response: High-frequency I2C communication and matrix processing consume more energy, necessitating the use of autonomous "Wake-on-Motion" features for battery-powered longevity.

These parameters and constraints establish the scientific framework within which the experimental system will operate, allowing the evaluation of theoretical models and algorithms discussed in the literature review.

3.5 Conclusion of the Scientific Review and Theoretical Formulation

The conclusion of the scientific review and theoretical formulation showed that the preceding chapters established a strong technical base which supports high-precision people counting through VL53L7CX sensor technology. The research has successfully:

- Identified Time-of-Flight (ToF) as the superior modality for balancing directional counting accuracy with absolute user privacy.
- Characterized the VL53L7CX architecture, specifically its multi-zone SPAD array and wide diagonal FoV, as a key enabler for narrow passageway monitoring.
- Systematically analyzed the mounting topologies, contrasting the pre-entry awareness of inclined wall-mounts with the biometric clarity of orthogonal ceiling-mounts.
- Defined the methodology for disturbance rejection, highlighting the use of planar coherence analysis and sequence-driven state machines to filter door movements.

The implementation phase needs this state-of-the-art literature synthesis because it provides essential elements which complete the required components. The identified gaps—namely the lack of standardized performance comparisons across inclined, vertical, and lateral positions—set the stage for the researcher's upcoming experimental validation. The thesis will use synchronized depth and video data to determine the best setup for people counting systems which protect user privacy in contemporary urban residential spaces.

Chapter 4

Research Focus and Dataset-Centric Methodology

The initial conceptual direction of this research aimed to design and implement a complete bidirectional people counting system based on multi-zone Time-of-Flight sensing. However, preliminary experimental observations quickly revealed that the fundamental difficulty of doorway monitoring does not primarily lie in implementing counting logic, but rather in understanding how depth sensing behaves under realistic environmental disturbances and geometric constraints.

For this reason, the scope of the research was intentionally refined to focus on the creation of a rigorous, well-documented experimental dataset that characterizes the behavior of a low-cost multi-zone infrared Time-of-Flight sensor in challenging doorway scenarios. This dataset-oriented approach allows the investigation to move beyond simplified laboratory demonstrations and instead capture the complex edge cases that strongly influence the reliability of real-world occupancy monitoring systems.

The resulting contribution of this thesis is therefore not limited to algorithmic implementation, but rather the systematic documentation and analysis of sensing behavior under a wide range of controlled experimental conditions. Such datasets are essential for developing robust algorithms and for enabling reproducible research within the domain of privacy-preserving indoor sensing.

4.1 Motivation for a Dataset-Centric Approach

In recent years, many publications have proposed algorithmic frameworks for occupancy detection using Time-of-Flight sensors. However, a common limitation in the literature is the lack of publicly available datasets that capture the sensor response in realistic conditions, particularly in narrow passageways where environmental disturbances and human behaviors interact in complex ways.

Most existing studies evaluate algorithms using limited test sequences or controlled demonstrations that do not adequately represent the full range of real-world situations. For example, many works consider only simple entry and exit events, while ignoring critical edge cases such as multiple individuals crossing simultaneously, door movement interference, or subjects carrying objects that alter their geometric profile.

The absence of detailed datasets creates a major barrier for scientific progress in this field. Without standardized experimental observations, it becomes difficult to compare algorithmic approaches or understand the fundamental limitations of the sensing modality.

The dataset created in this research therefore aims to address several key needs:

- **Comprehensive scenario coverage**: capturing a wide range of human movement patterns and environmental conditions.
- **High temporal fidelity**: synchronizing depth sensor data with video ground truth to allow precise event analysis.
- **Mounting configuration comparison**: documenting how sensor placement affects sensing reliability.
- **Edge-case documentation**: identifying situations that produce false detections, missed detections, or ambiguous measurements.

By focusing on systematic data acquisition rather than solely algorithmic development, the study provides a foundation that can support future work on signal processing, machine learning, and sensor fusion for occupancy monitoring.

4.2 Scope Refinement and Research Direction

The refined objective of the thesis can therefore be summarized as follows:

> *To experimentally characterize the behavior of a multi-zone infrared Time-of-Flight sensor when monitoring doorway environments under a diverse set of realistic human and environmental scenarios, and to construct a structured dataset enabling detailed analysis of sensing reliability and disturbance rejection.*

This refined focus leads to three principal research directions.

4.2.1 Experimental Characterization of Depth Sensing

The first direction investigates how the depth measurements produced by the sensor evolve as people interact with the monitored doorway. Particular attention is given to situations that produce ambiguous depth patterns, such as when two individuals pass simultaneously or when a moving door partially occludes the sensing volume.

Understanding these phenomena requires detailed observation of the spatial-temporal patterns present within the multi-zone depth matrix. The dataset therefore records complete depth frames rather than only processed detection results, allowing future analysis of raw sensing behavior.

4.2.2 Evaluation of Disturbance Sources

A second objective is the identification and analysis of disturbances that affect the reliability of infrared depth sensing in narrow passageways. Among the most significant disturbances are:

- Mechanical motion of the door during opening and closing cycles
- Partial occlusions caused by subjects walking closely together

- Variations in body geometry due to carried objects such as chairs or a fan
- Direction changes where a subject begins entering but reverses direction

Each of these situations produces distinct spatial patterns within the depth matrix, and documenting them provides valuable insight into the limits of simple threshold-based detection methods.

4.2.3 Critical Evaluation of Disturbance Rejection Approaches

The third research direction involves evaluating several disturbance rejection strategies reported in the literature. These approaches include geometric clustering, temporal filtering, and sequence-based state machines.

Rather than proposing a fully optimized algorithm, the goal of this study is to examine how these techniques behave when applied to real experimental data. The dataset therefore provides a platform for comparing different approaches and identifying their strengths and weaknesses.

4.3 Importance of Synchronization and Ground Truth

For experimental data to be scientifically meaningful, it must be possible to determine the precise relationship between sensor measurements and real-world events. This requirement is particularly critical in occupancy sensing, where timing differences of even a fraction of a second can lead to incorrect interpretations of entry or exit events.

To address this requirement, the experimental platform includes a visible-light camera that records the monitored scene simultaneously with the Time-of-Flight sensor measurements. The camera does not participate in the counting algorithm itself but instead serves as a ground-truth reference for human interpretation of events.

The synchronization of these two data streams is therefore a fundamental aspect of the experimental methodology. The dataset records both the depth measurements and the corresponding video sequences, enabling detailed analysis of how sensor responses correspond to actual human motion.

4.4 Scientific Contribution of the Dataset

The dataset generated through this experimental campaign provides several contributions to the research community.

First, it offers a detailed characterization of the sensing behavior of a widely used low-cost infrared ToF sensor when deployed in a realistic indoor environment. Such characterization is essential for understanding the physical limitations of the sensing modality.

Second, the dataset documents a wide variety of edge cases that are often overlooked in simplified demonstrations but are critical for practical deployments. These include complex motion patterns, environmental disturbances, and variations in sensor mounting geometry.

Third, the dataset establishes a reproducible experimental framework that can be extended in future research. By carefully documenting the test scenarios, sensor configurations, and synchronization procedures, the work provides a blueprint for future studies investigating privacy-preserving indoor sensing technologies.

4.5 Transition to Experimental Scenario Design

Having established the dataset-oriented research methodology, the next step is the systematic design of experimental scenarios that can capture the diverse conditions encountered in real-world doorway monitoring.

The following chapter therefore presents the detailed experimental design used to generate the dataset. It describes the movement scenarios performed by test subjects, the different door states that were considered, the mounting configurations used for the sensor, and the synchronization procedure employed to align depth measurements with video ground truth.

This structured experimental framework ensures that the resulting dataset captures both typical and challenging situations, providing a comprehensive basis for the analysis presented in the later chapters of this thesis.

Chapter 5

Experimental Scenario Design and Dataset Construction

5.1 Dataset-Centric Contribution and Scope

This thesis is refocused around the creation of a rigorous, well-documented dataset characterizing the behavior of a low-cost multizone infrared Time-of-Flight (ToF) sensor in a constrained doorway environment. Rather than positioning the work primarily as a single optimized people-counting algorithm, the core contribution is the systematic acquisition, documentation, and analysis of raw multizone depth measurements across (i) multiple sensor mounting configurations, (ii) controlled door states, and (iii) a structured suite of motion scenarios and disturbances. The dataset is designed to support reproducible comparative evaluation of door-rejection approaches and event-timing reliability.

5.2 Sensing Hardware and Acquisition Configuration

5.2.1 Time-of-Flight Sensor and Operating Point

All trials were recorded using the **STMicroelectronics VL53L7CX** in 8×8 **multizone mode** (64 zones). The sampling frequency was fixed at:

$$f_s = 15 \text{ Hz}, \qquad T_s \approx 66.7 \text{ ms}.$$

The sensor was operated in continuous ranging mode. Integration time was not tuned in this dataset release; instead, the configuration was held constant across all experiments to isolate the effects of geometry (mounting) and scenario variation.

5.2.2 Embedded Platform and Data Logging

The VL53L7CX Time-of-Flight sensor was interfaced using an STM32 Nucleo development board based on the STM32F4 microcontroller family (NUCLEO-F401RE) together with the X-NUCLEO-53L7A1 expansion board. This hardware platform provides a convenient ecosystem for rapid prototyping and evaluation of multizone ToF sensing applications. The sensor communicates with the microcontroller through the I^2C interface, which is used to configure the sensor operating parameters and retrieve the distance measurements generated by the device.

The VL53L7CX operates in multizone ranging mode and provides spatially resolved depth information in the form of an 8×8 grid of measurement zones, resulting in 64 independent distance readings per frame. Each zone corresponds to a small angular region within the sensor field of view, allowing the system to capture a coarse depth map of the observed area. Distance measurements are expressed in millimeters and represent the distance between the sensor and the nearest detected surface within each zone.

During operation, the STM32 microcontroller continuously acquires the sensor frames at a configured frequency of 15 Hz. The collected data are then transmitted to a host computer through a UART serial interface configured at a baud rate of 460800 bps. On the host computer, the serial data stream is captured and recorded using the Tera Term terminal software. Tera Term provides a logging functionality that stores the incoming serial stream directly to a text file while appending a precise timestamp to each received line.

The generated log file therefore contains both the sensor measurements and their corresponding timestamps with millisecond precision. Each measurement frame is printed sequentially as an 8×8 matrix of distance values, representing the spatial depth distribution observed by the sensor at a given instant. This timestamped logging mechanism is essential for the experimental validation stage, as it enables accurate temporal synchronization between the sensor measurements and the video recordings used as ground truth in the conducted experiments.

5.2.3 Ground-Truth Camera

Ground-truth video data were recorded using an iPhone 13 smartphone operating in ultra-wide camera mode with a resolution of 4K at 60 frames per second. This configuration was selected in order to obtain a large field of view while maintaining sufficient spatial and temporal resolution to observe the movements of participants within the monitored area. The high frame rate also allows accurate temporal comparison between the video frames and the measurements produced by the Time-of-Flight sensor.

A third-party camera application was used during the recordings to overlay a precise time reference directly onto the video stream. The timestamp is displayed in the format `HH:MM:SS.mmm`, providing millisecond-level precision. This embedded visual timestamp allows manual synchronization between the video data and the sensor measurements recorded in the serial log. By comparing the timestamps visible in the video with the timestamps present in the sensor log file, the two data streams can be aligned during post-processing.

For the primary experimental configurations, the main camera was mounted on the ceiling and oriented in a nadir (downward-facing) direction. The camera was positioned approximately above the center of the monitored region of interest (ROI), which corresponds to a floor area of 2×2 m located near the doorway. This placement ensures that the entire experimental area is clearly visible in the recorded video and provides a reliable ground-truth reference for the presence and movement of participants within the monitored space.

For the side-wall sensor configuration, an additional camera was introduced in order to capture the scene from a viewpoint similar to that of the sensor. In this case, the second camera was placed next to the ToF sensor so that both devices shared approximately the same perspective of the monitored area. This additional viewpoint facilitates the interpretation of the sensor measurements and improves the reliability of the manual

comparison between the sensor detections and the ground-truth observations.

5.2.4 Camera Test Validation

1. We tested first the fps of a random video captured from the camera mounted on the ceiling and this was the result:

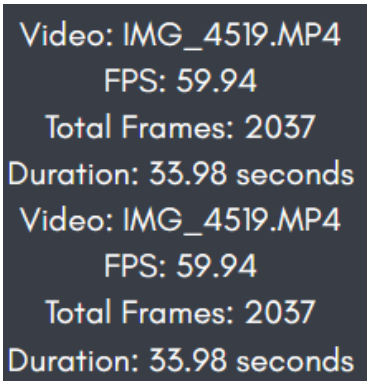


Figure 5.1: camera fps test on MATLAB.

As we see, The MATLAB result shown validate a fixed frame of 59.94 that is enough and good for our experiments.

2. To test the resolution, we used a chackerboard and put it in the middle of the FOV to check if the MATLAB can adapt and analyze it:

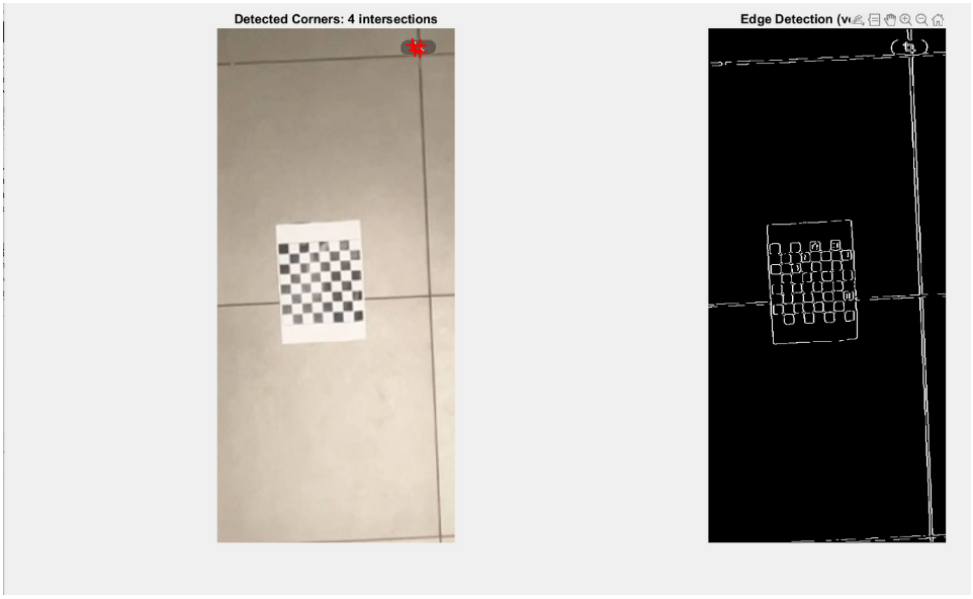


Figure 5.2: Resolution_test.

The Image on the left is a frame extracted from the video of the camera mounted. On the right you can see how the MATLAB processed the frame and succeeded to draw the cubes and especially the corners of a checkerboard putted on the center of our 2x2 area. That's a very good result to validate the resolution of our camera.

3. A video of the camera has been tested to verify its quality:

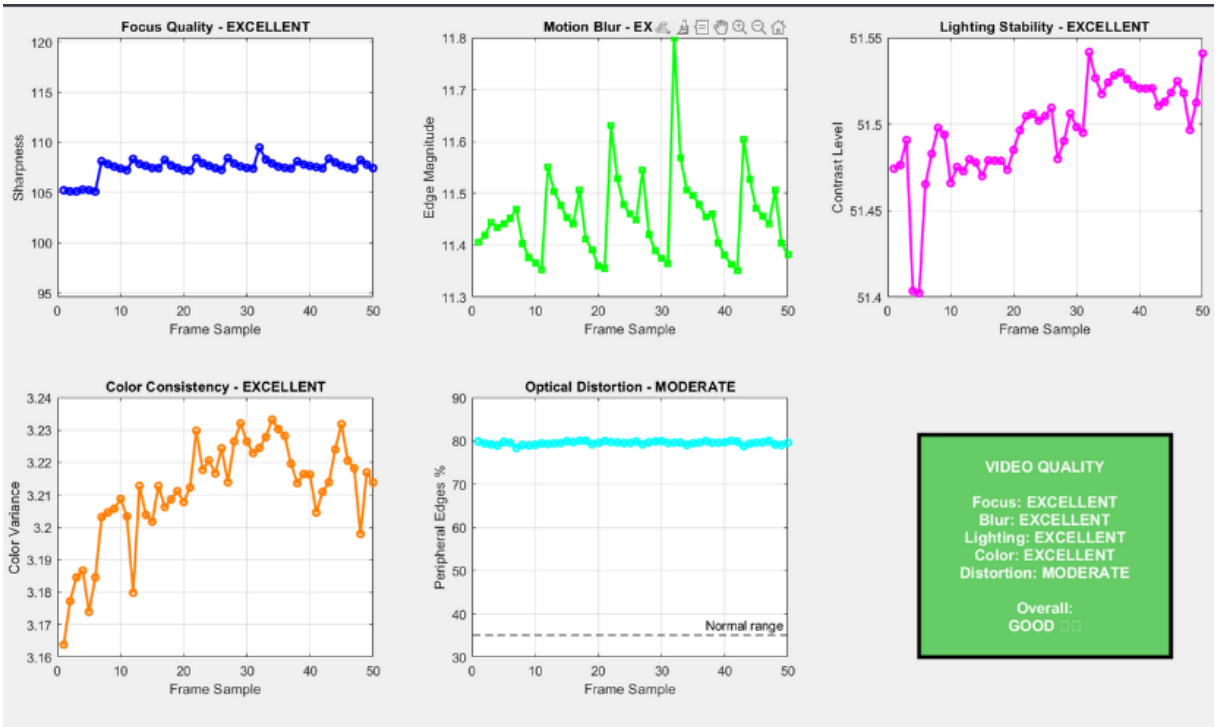


Figure 5.3: quality assessment.

The video quality assessment shows that the camera provides stable and reliable image acquisition suitable for sensor validation experiments. The focus quality is classified as excellent, with high and consistent sharpness values across sampled frames, indicating correct focusing and absence of defocus drift. Motion blur is also rated excellent, as the edge magnitude remains stable over time, confirming that camera motion and exposure settings do not introduce significant blur. Lighting stability exhibits minimal variation in contrast levels, demonstrating consistent illumination during recording. Color consistency remains stable across frames, indicating no automatic white balance or color drift effects. Optical distortion is classified as moderate, as a higher concentration of edges appears toward the image periphery; however, this level of distortion is typical for wide field-of-view lenses and remains acceptable for the defined 2×2 m monitoring area. Overall, these results confirm that the recorded video data are of sufficient quality to be used as a reliable reference for validating the depth sensor measurements.

5.3 Coordinate Frame, Doorway Geometry, and Environment

5.3.1 Doorway Coordinate System

A global reference frame $\mathcal{W}$ is defined:

- Origin O: center of the doorway threshold line on the floor.

- x: across the door width, left-to-right when facing into the room.
- y: into the room, normal to the doorway plane.
- z: upward.

5.3.2 Door Geometry and Door-State Parameterization

The doorway geometry is:

$$W_d = 1.0 \text{ m}, \qquad H_d = 2.2 \text{ m}, \qquad T_d \approx 0.05 \text{ m}.$$

The door opens into the room with hinges on the right side (facing into the room). Door state is parameterized by the door opening angle:

$$\alpha_d \in \{90^\circ, 60^\circ, 45^\circ, 30^\circ\}.$$

Door angles were measured and held using a protractor and manual positioning.

5.3.3 Region of Interest and Environment

The primary analysis region is a 2×2 m ROI marked on the floor using tape called also FOV :Field Of View). The ROI is centered approximately at:

$$(x, y) = (0, 1.0) \text{ m},$$

and is used consistently across all configurations. The floor material is gray ceramic tile. Experiments were conducted under light.

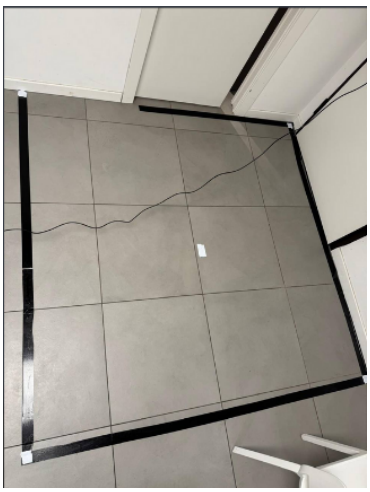


Figure 5.4: Field Of View.

5.4 Positioning Matrix: Sensor and Camera Placements

This study evaluates three sensor placement strategies: ceiling mount (three offsets), inclined above-door corridor-side mount, and side mount (implemented using a chair as a surrogate wall fixture).

5.4.1 Configuration A: Ceiling Mount (Three Offsets, Nadir)

The ceiling height of the experimental room is $H_c = 2.7$ m. In the ceiling-mounted configuration, the sensor is positioned approximately at the center of the doorway width and oriented in a nadir direction, meaning that the optical axis of the sensor is pointing vertically downward toward the floor. To evaluate the influence of the sensor placement relative to the doorway, three different offsets along the room depth direction were investigated. These positions correspond to translations along the y-axis of the adopted coordinate system and are defined as:

$$t_{A1} = (0, 0.30, 2.70), \quad t_{A2} = (0, 0.60, 2.70), \quad t_{A3} = (0, 0.90, 2.70) \text{ [m]}.$$

In this notation, the first coordinate represents the lateral displacement across the door width (x), the second coordinate represents the offset into the room (y), and the third coordinate represents the vertical height above the floor (z). Since the sensor is centered with respect to the doorway, the x coordinate remains equal to zero for all tested configurations.

To verify that the monitored region of interest (ROI) falls within the observable area of the sensor, a first-order geometric estimate of the footprint of the sensor field of view (FoV) can be performed. The VL53L7CX sensor has an approximate field of view of $60^\circ \times 60^\circ$. When the sensor is pointing directly downward (nadir orientation), the projected footprint on the floor can be approximated using simple trigonometric relations. The side length S of the square footprint can be estimated as

$$S \approx 2H_c \tan(30^\circ)$$

which yields

$$S \approx 2 \times 2.7 \times \tan(30^\circ) \approx 3.12 \text{ m}.$$

This estimate indicates that the sensor can observe an area of approximately 3.12×3.12 m on the floor when mounted at the ceiling height. Therefore, the selected 2×2 m region of interest used in the experiments lies comfortably within the observable footprint of the sensor. This ensures that the entire monitored area remains inside the sensor field of view during the evaluation of the ceiling-mounted configurations.

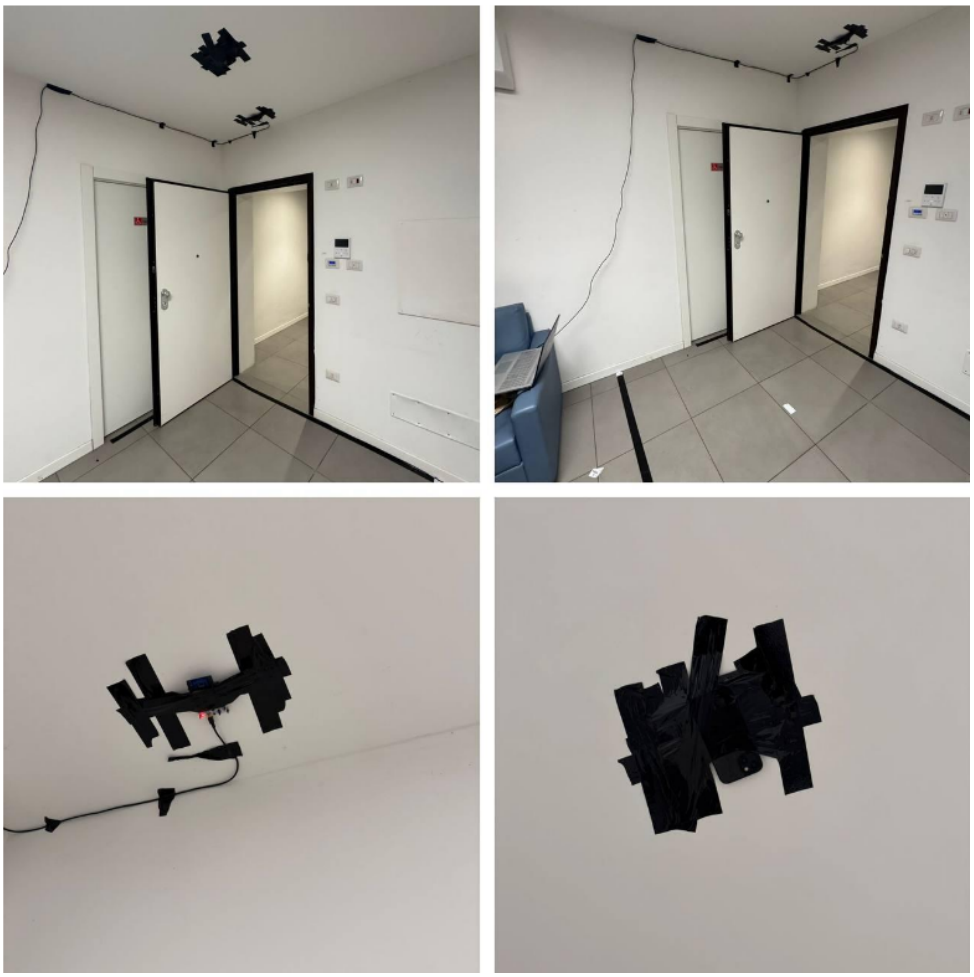


Figure 5.5: Ceiling Mount Setup.

5.4.2 Configuration B: Inclined Above-Door Mount (Corridor Side)

The inclined mounting configuration was implemented on the wall, with the sensor positioned above the door frame. The location of the sensor in the adopted coordinate system is defined as

$$t_B = (0, 0, 2.38) \text{ [m]}.$$

In this configuration, the sensor is centered along the doorway width and placed directly above the doorway threshold. The vertical coordinate corresponds to a height of 2.38 m above the floor, which approximately matches the upper region of the door frame. This placement allows the sensor to observe the passage area while remaining physically unobtrusive and representative of realistic installation constraints found in corridors, apartments, or hotel rooms.

Unlike the ceiling-mounted configuration, the sensor is not oriented in a nadir direction. Instead, the device is tilted downward toward the monitored region using a pitch angle of $\theta = 49°$. In this work, the pitch angle is defined with respect to the horizontal plane, meaning that the sensor optical axis forms an angle of 49° below the horizontal direction.

This inclination was selected empirically during preliminary testing in order to achieve a suitable alignment between the sensor field of view and the doorway region. In particular,

the chosen angle allows one full row of the 8×8 sensing zones to approximately align with a region parallel to the doorway plane. This alignment is useful for detecting objects or people crossing the doorway because it increases the likelihood that the transition between indoor and outdoor regions is captured by a consistent set of sensor zones.

By directing the sensor toward the doorway plane rather than vertically downward, this configuration emphasizes the detection of motion occurring along the entrance path. Consequently, the inclined mounting arrangement enables the evaluation of how well the sensor can detect entry and exit events when observing the doorway from an oblique viewpoint, which is a common installation scenario in real-world deployments.

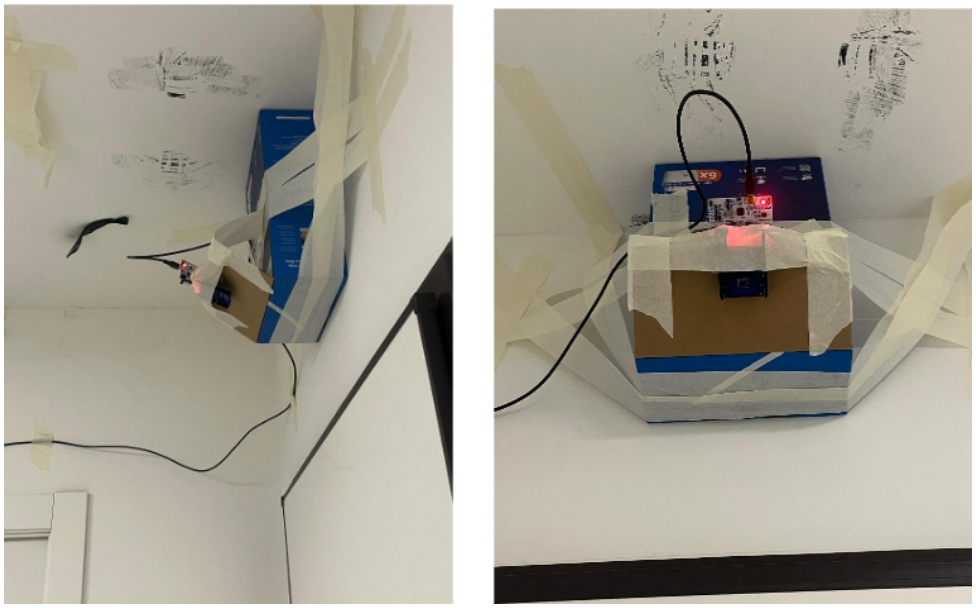


Figure 5.6: Above the door setup.

5.4.3 Configuration C: Side Mount (Chair-Based Surrogate Wall)

The side-mounted configuration was implemented using a chair as a surrogate mounting structure, since no permanent wall support was available at the desired location. This arrangement was adopted in order to reproduce, as closely as possible, a lateral wall-mounted sensing geometry within the available experimental environment. The sensor position in the adopted coordinate system is defined as

$$t_C = (-1.5, 1.0, 1.0)\ [\text{m}],$$

which corresponds to a location approximately 1 m inside the room along the depth direction, 1 m to the left of the left door frame, and 1 m above the floor. This placement was selected to represent a side-view sensing configuration in which the monitored doorway region is observed laterally rather than from above.

In contrast to the inclined configuration, the sensor in the side-mounted setup was not intentionally tilted downward by a predefined pitch angle. Instead, it was mounted directly against the supporting surface so that its optical axis was oriented straight ahead with respect to the local mounting plane. Therefore, the sensing direction was determined primarily by the physical orientation of the chair-based support rather than by a controlled angular adjustment. The sensor was aimed toward the monitored region of interest (ROI) without introducing a deliberate leftward or rightward bias.

This configuration is particularly relevant for studying the behavior of the ToF sensor under lateral observation conditions, where the measured depth patterns are strongly affected by body profile, partial occlusion, and the relative orientation of the person with respect to the sensor. Compared with ceiling-mounted or above-door configurations, the side-mounted geometry is observed to be more sensitive to variations in body posture and object carrying, since the sensor observes the subject from a side perspective rather than from a top-down or frontal oblique viewpoint. For this reason, the side-mounted setup provides an important experimental condition for evaluating the robustness and limitations of the sensor in less favorable but still practically relevant installation scenarios.

Figure 5.7: Side wall setup.

5.4.4 Main Camera Placement

The main camera is ceiling-mounted and aligned above the ROI center:

$$\mathbf{t}_{cam} \approx (0, 1.0, 2.7) \quad [\text{m}],$$

with nadir orientation. A secondary camera is used in side-mount tests and placed adjacent to the ToF sensor to provide a near-sensor viewpoint.

5.4.5 Positioning Summary Table

Field	Ceiling A1/A2/A3	Inclined B	Side C
ToF position $\mathbf{t}$ [m]	$(0, 0.30, 2.70)$; $(0, 0.60, 2.70)$; $(0, 0.90, 2.70)$	$(0, 0, 2.38)$	$(-1.5, 1.0, 1.0)$ (chair-based)
ToF pointing	Nadir	Pitch 49° down	Aim to ROI center.
Door state factors	α_d varied (S1,S2)	α_d varied (S1,S2)	α_d varied (S1,S2)
Expected primary risk	short dwell under fast motion	door-plane disturbance in slanted FoV	geometry-induced timing offset + surrogate mount vibration

Table 5.1: Mounting configurations and measured placement parameters in the doorway coordinate frame.

5.5 Scenario Suite: Experimental Protocols and Repetition Structure

Each mounting configuration includes five scenarios. All scenarios are single-person trials. Unless explicitly stated otherwise, each scenario condition is repeated 8 times.

5.5.1 Scenario S1: Entry With Static Door Angle

Purpose: quantify entry signatures and door-plane static disturbance effects.
Conditions: $\alpha_d \in \{90°, 60°, 45°, 30°\}$ with 8 trials per angle (32 trials per mount).
Protocol:

1. Set door angle α_d using a protractor; keep door fixed.
2. Subject starts in corridor at a marked position (outside the room).
3. On cue, subject crosses the threshold ($y = 0$) into the room at normal speed and continues into/through the ROI.
4. Repeat steps 1–3 for 8 trials at the same α_d.
5. Change α_d and repeat.

Expected ToF signature: coherent multizone cluster evolution, plus static planar components from door depending on α_d and mounting geometry.

5.5.2 Scenario S2: Exit With Static Door Angle

Purpose: symmetric counterpart to S1, establishes direction-dependent patterns.
Conditions: $\alpha_d \in \{90°, 60°, 45°, 30°\}$ with 8 trials per angle (32 trials per mount).
Protocol: identical to S1 but subject starts inside the room and exits to corridor.
Expected ToF signature: time-reversed zone transition relative to S1.

5.5.3 Scenario S3: Near-Door Non-crossing Behavior

Purpose: stress false-positive rejection when motion occurs near the portal without entry/exit.
Trials: 8 trials per mount.
Composite protocol (single trial):

1. Subject approaches the doorway, reaches near the threshold, and performs a U-turn before completing a crossing; returns to the starting location.
2. Subject traverses laterally from left-of-door to right-of-door near the doorway, pauses for several seconds, then continues without crossing the threshold.

Expected signature: partial activation and local motion energy without a full threshold transition.

5.5.4 Scenario S4: Fast Walking

Purpose: evaluate undersampling risk at 15 Hz under high-speed motion.
Speed: approximately 8 km/h (≈ 2.22 m/s).
At 15 Hz the displacement per frame is:

$$\Delta d \approx \frac{2.22}{15} \approx 0.148 \text{ m/frame.}$$

Trials: 8 trials per mount.
Protocol: subject performs a fast entry crossing similar to S1 but at approximately 8 km/h.
Expected signature: short-duration activation (few frames) and potentially discontinuous cluster observations.

5.5.5 Scenario S5: Crossing With Carried Object

Purpose: test silhouette/geometry disturbances from large carried objects.
Objects: portable fan or chair (object type recorded per trial).
Trials: 8 trials per mount.
Protocol: subject performs an entry crossing while carrying the object.
Expected signature: larger footprint cluster, possible phantom surfaces, and increased invalid/occluded zones.

5.5.6 Scenario Matrix Table

ID	Name	Primary factor(s)	Repetitions
S1	Entry + door angle	$\alpha_d \in \{90, 60, 45, 30\}$	8 per angle
S2	Exit + door angle	$\alpha_d \in \{90, 60, 45, 30\}$	8 per angle
S3	Non-crossing	u-turn + pause + lateral pass-by	8
S4	Fast crossing	≈ 8 km/h	8
S5	Carry object	fan or chair	8

Table 5.2: Scenario suite used to populate the dataset across mounting configurations.

5.6 Synchronization and Event Timing

5.6.1 Manual Synchronization Procedure

The synchronization between the Time-of-Flight (ToF) sensor measurements and the ground-truth video recordings was performed manually using the millisecond timestamps available in both data sources. Since the sensor and the camera were not triggered simultaneously by a shared hardware clock, a post-processing alignment procedure was required to establish a common temporal reference between the two streams.

The synchronization process was carried out according to the following steps:

1. The ToF sensor measurements were recorded through the serial interface and logged on the host computer using Tera Term. Each line of the serial log file contains a timestamp automatically appended by the logging software with millisecond precision. These timestamps correspond to the exact time at which the data were received by the computer and therefore provide a reliable temporal reference for the sensor measurements.

2. The ground-truth video recorded by the iPhone camera includes a visual timestamp overlay generated by a third-party application. The timestamp is displayed directly within the video frames using the format `HH:MM:SS.mmm`, which provides hour, minute, second, and millisecond resolution. This embedded time reference allows the experimenter to identify the exact moment at which a particular event occurs in the video.

3. For each experimental trial, the ToF log file and the corresponding video recording were inspected simultaneously. The beginning and end of each sequence were trimmed so that both the sensor data and the video frames correspond to the same temporal interval. This alignment ensures that the analyzed portion of the sensor measurements matches the same time window observed in the video recording.

Through this procedure, a common time window was established for both the ToF sensor data and the ground-truth video. This synchronized dataset enables accurate comparison between the detected sensor events and the actual movements observed in the recorded video sequences.

5.6.2 Formal Event Definitions and Measured Latency

To avoid ambiguity, event timing is defined as:

- t_{GT}: video time when torso center crosses threshold plane ($y = 0$) for crossing scenarios (S1,S2,S4,S5).

- t_{ToF}: first ToF frame meeting a detection criterion (defined in analysis code; e.g., motion energy threshold or baseline difference threshold with sufficient zone count).

Latency is:

$$\Delta t = t_{\mathrm{ToF}} - t_{\mathrm{GT}}.$$

Observed delays are interpreted as a combination of geometric sensing-volume mismatch and logging uncertainty, rather than assumed clock drift.

5.7 Raw Data Format and Preprocessing

5.7.1 Raw Time-of-Flight Log Format

Each ToF frame is recorded as a labeled block:

- one line: sensor label (CENTER),
- eight lines: eight integer distances per line (mm),
- one blank separator line.

Invalid/no-target zones are encoded as the sentinel value:

$$D_{ij} = 99999.$$

5.7.2 Cleaning Rules

Preprocessing converts invalid values to missing:

$$99999 \rightarrow \text{NaN}.$$

A frame validity ratio is defined:

$$\rho(t) = \frac{1}{64} \sum_{i,j} \mathbf{1}[D_{ij}(t) \neq \text{NaN}],$$

enabling coverage comparisons across mounts and door angles.

5.7.3 Baseline-Based and Motion-Based Representations

Two derived representations are used:

1. Baseline (height-map): $H_{ij}(t) = D_{ij}^{\text{empty}} - D_{ij}(t)$
2. Motion energy: $M_{ij}(t) = |D_{ij}(t) - D_{ij}(t-1)|$

Motion energy supports robust visualization even when the trial does not include a clean empty-scene baseline (notably S3).

5.8 Reproducibility and Limitations

5.8.1 Reproducibility Checklist

- Photos/diagrams of each mount with measured distances.
- Protractor measurement procedure for door angles.
- Consistent ROI definition and floor marking.
- Public parsing rules for ToF log blocks and invalid sentinel.

5.8.2 Known Limitations

- Manual synchronization (no hardware trigger).
- Side mount implemented with a chair as surrogate wall.
- Participant coverage not fully balanced across mounting configurations.

Chapter 6

Experimental Results: Representative Synchronized Trials

This chapter reports the experimental results obtained with the VL53L7CX multizone Time-of-Flight sensor under the three mounting configurations introduced in Chapter 5. In contrast to a purely statistical analysis over all trials, the present chapter adopts a *representative-trial* approach: for each mounting configuration and for each scenario (S1–S5), one synchronized trial is selected and discussed in detail. For each case, the synchronized evidence consists of (i) a video frame (or short segment) from the ground-truth camera with millisecond timestamp overlay, and (ii) the corresponding ToF measurements (8×8 depth map and/or temporal signal extracted from the log).

The purpose of this presentation is twofold. First, it provides an interpretable, human-verifiable description of the characteristic ToF signatures associated with each scenario. Second, it highlights the qualitative differences among mounting configurations, with emphasis on typical failure modes (e.g., delayed detection, partial occlusion, reduced valid-zone coverage, and ambiguity introduced by door geometry and carried objects).

Throughout this chapter, invalid ToF measurements are encoded using the sentinel value 99999 (mapped to missing values during visualization). All timestamps are reported in milliseconds and are aligned using the manual synchronization procedure described in Section 5.6.1.

6.1 How to Read the Synchronized Figures

Each synchronized figure pair follows the same interpretation logic:

- The ground-truth video provides the reference for the true physical event (crossing of the threshold plane, U-turn, pause, or object entry into the ROI). The relevant event timestamp is denoted t_{GT}.
- The ToF log provides the first measurement frame in which the event is observable according to ToF physics (appearance of a compact cluster, a large depth discontinuity, or an increase in motion energy). The corresponding time is denoted t_{ToF}.
- The observed delay (if any) is defined as:

  $$\Delta t = t_{\mathrm{ToF}} - t_{\mathrm{GT}}.$$

  A positive Δt indicates that the ToF sensor response appears later than the first visual evidence in the camera.

6.2 Dataset Table

The values reported in Table 6.1 were computed in MATLAB by scanning the dataset directory containing the trimmed, synchronized trials. For each mounting configuration, MATLAB automatically enumerated the corresponding video files and summed their durations using the `VideoReader` interface. In parallel, the total number of ToF frames was estimated from the logged sensor data by parsing each text log and counting the number of recorded 8×8 frames (identified by the per-frame header/tag and the fixed 8-line matrix structure). This procedure provides a reproducible and file-based quantification of dataset size in terms of trial count, total synchronized video time, and total ToF samples.

Mounting configuration	**# Trials**	**Total synced video time (min)**	**Total ToF frames**
Ceiling-mounted (A1/A2/A3)	264	17.20	15213
Side-wall mounted (C)	88	12.86	5843
Above-door inclined (B)	88	7.99	8357

Table 6.1: Dataset size summary by mounting configuration.

6.3 Configuration A: Ceiling-Mounted Sensor (Nadir)

The ceiling-mounted configuration provides an approximately orthogonal view of the region of interest. In this geometry, the subject typically produces a compact multi-zone activation pattern corresponding to the head-and-shoulders profile, while the floor remains comparatively stable. Since the ROI is fully contained within the sensor footprint (Chapter 5), the ceiling configuration is observed to provide the most reliable coverage and the lowest sensitivity to partial occlusion.

6.3.1 Scenario S1: Entry With Controlled Door Angles

In this entry scenario, the subject crosses the threshold plane and enters the ROI. The synchronized figure shows that the ToF sensor response appears as a localized cluster of valid depth values (non-99999) moving through the central zones as the person transitions from the doorway into the room. The depth values decrease relative to the empty-scene baseline, reflecting the shorter sensor-to-subject distance compared to the floor/background. The primary qualitative observation is that the ceiling geometry yields a clean separation between the moving human cluster and static background returns.

6.3.2 Scenario S2: Exit With Controlled Door Angles

In the exit scenario, the subject starts inside the room and moves outward through the doorway. The ToF signature is observed to be approximately the time-reversal of the entry case: the activated zone cluster first appears in the ROI and then migrates toward the doorway boundary. In the synchronized example, the ToF frames show the cluster shrinking and disappearing as the subject exits the sensor footprint. Any delay between the camera and the ToF detection was interpretd as the time required for the subject to occupy a sufficient portion of the multi-zone grid to produce a stable return.

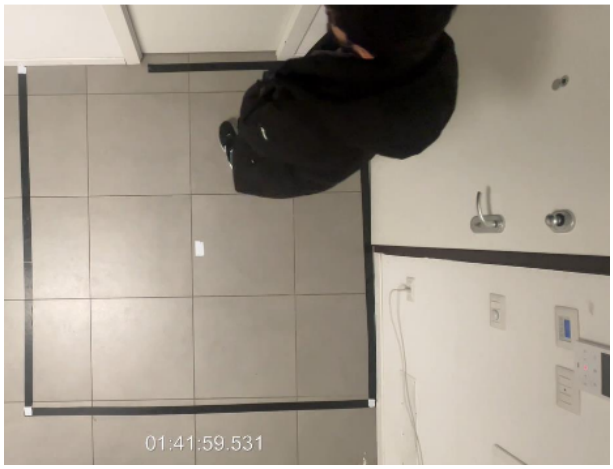
Figure 6.1: first frame captured of 30,60 and 90 cm mounted sensors

```
[2026-02-05 02:54:56.642] CENTER
[2026-02-05 02:54:56.642] 510 99999 99999 99999 490 488 480 461
[2026-02-05 02:54:56.642] 1189 99999 99999 99999 99999 99999 847 689
[2026-02-05 02:54:56.642] 99999 99999 99999 99999 99999 99999 1047 955
[2026-02-05 02:54:56.642] 99999 99999 99999 99999 99999 99999 99999 99999
[2026-02-05 02:54:56.642] 1169 99999 99999 99999 99999 99999 99999 99999
[2026-02-05 02:54:56.642] 99999 99999 99999 2734 2652 99999 99999 99999
[2026-02-05 02:54:56.642] 99999 99999 99999 99999 99999 99999 99999 99999
[2026-02-05 02:54:56.642] 99999 99999 99999 99999 1664 99999 99999 99999
[2026-02-05 02:54:56.642]
[2026-02-06 23:45:14.329] CENTER
[2026-02-06 23:45:14.330] 99999 99999 99999 99999 99999 99999 991 956
[2026-02-06 23:45:14.330] 99999 99999 99999 99999 99999 99999 1442 1395
[2026-02-06 23:45:14.338] 99999 99999 99999 99999 99999 99999 99999 99999
[2026-02-06 23:45:14.338] 1398 99999 99999 2673 99999 2741 99999 99999
[2026-02-06 23:45:14.338] 99999 99999 99999 99999 99999 99999 99999 99999
[2026-02-06 23:45:14.338] 99999 99999 99999 99999 99999 99999 99999 99999
[2026-02-06 23:45:14.347] 99999 99999 99999 99999 99999 99999 99999 99999
[2026-02-06 23:45:14.347] 1659 99999 99999 99999 99999 99999 1760 99999
[2026-02-07 01:13:13.088] CENTER
[2026-02-07 01:13:13.088] 37 99999 99999 99999 99999 1443 1371 99999
[2026-02-07 01:13:13.088] 1345 99999 99999 99999 99999 99999 99999 99999
[2026-02-07 01:13:13.088] 99999 99999 99999 99999 99999 99999 99999 99999
[2026-02-07 01:13:13.088] 99999 99999 99999 2713 99999 99999 99999 99999
[2026-02-07 01:13:13.088] 99999 99999 99999 99999 99999 99999 99999 99999
[2026-02-07 01:13:13.088] 99999 99999 99999 99999 99999 99999 99999 99999
[2026-02-07 01:13:13.089] 99999 99999 99999 99999 99999 99999 99999 99999
[2026-02-07 01:13:13.091] 99999 99999 99999 99999 1377 1373 99999 99999
```

Figure 6.2: data synchronized with the video

6.3.3 Scenario S3: Non-crossing Motion Near the Doorway

Scenario S3 is designed to test false-positive susceptibility by introducing motion near the doorway without a true entry/exit event. The representative example includes (i) a U-turn close to the threshold plane and/or (ii) lateral motion parallel to the doorway followed by a pause. In a robust counting system, these motions should not be misinterpreted as an entry/exit event. The ceiling-mounted ToF typically captures partial activation in peripheral zones without a consistent threshold-crossing pattern. The synchronized figure is used to describe how the observed ToF activation remains localized and does not exhibit a full transition across the ROI. Any spurious activation persisting over several frames should be discussed as a potential false-positive mechanism, particularly if it resembles a doorway-crossing pattern.




```
2026-02-07 01:41:59.509] CENTER
2026-02-07 01:41:59.509] 1480 1484 1474 1582 1394 1424 1384 99999
2026-02-07 01:41:59.509] 1708 99999 99999 99999 99999 99999 1851 99999
2026-02-07 01:41:59.509] 99999 99999 1050 99999 99999 99999 99999 99999
2026-02-07 01:41:59.509] 99999 99999 1052 1123 99999 99999 99999 99999
2026-02-07 01:41:59.509] 99999 1104 1183 99999 99999 99999 99999 99999
2026-02-07 01:41:59.509] 99999 99999 99999 99999 2655 99999 99999 99999
2026-02-07 01:41:59.509] 99999 99999 99999 99999 99999 99999 99999 99999
2026-02-07 01:41:59.509] 99999 2611 99999 99999 99999 99999 99999 99999
2026-02-07 01:41:59.509]
```

Figure 6.3: Synchronized data with the video of the U turn

For the stop:

```
2026-02-07 01:43:26.094] CENTER
2026-02-07 01:43:26.094] 15 1446 1475 1553 1434 1463 1395 99999
2026-02-07 01:43:26.104] 1739 99999 99999 99999 99999 99999 99999 99999
2026-02-07 01:43:26.104] 99999 99999 99999 99999 99999 99999 99999 99999
2026-02-07 01:43:26.104] 99999 99999 1057 1079 99999 99999 99999 99999
2026-02-07 01:43:26.104] 99999 99999 994 1091 99999 99999 99999 99999
2026-02-07 01:43:26.104] 99999 99999 1105 99999 99999 99999 99999 99999
2026-02-07 01:43:26.111] 99999 99999 99999 99999 99999 99999 99999 99999
2026-02-07 01:43:26.111] 99999 2579 99999 99999 99999 99999 99999 99999
2026-02-07 01:43:26.111]
2026-02-07 01:43:26.173] CENTER
2026-02-07 01:43:26.173] 1473 1510 99999 1474 1476 1428 1392 99999
2026-02-07 01:43:26.173] 1618 99999 99999 99999 99999 99999 1997 99999
2026-02-07 01:43:26.173] 1745 99999 1808 99999 99999 99999 99999 99999
2026-02-07 01:43:26.173] 99999 99999 1065 1046 99999 99999 99999 99999
2026-02-07 01:43:26.173] 99999 99999 1038 1086 99999 99999 99999 99999
2026-02-07 01:43:26.173] 99999 99999 1105 99999 2665 99999 99999 99999
2026-02-07 01:43:26.173] 99999 99999 99999 99999 99999 99999 2630 99999
2026-02-07 01:43:26.173] 99999 2580 99999 99999 99999 99999 99999 99999
2026-02-07 01:43:26.173]
2026-02-07 01:43:26.233] CENTER
2026-02-07 01:43:26.233] 1530 1460 99999 99999 1407 1447 1407 99999
2026-02-07 01:43:26.233] 1696 99999 99999 99999 99999 99999 1835 99999
2026-02-07 01:43:26.233] 99999 99999 1108 99999 99999 99999 99999 99999
2026-02-07 01:43:26.233] 99999 99999 1048 1078 99999 99999 99999 99999
2026-02-07 01:43:26.233] 99999 99999 1056 1064 99999 99999 99999 99999
2026-02-07 01:43:26.234] 99999 99999 1114 99999 99999 99999 99999 99999
2026-02-07 01:43:26.234] 99999 99999 99999 99999 99999 99999 99999 99999
2026-02-07 01:43:26.241] 99999 2643 99999 99999 99999 99999 99999 99999
2026-02-07 01:43:26.241]
2026-02-07 01:43:26.296] CENTER
2026-02-07 01:43:26.296] 1508 1473 1390 99999 1444 1433 1396 99999
2026-02-07 01:43:26.296] 1734 99999 99999 99999 99999 99999 1804 99999
2026-02-07 01:43:26.296] 99999 99999 1111 99999 99999 99999 99999 99999
2026-02-07 01:43:26.296] 99999 99999 1032 1078 99999 99999 99999 99999
2026-02-07 01:43:26.298] 99999 99999 1029 1093 99999 99999 99999 99999
2026-02-07 01:43:26.299] 99999 99999 1108 99999 99999 99999 99999 99999
2026-02-07 01:43:26.300] 99999 99999 99999 99999 99999 99999 99999 99999
2026-02-07 01:43:26.300] 99999 2617 99999 99999 99999 99999 99999 99999
2026-02-07 01:43:26.302]
```

Figure 6.4: Data of the stopping process

6.3.4 Scenario S4: Fast Crossing (Approximately 8 km/h)

In the fast-walking scenario, the subject crosses the ROI at approximately $v \approx 8$ km/h (≈ 2.22 m/s). Given the sensor sampling period $T_s \approx 66.7$ ms, the expected displacement per frame is on the order of $\Delta d \approx vT_s \approx 0.15$ m. Therefore, the subject may traverse the doorway region in only a few frames, increasing the risk of undersampling and discontinuous cluster observation. The synchronized example illustrates how the multi-zone pattern becomes shorter in time (fewer frames activated) and may exhibit higher variability in the number of valid zones. Any missed detection (e.g., only one frame showing activation) is important evidence of the frame-rate limitation under high-speed motion.

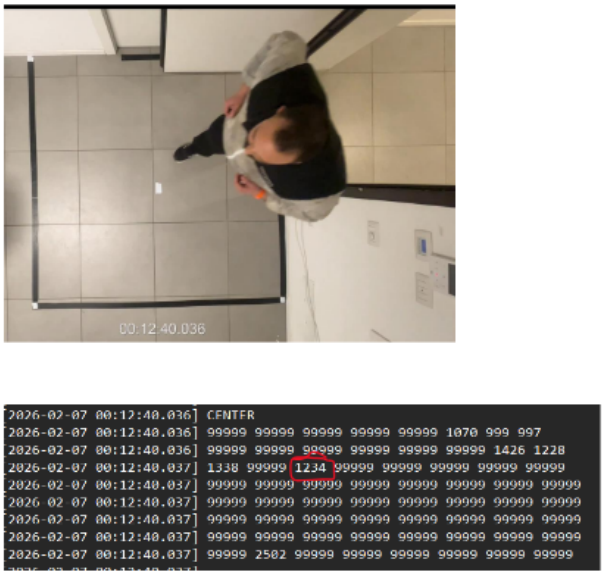


Figure 6.5: Data and video synchronized of the speed

6.3.5 Scenario S5: Entry With Carried Object (Fan/Chair)

Carrying large objects alters the subject's effective shape and reflectivity. In the ToF frames, this typically appears as an enlarged cluster footprint and/or additional depth surfaces at a different distance than the torso (e.g., the object leading the motion). The synchronized trial showed additional activation associated with the carried object and occasional partial occlusion of the body cluster. In ceiling mounting, such effects are often reduced compared to lateral mounting because the sensor sees the top-down projection of the combined silhouette.

Figure 6.6: Data and video synchronized

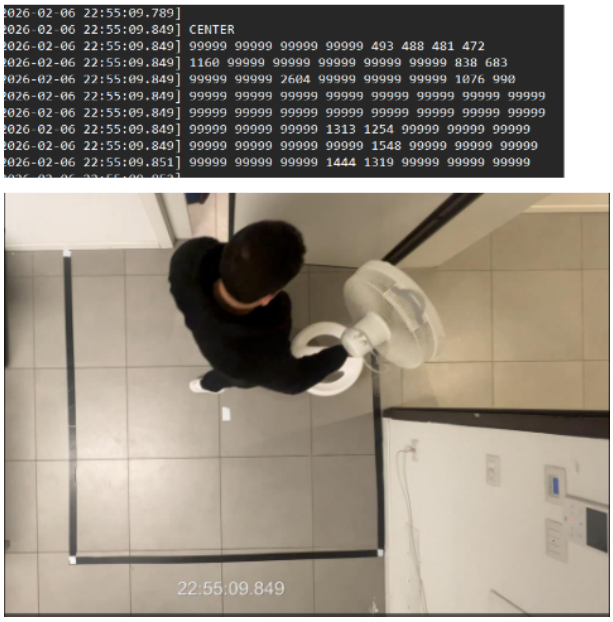


Figure 6.7: Data and video synchronized

6.4 Configuration B: Inclined Above-Door Sensor (Corridor Side)

The inclined configuration observes the doorway region from an oblique angle. Compared to the ceiling mount, this geometry increases the sensitivity to the door plane and to perspective distortions: a planar surface (door) may occupy a larger portion of the FoV and produce spatially coherent activations. As a result, door angle and occlusion effects are expected to play a stronger role.

6.4.1 Scenario S1: Entry With Controlled Door Angles

In the inclined entry scenario, the subject approaches the doorway and enters the ROI. Because the optical axis is oblique, the person may appear in the sensor measurements only after moving sufficiently into the sensing cone. The synchronized example represents (i) the door plane return (often a contiguous region of zones at a near-constant distance) and (ii) the subject return (localized changing cluster). If the door occupies part of the FoV, the ToF detection may be delayed relative to the camera, producing a larger Δt than in the ceiling case. The representative trial documents the observable separation (or overlap) between door and person signatures.

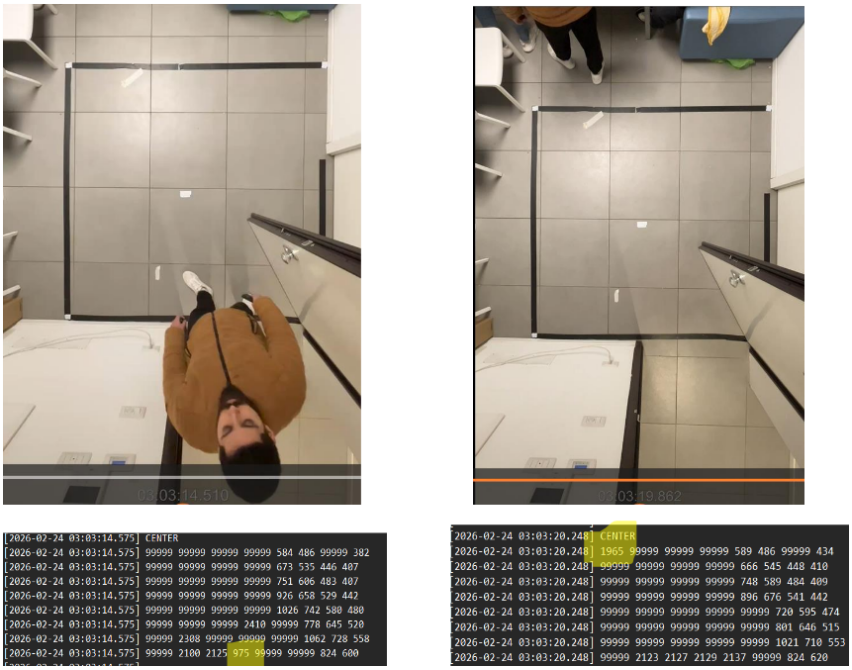


Figure 6.8: Data and video synchronized

6.4.2 Scenario S2: Exit With Controlled Door Angles

For exit motion, the inclined sensor may detect the subject later than the camera if the subject exits through regions partially masked by the door plane or if the subject presents a smaller reflective cross-section. The synchronized figure highlights that the subject remains detectable until crossing the threshold, and the signal drops early due to occlusion by the door or reduced reflectance.

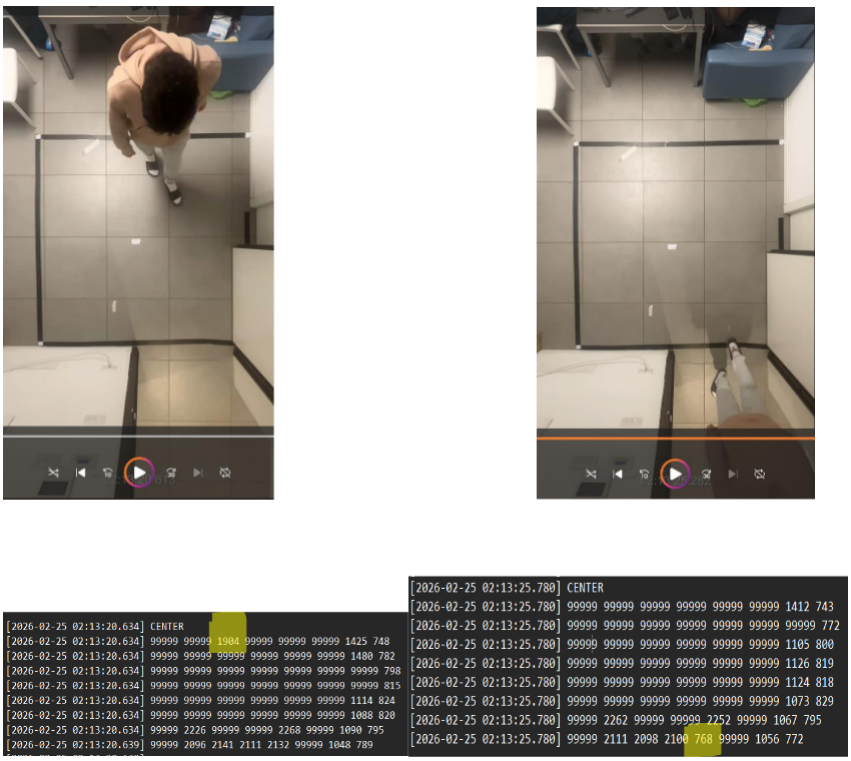
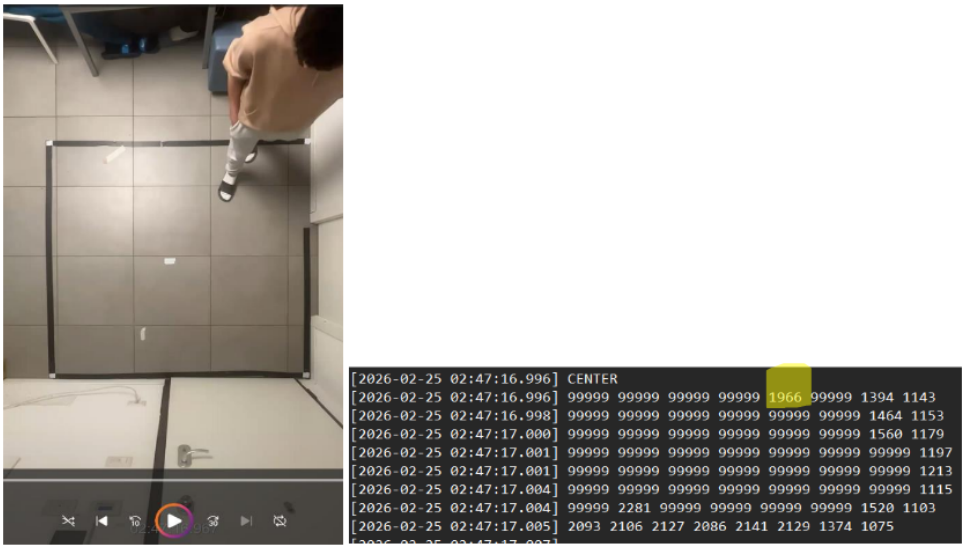


Figure 6.9: Data and video synchronized

6.4.3 Scenario S3: Non-Crossing Motion Near the Doorway

In S3, the inclined configuration is observed to be more vulnerable to false positives because motion near the door plane may resemble genuine threshold crossing when viewed obliquely. In the representative example, the U-turn and lateral motion produce partial activation patterns that do not complete a consistent entry/exit sequence. If the door is partially open, reflections from the door plane can generate apparent motion in multiple zones when the subject is nearby, which is discussed as a failure mode for door-rejection algorithms.



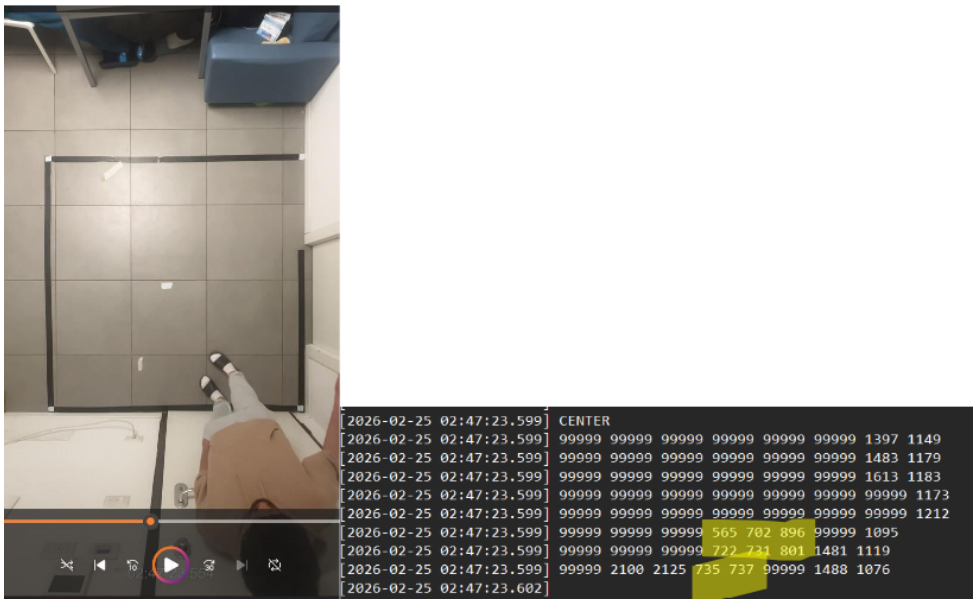


Figure 6.10: Data and video synchronized

For stopping:

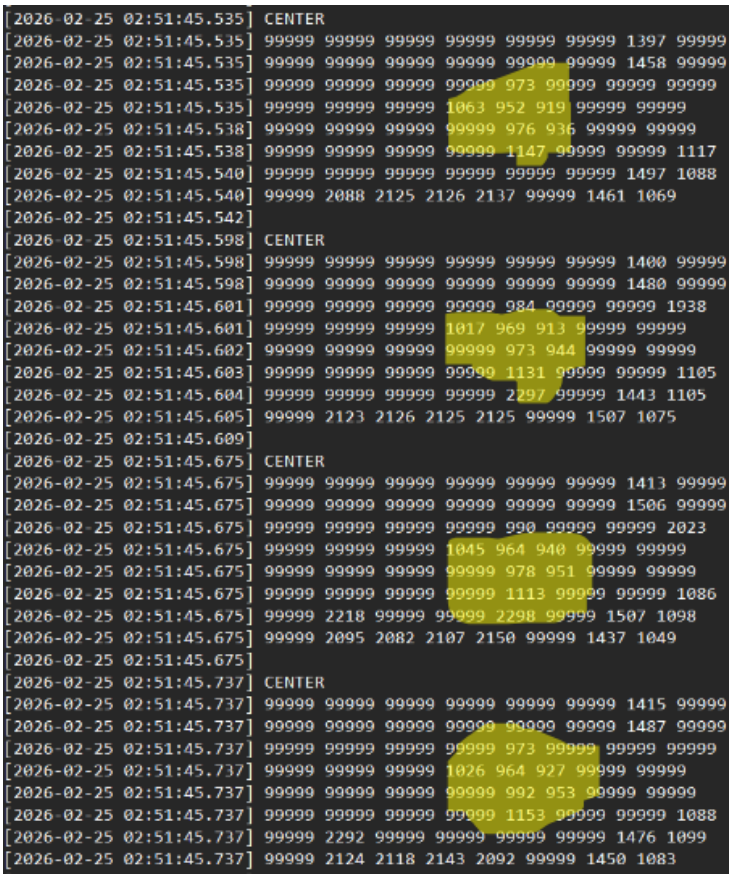


Figure 6.11: Data of the stopping scenario

6.4.4 Scenario S4: Fast Crossing

The fast crossing scenario under inclined viewing typically produces fewer valid frames than in ceiling mounting because the subject may traverse the sensing cone rapidly and only intersect a subset of zones. The representative trial shows that the ToF activation appears delayed, and the detected event time t_{ToF} shifts significantly relative to the ground-truth t_{GT}.

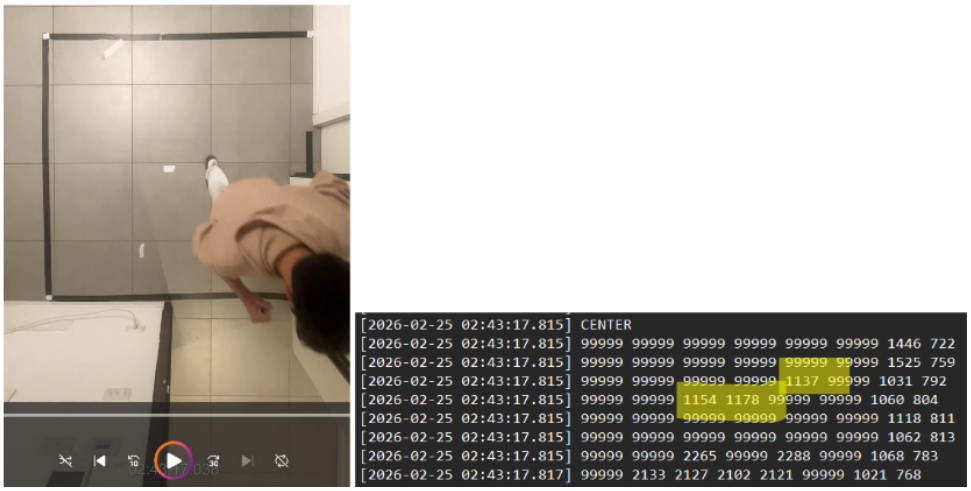


Figure 6.12: synchronized data and video

6.4.5 Scenario S5: Carried Object

In the inclined configuration, carried objects may generate strong planar returns that interact with the door plane, potentially producing ambiguous depth structures. The representative example shows the object surface is interpreted as part of the person cluster or as a separate activation region. This case is important for failure-mode analysis because real deployment environments frequently include bags, chairs, or household objects being moved through doors.

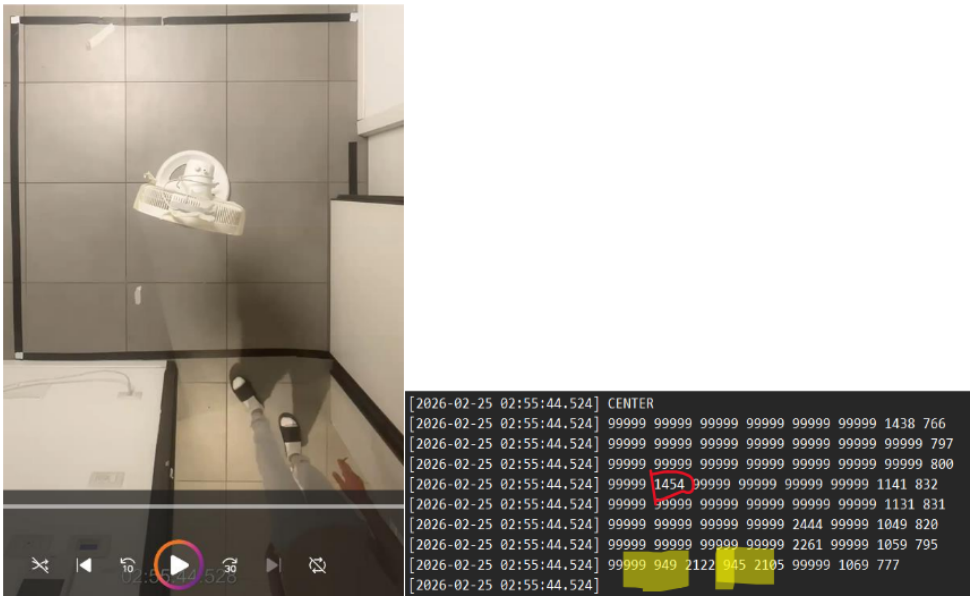


Figure 6.13: synchronized data and video for the object.

6.5 Configuration C: Side Sensor (Chair-Based Surrogate Wall)

The side configuration observes the ROI from a lateral viewpoint, which strongly emphasizes body profile and increases susceptibility to occlusion and geometric misalignment. Furthermore, because the sensor was mounted using a chair as a surrogate wall, small variations in stability and alignment may increase frame-to-frame variability compared to rigid ceiling/wall installations. The additional near-sensor camera used in this configuration supports interpretation by providing a similar perspective to the ToF sensor.

6.5.1 Scenario S1: Entry With Controlled Door Angles

In side-view entry, the camera can observe the subject immediately upon doorway appearance, while the ToF sensor may respond later because the subject must move into the sensor cone and present sufficient reflective area to multiple zones. This often produces a larger apparent delay Δt than in ceiling mounting. The synchronized trial describes how the subject appears gradually as a profile cluster, often first in a small subset of zones, then expanding as the subject moves deeper into the ROI. If the door angle modifies the visible corridor-to-room geometry, it may also alter which zones receive valid returns.

Figure 6.14: synchronized data and video for people entering.

6.5.2 Scenario S2: Exit With Controlled Door Angles

In exit motion, the subject may be partially occluded by the door or by self-occlusion (body segments aligned along the line of sight). The synchronized example shows the ToF measurement disappears earlier than expected, which would manifest as a premature reduction in valid zones. Such effects are relevant to missed detection analysis.

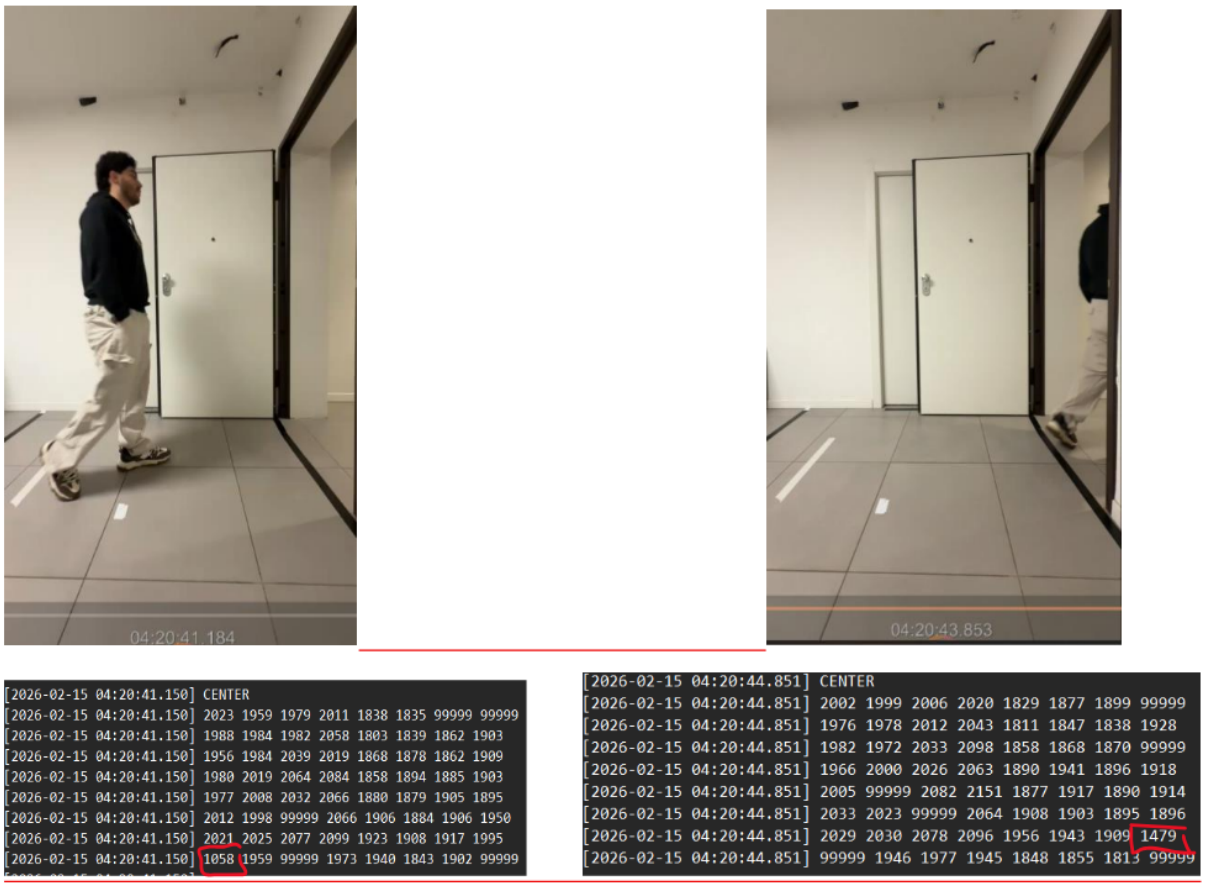
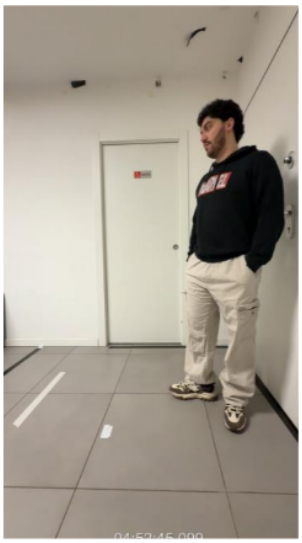


Figure 6.15: synchronized data and video for people exiting.

6.5.3 Scenario S3: Non-Crossing Motion Near Doorway

The side configuration is particularly sensitive to near-door motions that do not correspond to a true crossing event, since lateral movements can resemble an entry or exit when only partial profile returns are visible. The representative trial discusses the ToF patterns remain localized without completing a consistent crossing sequence.



```
[2026-02-15 04:52:45.040] CENTER
[2026-02-15 04:52:45.041] 1995 1993 1985 2030 2009 2043 99999 99999
[2026-02-15 04:52:45.041] 1980 1993 2008 2033 2055 1088 1150 99999
[2026-02-15 04:52:45.041] 1947 2008 2058 2060 2083 1050 1104 1697
[2026-02-15 04:52:45.041] 1978 2011 2039 2093 2091 1064 1125 1696
[2026-02-15 04:52:45.041] 1993 2048 2056 2121 2108 1162 1083 1713
[2026-02-15 04:52:45.041] 1989 2038 2082 2109 2172 2071 1084 1747
[2026-02-15 04:52:45.041] 2012 2045 2067 2106 2106 2071 1104 1726
[2026-02-15 04:52:45.041] 99999 99999 99999 99999 99999 1228 1138 99999
```

Figure 6.16: synchronized data and video for u turn.

For stopping:


```
[2026-02-15 04:58:03.631] CENTER
[2026-02-15 04:58:03.631] 2021 2011 984 2058 2039 2041 2043 99999
[2026-02-15 04:58:03.631] 1989 833 902 998 2089 2009 2038 99999
[2026-02-15 04:58:03.631] 1968 821 928 987 2070 2045 2065 1813
[2026-02-15 04:58:03.631] 2027 795 909 962 2073 2094 2087 1788
[2026-02-15 04:58:03.632] 1997 786 925 979 2103 2067 2120 1739
[2026-02-15 04:58:03.632] 2013 793 903 1032 2120 2113 2089 1707
[2026-02-15 04:58:03.632] 2002 917 905 1119 2113 2093 2111 1729
[2026-02-15 04:58:03.632] 99999 928 956 1104 99999 99999 99999 999
[2026-02-15 04:58:03.643]
[2026-02-15 04:58:03.693] CENTER
[2026-02-15 04:58:03.693] 2028 2004 998 2041 2054 2029 2077 99999
[2026-02-15 04:58:03.693] 1958 840 912 960 2069 2030 2026 99999
[2026-02-15 04:58:03.693] 1973 822 909 999 2078 2056 2063 1722
[2026-02-15 04:58:03.693] 2002 807 904 969 2120 2095 2054 1824
[2026-02-15 04:58:03.703] 2061 793 897 980 2086 2128 2066 1894
[2026-02-15 04:58:03.703] 2016 781 892 1018 2111 2143 2091 1696
[2026-02-15 04:58:03.703] 2053 902 903 1138 2154 2093 2120 99999
[2026-02-15 04:58:03.703] 99999 946 958 1166 99999 99999 1883 9999
[2026-02-15 04:58:03.703]
[2026-02-15 04:58:03.755] CENTER
[2026-02-15 04:58:03.755] 2001 2015 966 2026 2055 2045 2057 99999
[2026-02-15 04:58:03.756] 1974 837 893 1011 2010 2027 2091 99999
[2026-02-15 04:58:03.757] 1995 815 912 991 2099 2068 2072 1705
[2026-02-15 04:58:03.757] 2004 801 911 959 2066 2046 2093 1928
[2026-02-15 04:58:03.759] 2026 795 898 961 2103 2088 2080 1916
[2026-02-15 04:58:03.761] 2009 788 905 1032 2141 2085 2094 1811
[2026-02-15 04:58:03.761] 2023 888 902 1114 2117 2105 2113 1717
[2026-02-15 04:58:03.762] 99999 929 929 1195 99999 99999 1917 1605
[2026-02-15 04:58:03.766]
[2026-02-15 04:58:03.833] CENTER
[2026-02-15 04:58:03.833] 2028 1986 981 2042 2041 2026 2084 99999
[2026-02-15 04:58:03.833] 1992 841 902 1012 2036 2021 2042 99999
[2026-02-15 04:58:03.833] 1994 814 916 999 2071 2081 2042 1774
[2026-02-15 04:58:03.833] 1983 807 915 984 2112 2079 2044 1773
[2026-02-15 04:58:03.833] 1972 803 915 973 2122 2110 2097 1794
[2026-02-15 04:58:03.833] 2023 827 893 1016 2113 2096 2106 1823
[2026-02-15 04:58:03.833] 2044 872 910 1097 2131 2105 2097 1654
[2026-02-15 04:58:03.833] 1940 928 940 1127 99999 99999 1699 1601
```

Figure 6.17: synchronized data and video for stopping scenario.

6.5.4 Scenario S4: Fast Crossing

Under fast crossing, the lateral viewpoint combined with 15 Hz sampling increases the risk of missing the event entirely (few frames intersecting the subject) or detecting only a fragment of the trajectory. The representative example describes the number of frames in which the subject is visible.

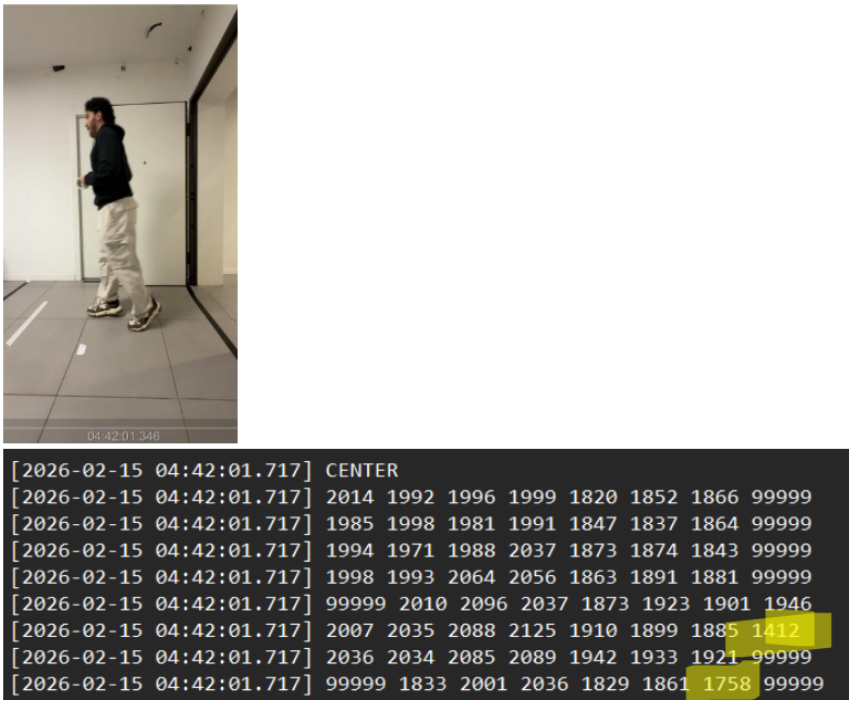


Figure 6.18: synchronized data and video for speed scenario.

6.5.5 Scenario S5: Carried Object

Carried objects are expected to have the strongest negative impact in side view, where the object may fully occlude the torso or create multiple depth surfaces aligned along the line of sight. The representative synchronized trial describes how the sensor returns become fragmented (many invalid zones), the object dominates the depth map, and the event timing shifts compared to the no-object case.

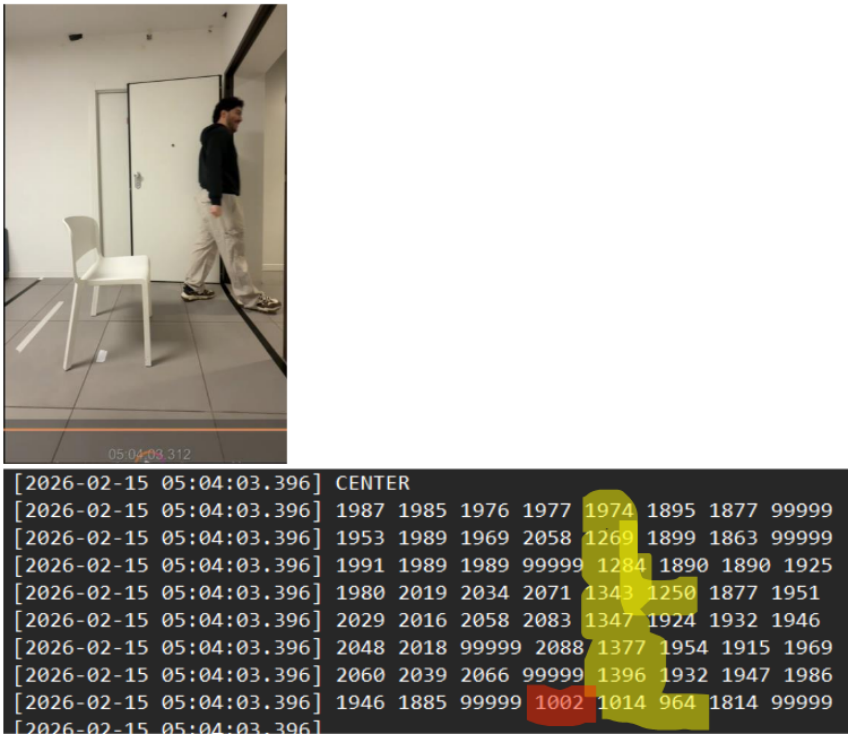


Figure 6.19: synchronized data and video for object scenario.

6.6 Summary of Qualitative Observations

Across the representative synchronized trials, three qualitative trends were observed.

1. **Ceiling mounting** provides the cleanest and most spatially stable ToF signature in the ROI, with compact clusters and limited door-plane interference.

2. **Inclined mounting** introduces stronger sensitivity to door geometry and planar surfaces, which can increase ambiguity and delay.

3. **Side mounting** exhibits the largest apparent delays and the greatest susceptibility to profile occlusion and carried-object disturbances, especially under fast motion at 15 Hz.

6.6.1 Observed Acquisition Limitations and Typical Failure Cases

Although the experimental protocol was designed to ensure repeatable trials across mounting configurations and scenarios, several practical limitations of the VL53L7CX multizone acquisition were observed during data collection. These limitations do not invalidate the dataset; rather, they represent realistic sensing constraints that must be documented explicitly, as they directly affect the interpretation of both the ToF frames and the synchronization against the video ground truth.

Low-reflectance clothing and reduced signal stability. A recurring observation was that subjects wearing dark clothing (e.g., black jackets or dark trousers) occasionally produced weaker and less stable depth returns in specific zones. In these cases, the ToF log

exhibited an increased number of invalid readings (encoded as 99999) and a reduced spatial extent of the detected cluster. This behavior is consistent with reduced infrared reflectance, which decreases the returned photon count and therefore degrades the signal-to-noise ratio. The effect was most evident in zones near the periphery of the field of view and under natural sunlight conditions, where ambient infrared can further reduce ranging confidence. As a consequence, event timing based on a first-detection criterion can shift later with respect to the video timestamp, not due to synchronization error but due to delayed physical observability in the ToF measurements.

Small objects and limited spatial observability. A second limitation concerns the detection of small objects. Due to the coarse spatial resolution of the 8×8 grid, each zone corresponds to a relatively large footprint on the monitored floor area. If an object occupies only a small fraction of a zone (for example, a narrow item, a small bag, or a partial limb segment), the depth change may remain below the effective detection threshold of the sensor or be filtered internally as an unreliable return. In such cases, the log may report no-target values (99999) even though the object is visible in the camera. This phenomenon is particularly relevant for edge-case scenarios involving carried items, where the object may be partially occluded or may not present a sufficiently large reflective surface toward the sensor. Therefore, the dataset includes instances where the ground-truth video confirms object presence while the ToF depth map shows only weak or intermittent activation.

High-speed motion and temporal undersampling. Finally, very fast walking introduced occasional partial or missed detections. The sensor operates at 15 Hz, corresponding to a sampling period of approximately $T_s \approx 66.7$ ms. At high walking speed (scenario S4, approximately 8 km/h), the subject can move a significant distance between consecutive ToF frames. As a result, the subject may traverse the region of interest in only a small number of frames, producing a short activation window and potentially discontinuous zone transitions. Depending on the mounting configuration and the effective overlap between the sensing cone and the ROI, this undersampling may lead to fragmented clusters or, in the worst case, a missed detection if the subject passes through zones that already contain invalid or noisy returns. This limitation is important for interpreting timing metrics and for understanding why high-speed scenarios can appear less consistent in the ToF stream compared to the video ground truth.

Overall, these failure cases highlight the sensitivity of low-cost multizone ToF sensing to target reflectivity, object scale relative to zone footprint, and temporal sampling limits. They provide an important motivation for robust post-processing strategies (e.g., confidence-based filtering, adaptive thresholds, and temporal consistency checks) when developing doorway monitoring algorithms based on the collected dataset.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

This thesis investigated privacy-preserving indoor people monitoring using a low-cost multizone infrared Time-of-Flight (ToF) sensor, with a particular focus on doorway environments where human motion and door dynamics create challenging sensing conditions. The work was intentionally refocused toward a dataset-centric contribution: rather than developing a single end-to-end counting algorithm, the main objective became the creation of a rigorous and well-documented experimental dataset that captures the behavior of the VL53L7CX under a structured set of scenarios and mounting configurations.

A complete experimental testbed was implemented using the VL53L7CX operating in 8×8 mode at 15 Hz, interfaced through an STM32 Nucleo-F401RE platform and logged via serial to a PC. Ground-truth video was recorded using an iPhone 13 in ultra-wide mode at 4K resolution and 60 fps. A millisecond timestamp overlay in the video, together with millisecond timestamps in the serial log, enabled post-processing alignment between the ToF stream and the visual ground truth. The dataset design was structured around three physically distinct mounting configurations (ceiling-mounted, inclined above-door, and side-mounted) and a set of five scenarios intended to stress both typical and adverse operating conditions, including controlled door angles, fast motion, non-crossing behavior, and object carrying. This structure enables controlled comparisons across geometry and scenario, supporting repeatable analysis.

The results chapter presented representative synchronized trials illustrating how ToF signatures evolve across scenarios and mounting configurations. The ceiling-mounted configuration provided the clearest depth signatures within the defined 2×2 m ROI, with comparatively stable coverage and reduced occlusion effects. In contrast, the inclined above-door configuration showed stronger sensitivity to planar surfaces and door geometry, which can introduce ambiguity in the ToF patterns depending on the door opening angle. The side-mounted configuration was the most sensitive to profile-based occlusion and exhibited more pronounced timing offsets relative to the camera ground truth, reflecting the fundamental mismatch between camera visibility and the effective sensing volume of the ToF sensor in lateral viewing. In addition, several practical limitations were observed and documented, including reduced stability for low-reflectance clothing, unreliable detection of small objects due to coarse zone footprints, and occasional undersampling under high-speed motion at 15 Hz.

Overall, the thesis delivers a complete experimental methodology, a structured dataset, and an analysis framework for interpreting VL53L7CX multizone ToF behavior in doorway

monitoring. The dataset and documentation are intended to support further research on disturbance rejection, robust event timing, and algorithmic development under realistic constraints.

7.2 Future Work

The dataset produced in this thesis provides a foundation that can be extended and exploited by future researchers, particularly data scientists and signal processing engineers, to develop and validate improved analysis techniques and counting algorithms. Several directions are recommended.

7.2.1 Algorithmic Refinement and Robust Door-Rejection Strategies

A natural continuation of this work is the implementation and evaluation of advanced door-rejection methods based on the recorded ToF measurements. Candidate approaches include spatial planar-consistency checks, temporal state machines with stronger transition constraints, and confidence-based filtering using per-zone validity patterns. Since the dataset includes multiple door opening angles and both entry/exit trials, it is suitable for systematic testing of how door motion and door geometry can be distinguished from human motion.

7.2.2 Machine Learning and Data-Driven Modeling

The dataset can be used by machine learning practitioners to train and evaluate data-driven models, such as temporal classifiers (e.g., LSTM/GRU architectures), convolutional models operating on the 8×8 depth maps, or hybrid approaches that combine engineered features (motion energy, validity ratios, cluster size) with learned representations. Because the dataset includes synchronized video, the ground truth can support event labeling and supervised learning. In addition, the dataset can enable research on knowledge distillation approaches, where a vision-based model provides pseudo-labels to train a ToF-only model without requiring manual annotation of every frame.

7.2.3 Improved Synchronization and Timing Accuracy

In this thesis, synchronization was performed manually using millisecond timestamps in the serial log and in the video overlay. Future experiments could significantly improve timing accuracy by implementing a hardware-based synchronization trigger, for example through a microcontroller-controlled LED visible to the camera and simultaneously logged in the ToF stream, or through a dedicated trigger signal if the camera supports external control. This would reduce uncertainty in event timing and enable more precise latency analysis across mounting configurations.

7.2.4 Extended Scenario Coverage and Dataset Expansion

The current dataset focuses on single-person trials with controlled disturbances. Future dataset expansions considers additional edge cases that are critical in real deployments, including two-person crossing, tailgating, groups entering together, children or shorter

subjects, and more systematic variations of clothing materials and carried object types. Further extensions could include additional environmental conditions such as different lighting regimes (night conditions, strong direct sunlight), different floor reflectance, and diverse doorway geometries. Such expansions would increase the generality of the dataset and enable stronger conclusions about sensor robustness.

7.2.5 Sensor Configuration Exploration

The experiments were conducted with fixed acquisition parameters (8×8 mode, 15 Hz) without tuning timing budget or integration settings. Future work may explore the effect of these parameters on detection robustness, noise stability, and motion capture performance, for example by comparing 15 Hz vs. higher frequencies or evaluating the trade-off between signal quality and temporal resolution. This would support principled guidelines for optimal configuration depending on the application constraints.

7.2.6 Practical Exploitation of the Dataset

Finally, the dataset and synchronized videos are explicitly intended to be reusable. A data scientist or researcher can use this dataset to develop alternative preprocessing pipelines, define improved feature sets, benchmark detection and timing metrics, or perform comparative experiments across mounting configurations. In this sense, the work is structured not only as a thesis outcome but also as a reusable experimental resource that enables further refinement, validation, and innovation in privacy-first indoor people monitoring using low-cost ToF sensing.

Bibliography

[1] European Commission, "Revised energy performance of buildings directive (eu/2024/1275)," European Commission, Energy, 2024, entered into force 28 May 2024; transposition deadline 29 May 2026.

[2] European Union, "Regulation (eu) 2016/679 (general data protection regulation)," 2016, official Journal of the European Union.

[3] STMicroelectronics, "Vl53l7cx: Time-of-flight (tof) 8x8 multizone ranging sensor with 90 degrees fov," ST product page, 2026.

[4] ——, "Stsw-img036: Vl53l7cx ultra lite driver (uld) api," ST embedded software page, 2026.

[5] ——, *UM3038: A guide to using the VL53L7CX Time-of-Flight multizone ranging sensor with the Ultra Lite Driver (ULD) API*, 2022.

[6] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.

[7] K. Bernardin and R. Stiefelhagen, "Evaluating multiple object tracking performance: The clear mot metrics," in *EURASIP Journal on Image and Video Processing*, 2008, defines MOTA and related tracking metrics.

[8] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," in *International Journal of Computer Vision*, 2000.

[9] "A zone-level occupancy counting system for commercial office spaces using low-resolution time-of-flight sensors," *Energy and Buildings*, vol. 252, p. 111390, 2021.

[10] G. Caroleo, A. Albini, and P. Maiolino, "On the characterisation of the time-of-flight vl53l5cx sensor by stmicroelectronics for indoor robotics applications," *Sensors*, vol. 26, no. 5, p. 1639, 2026.

[11] "Energy savings and ventilation performance from co2-based demand controlled ventilation: Simulation results from ashrae rp-1747," *Science and Technology for the Built Environment*, 2019.

[12] STMicroelectronics, "X-nucleo-53l7a1: Vl53l7cx-based expansion board for stm32 nucleo," https://www.st.com/en/ecosystems/x-nucleo-53l7a1.html, 2026, accessed: 2026-03-15.

[13] European Commission, "Energy performance of buildings directive (epbd)," https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficient-buildings/energy-performance-buildings-directive_en, 2024, accessed: 2026-03-15.

[14] European Union, “Regulation (eu) 2016/679 (general data protection regulation),” https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng, 2016, accessed: 2026-03-15.

[15] MathWorks, “Camera calibrator app (computer vision toolbox),” https://www.mathworks.com/help/vision/ref/cameracalibrator-app.html, 2026, accessed: 2026-03-15.

[16] ——, “Videoreader (matlab),” https://www.mathworks.com/help/matlab/ref/videoreader.html, 2026, accessed: 2026-03-15.

[17] TeraTerm Project, “Tera term open source project (official website),” https://teratermproject.github.io/, 2026, accessed: 2026-03-15.