SituGuard: LLM-based Fine-grained Smart Glass Privacy Control in Home Environments

Shuning Zhang
zsn23@mails.tsinghua.edu.cn
Tsinghua University
Beijing, China

Ying Ma
ying.ma1@student.unimelb.edu.au
School of Computing and Information Systems
University of Melbourne
Melbourne, Australia

Qucheng Zang
3190602072@caa.edu.cn
China Academy of Art
Hangzhou, China

Yongquan 'Owen' Hu
yongquan@ahlab.org
Augmented Human Lab
National University of Singapore
Singapore

Abstract

The rapid development of smart glasses, especially used in home context brings significant privacy risks. Current privacy controls are often device-specific, lacking the granularity and automation needed for the dynamic home environment. We introduce SituGuard, a novel privacy management framework designed to provide fine-grained, contextual and automatic control over smart glasses' visual data in home environments. SituGuard utilizes a multi-dimensional privacy categorization schema and an adaptive policy engine that leverages LLMs to dynamically assess privacy acceptability and enforce user-defined rules. It automatically detects sensitive objects locally, adopts privacy engine and finally automatically obfuscates the objects. An evaluation study (N=12) proved the feasibility of SituGuard, its automatic control and user feedback modules.

CCS Concepts

• **Security and privacy** → **Privacy protections**; • **Computing methodologies** → *Computer vision.*

Keywords

Visual privacy, Home environment, Large Language Model, Privacy

ACM Reference Format:
Shuning Zhang, Ying Ma, Qucheng Zang, and Yongquan 'Owen' Hu. 2025. SituGuard: LLM-based Fine-grained Smart Glass Privacy Control in Home Environments. In *Companion of the the 2025 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Companion '25), October 12–16, 2025, Espoo, Finland*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3714394.3750590

1 Introduction

The proliferation of smart glasses, exemplified by products like Ray-Ban Meta and Orion AR, marks their increasing integration into users' daily lives [3]. These devices hold significant potential for in-home applications, such as health self-monitoring [2, 14], intelligent household management [2], and social interaction [14].

However, the convenience of smart glasses is shadowed by unprecedented privacy risks associated with their always-on visual sensors. These sensors capture vast amounts of data [3, 8], rendering traditional permission models ineffective [4]. The context-dependent nature of privacy necessitates a dynamic trade-off between privacy and utility [18, 19], making static redaction techniques sub-optimal. Furthermore, the continuous nature of video streams presents a significant control challenges for users who can often only intervene asynchronously [15].

Prior work on dynamic privacy control has extended traditional permission models [4] with fine-grained approaches [1, 7, 12, 15]. Yet, object-based controls often require repetitive manual verification, which is difficult to scale in complex home environments with multiple salient objects [15]. Group-based controls can impose a significant cognitive load on users, hindering their ability to achieve satisfactory outcomes [1]. Consequently, both approaches fall short in minimizing user effort particularly for smart glasses.

The advent of Large Language Models (LLMs) offers a promising avenue for advanced image understanding and automated privacy control [18]. Despite this potential, there has been limited research on integrating LLMs to create automatic privacy-preserving techniques. This paper takes a first step through introducing SituGuard, a technique designed to automate privacy control for smart glasses in home environments through novel LLM-based techniques.

The core of SituGuard is an LLM-driven engine that generates contextual privacy protection rules. SituGuard first employs a localized object detection module to identify private objects within the visual streams. It then leverages Qwen3-8B[1] to reason about appropriate privacy protections based on the specific context, user preferences, and a multi-dimensional privacy classification of household objects. Qwen3-8B automatically generates rules such as redacting and retaining specific objects, which are then executed through targeted obfuscation. This significantly eases users' control efforts, especially in fine-grained privacy control. Through a user study (N=12), SituGuard outperformed alternative techniques without

[1] quantized to 4-bit.

UbiComp Companion '25, Espoo, Finland

ACM ISBN 979-8-4007-1477-1/2025/10
https://doi.org/10.1145/3714394.3750590

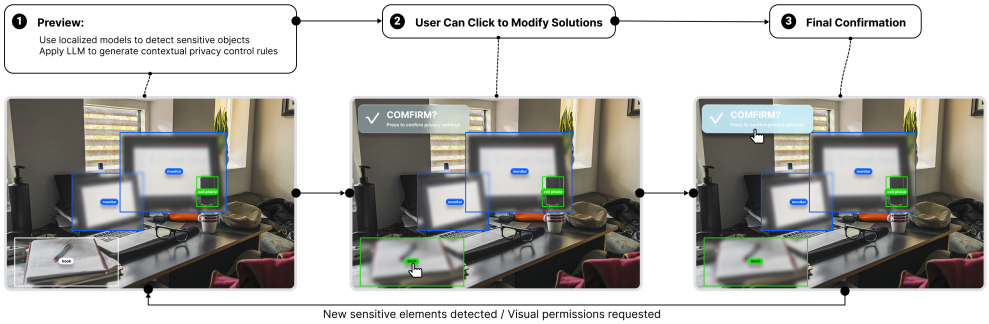

Figure 1: The interaction flow of SituGuard.

either LLM-based control or manual user adjustment. Participants reported high user satisfaction, a strong sense of perceived control and low cognitive load when using SituGuard.

2 Related Work

Our work builds upon three research areas: the privacy challenges of smart glasses, the evolution of permission models, and fine-grained visual privacy controls. **Existing protection mechanisms for smart glasses have been shown to be inadequate.** On-device indicators are often ineffective [3], and manual controls such as sliders [1] or automatic detection [15] impose a significant burden on users, particularly for managing fine-grained preferences. Our work addresses these limitations by leveraging LLMs to automatically reason about and protect sensitive visual information.

Traditional permission models, like install-time or run-time requests, are ill-suited for the continuous and passive data gathering inherent to smart glasses [4, 17]. While one-time permissions are an improvement, they still lack the granularity required for dynamic environments [17]. Context-aware models have been developed to reduce user cognitive load [5, 17], but they typically focus on device-level access rather than fine-grained control over specific objects [17]. SituGuard extends these dynamic approaches by enabling object-level privacy control based on the specific contexts of a user's home.

Fine-grained visual privacy control follows two main streams: modifying control interfaces [1, 12] and enabling object-level filtering [7, 9, 13, 15]. Interface-based approaches have moved beyond binary permissions to include graded controls, such as providing approximate location data [5] or using sliders to balance privacy and utility [1, 12]. However, these methods offer limited granularity. The second stream focuses on filtering specific objects, such as redacting bystanders [7]. Yet, these systems often require users to manually configure policies by selecting objects to block [13] or retain [9]. This manual configuration is cognitively demanding and impractical, as privacy management is often a secondary task [13]. While Aragorn automated object recognition, it still required manual selection, leaving unsolved the challenge of efficiently configuring policies for multiple objects [15]. SituGuard addresses this gap by using an LLM to automatically generate and adapt multi-object privacy rules, thereby minimizing user effort.

3 Design and Implementation

System Flow. As illustrated in Figure 2, SituGuard consists of four modules: sensitive object identification, privacy rule set construction, LLM-based policy generation and group-based erasure. Upon users' using smart glasses to complete tasks, their cameras would capture the environment. SituGuard would detect and selectively obfuscate the sensitive objects in the recordings with a frame-wise manner. The users could also express their preference to obfuscate or not obfuscate through pointing on the detected object boxes. The backend privacy setting around the object would correspondingly change. SituGuard at the backend would first track the image taken and apply a localized object detection model to identify sensitive objects. It pre-generates a dynamic classification engine to determine the privacy attributes and control rules for each object. It then uses LLM to reason given users' attributes and rules, resulting in contextually appropriate control settings. Finally, SituGuard operationalizes the settings through pixel-wise erasure.

System Design. We introduce the four modules in the system flow in this section separately. These interconnected modules are designed to achieve efficient localized obfuscation, while at the same time minimize users' cognitive load, without inducing high hardware cost that may hinder deployment.

Dynamic object identification. As the objects in the recorded view is rapidly changing, we identified the objects in a frame-wise manner, using localized fine-tuned detection models to balance accuracy, processing speed and hardware cost. This choice is because for home environments, the private objects are extensively benchmarked [16], and we could get reliable results and datasets on identifying objects. Different from Aragorn [15], we tracked multiple dynamic objects instead of only one object because, in most of the tasks the recorded view contain different saliency objects crucial for the task, or at least, different contextually sensitive objects. We selected Yolo-v10 for processing rather than Yolo-world (an open world model), other variations of yolos, or variants like detectrons for high detection accuracy and light-weighted implementation, which was verified in one of our pilot study. We chose the small sized model, Yolo-v10n, to make it compatible with the hardware specificity of mainstream smart glasses.

Dynamic rule construction. We combined different classifications to construct a multi-dimensional rule set, which were then input into LLMs as the context for classification. We first combined

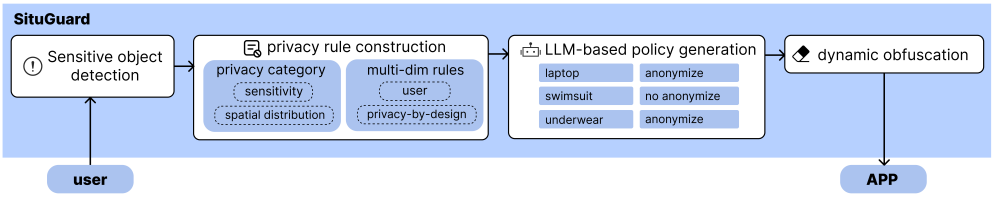

Figure 2: The system flow of SituGuard.

the results of previous works [16] to construct a comprehensive set of private objects. For classifications, we followed two primary aspects: sensitivity of the data and the spatial distribution of the data. For the sensitivity of the data, we followed previous work [16] and identified the spectrum of data sensitivity for each data type. This sensitivity guided the obfuscation, where more sensitive data corresponds to stricter rules in obfuscation, and less sensitive data corresponds to less strict rules in obfuscation. We classified the data into three categories: high sensitive, middle sensitive and low sensitive. We also grouped objects according to the places of their common appearance, such as living room objects (e.g., books, badges, TV screen), bedroom objects (e.g., underwear, swimming suit). Besides the classification, to adaptively guide the judgment, the LLM took two set of rules. The first set followed privacy by design principle where all private objects would be erased. This guarantees users' privacy to a maximum level. The second set is defined by users, where they control the setting through clicking on the interface to allow or disallow the collection of a specific data type under specific context. These multi-dimensional attributes and rules serve as the basis for SituGuard's policy decisions.

LLM-based assessment. We used Qwen3-8B for dynamic assessment based on the multi-dimensional attributes and rules constructed previously. Notably, we input LLMs with the context of the recordings, the task, the objects in the image, and multi-dimensional attributes and rules for decision. The decision would be effective in an asynchronous way until the next round of the assessment, as LLMs' latency are longer than the frame rates. We used LLMs for generating the data policy of each data type, as LLMs exhibited privacy-related contextual awareness in previous work [10].

Adaptive object obfuscation. After generating privacy rules, we used pixel-wise replacement to execute adaptive object obfuscation. SituGuard performed obfuscation according to LLMs' assessment. If the object was decided to be obfuscated, SituGuard would replace it with white pixels to avoid revealing users' private information.

System Implementation. We implemented the system on Hololens2 smart glasses and a Lenovo R9000P laptop with a RTX 3060 graphical card (6GB Memory). This setting mimics daily usage, which facilitates deployment. We used C script and python for implementation. The detection on Hololens2 used onnx, while the Qwen3-8B ran with ollama framework. We used socket for communications between the smart glasses and the laptop.

For the dynamic object identification, we finetuned Yolov10 on LVIS [6] and MSCOCO2017 [11] datasets, which are commonly used datasets featuring daily objects, and contained ample home environment objects. We used Yolov10n with a precision of 0.6704 to balance accuracy, model size and inference latency.

For the LLM-based assessment, the prompt adopts a role-play manner, letting LLMs to act as privacy evaluator to generate appropriate data control rules. We input Qwen3-8B (1) a context description containing the task users entered, and the application the users are currently used, (2) multi-dimensional classifications and rules, (3) the object name and positions detected. The output generates a binary collection setting (yes or no) for each object, guiding the adaptive object obfuscation. For the obfuscation, we adopted the unity engine for pixelwise camera replacement before sending them to the application. Notably, to ease calculation, we used the bounding box detected by the yolo model as the object's borders, rather than to segment the object.

4 User Study

4.1 Study Setup

Participants and Apparatus. This IRB-approved study recruited 12 participants (5M, 7F, aged 19-26, SD=2.3) through distributing questionnaires on online platforms. 3 participants are from IT-related occupations, 2 participants are from finance-related occupations, and others are from other disciplinarian. Each participant received 100 RMB as compensation.

Study Design. The study employed a within-subjects design, with the two factors being technique and scenario. We compared three ablation variations of SituGuard:

- **SituGuard** automatically provided a recommended privacy configuration. Users can freely accept or manually adjust this recommendation by selecting or deselecting specific items to be anonymized.
- **Manual-only control** removed the context-aware recommendation feature. For each scenario, it has with no default obfuscation. Users were required to configure their privacy control settings from scratch by manually selecting all the items they wished to obscure. This condition helped evaluate automated suggestions.
- **Recommendation-only control** removed the manual adjustment feature, displaying its recommended rules and obfuscation for the scenario. Users could only accept or reject the entire recommendation, and could not make adjustments. This condition helped evaluate user agency and perceived control of SituGuard.

For scenario, we selected four representative scenarios according to the prior work on home environments: daily health monitoring and behavior management [2, 14], household and lifestyle management [2], social interaction and contextual awareness [14], multimodal learning and work assistance [14].

Procedure. After providing informed consent and completing a demographic questionnaire, participants were introduced to the study's premise and the three techniques they would interact with.

For each trial, participants were presented with an interactive mockup of one of the three system conditions. They were asked to use the given interface to set their preferred privacy level for that scenario. After completing the task, they completed questionnaires around the condition's usability, their perceived sense of control, cognitive load, and overall satisfaction. This process was repeated until all three techniques are evaluated across all scenarios.

Analysis Methods. We used Friedman non-parametric test to analyze the subjective ratings, comparing the usability scores (e.g., SUS, out of 100), cognitive load ratings (e.g., NASA-TLX, out of 21), and Likert scale (out of 7) for perceived control and satisfaction across the three techniques. Post-hoc pairwise comparisons with Bonferroni correction were used to examine specific differences between conditions. For privacy protection effectiveness, the ground truth count of sensitive items were pre-identified by one author manually. This author discussed with other authors intermittently to ensure the reliability of the results.

4.2 Results

4.2.1 Privacy Protection Effectiveness. To assess the effectiveness of each technique, we measured the number of sensitive items users chose to anonymize compared to a "ground truth" count of sensitive items pre-identified by the researchers for each scenario.

Our findings indicate that SituGuard enabled users to configure privacy settings that most closely matched the ground truth. Participants using the SituGuard consistently anonymized an appropriate number of sensitive objects across all scenarios. In the *Manual-only* condition, participants frequently overlooked items, resulting in under-protection. On average, participants in this condition missed 30% of the sensitive items in the *multimodal learning and work assistance* scenario, the most complex environment. Conversely, the *Recommendation-only* condition often led to a binary choice, either accept a potentially over-protective configuration or reject it entirely, leaving sensitive information exposed. For instance, in the *household and lifestyle management* scenario, 5/12 users rejected the recommendation because it obscured an item they needed for the task, thereby failing to anonymize any items at all.

4.2.2 Subjective Ratings. Our analysis found a statistically significant main effect of the technique across all four metrics: *usability* (System Usability Scale, SUS), *cognitive load* (NASA-TLX), *perceived control*, and *overall satisfaction* ($p < .05$ for all).

SituGuard received the highest scores for both usability (M=88.5, SD=5.2) and overall satisfaction (M=6.6, SD=0.7). It was rated significantly more usable and satisfying than both the Manual-only (M=65.3 for usability, M=4.5 for satisfaction), and Recommendation-only (M=72.1 for usability, M=5.0 for satisfaction) conditions. For cognitive load, the Manual-only condition imposed the highest cognitive load (M=16.8, SD=3.3), which was significantly higher than both SituGuard (M=7.5, SD=2.1) and Recommendation-only (M=6.2, SD=1.9). Participants in the manual condition reported feeling overwhelmed by the need to "scan everything from scratch". For perceived control, the Recommendation-only condition scored significantly lower on perceived control (M=4.1, SD=1.1) compared to SituGuard (M=6.5, SD=0.6) and Manual-only (M=6.2, SD=0.8).

Acknowledgments

This work was supported by the Natural Science Foundation of China under Grant No. 62472243 and 62132010.

References

[1] Melvin Abraham, Mark Mcgill, and Mohamed Khamis. 2024. What You Experience is What We Collect: User Experience Based Fine-Grained Permissions for Everyday Augmented Reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–24.

[2] Riku Arakawa, Hiromu Yakura, and Mayank Goel. 2024. PrISM-Observer: Intervention agent to help users perform everyday procedures sensed using a smartwatch. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–16.

[3] Divyanshu Bhardwaj, Alexander Ponticello, Shreya Tomar, Adrian Dabrowski, and Katharina Krombholz. 2024. In Focus, Out of Privacy: The Wearer's Perspective on the Privacy Dilemma of Camera Glasses. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–18.

[4] Adrienne Porter Felt, Serge Egelman, Matthew Finifter, Devdatta Akhawe, and David Wagner. 2012. How to ask for permission. *HotSec* (2012).

[5] Huiqing Fu and Janne Lindqvist. 2014. General area or approximate location? How people understand location permissions. In *Proceedings of the 13th Workshop on Privacy in the Electronic Society*. ACM, 117–120. doi:10.1145/2661435.2661441

[6] Agrim Gupta, Piotr Dollar, and Ross Girshick. 2019. Lvis: A dataset for large vocabulary instance segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 5356–5364.

[7] Rakibul Hasan, David Crandall, Mario Fritz, and Apu Kapadia. 2020. Automatically detecting bystanders in photos to reduce privacy risks. In *2020 IEEE Symposium on Security and Privacy (SP)*. IEEE, 318–335.

[8] Yongquan 'Owen' Hu, Jingyu Tang, Xinya Gong, Zhongyi Zhou, Shuning Zhang, Don Samitha Elvitigala, Florian 'Floyd' Mueller, Wen Hu, and Aaron J Quigley. 2025. Vision-based multimodal interfaces: A survey and taxonomy for enhanced context-aware system design. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–31.

[9] Suman Jana, David Molnar, Alexander Moshchuk, Alan Dunn, Benjamin Livshits, Helen J Wang, and Eyal Ofek. 2013. Enabling {Fine-Grained} permissions for augmented reality applications with recognizers. In *22nd USENIX Security Symposium (USENIX Security 13)*. 415–430.

[10] Siwon Kim, Sangdoo Yun, Hwaran Lee, Martin Gubri, Sungroh Yoon, and Seong Joon Oh. 2023. Propile: Probing privacy leakage in large language models. *Advances in Neural Information Processing Systems* 36 (2023), 20750–20762.

[11] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, and Piotr Dollár. 2015. Microsoft COCO: Common Objects in Context. arXiv:1405.0312

[12] Vivek C Nair, Gonzalo Munilla-Garrido, and Dawn Song. 2023. Going incognito in the metaverse: Achieving theoretically optimal privacy-usability tradeoffs in VR. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–16.

[13] Nisarg Raval, Animesh Srivastava, Ali Razeen, Kiron Lebeck, Ashwin Machanavajjhala, and Lanodn P Cox. 2016. What you mark is what apps see. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*. 249–261.

[14] Julian Steil, Marion Koelle, Wilko Heuten, Susanne Boll, and Andreas Bulling. 2019. Privaceye: privacy-preserving head-mounted eye tracking using egocentric scene image and eye movement features. In *Proceedings of the 11th ACM symposium on eye tracking research & applications*. 1–10.

[15] Hari Venugopalan, Zainul Abi Din, Trevor Carpenter, Jason Lowe-Power, Samuel T King, and Zubair Shafiq. 2024. Aragorn: A privacy-enhancing system for mobile cameras. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 4 (2024), 1–31.

[16] Yuntao Wang, Zirui Cheng, Xin Yi, Yan Kong, Xueyang Wang, Xuhai Xu, Yukang Yan, Chun Yu, Shwetak Patel, and Yuanchun Shi. 2023. Modeling the trade-off of privacy preservation and activity recognition on low-resolution images. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*.

[17] Primal Wijesekera, Joel Reardon, Irwin Reyes, Lynn Tsai, Jung-Wei Chen, Nathan Good, David Wagner, Konstantin Beznosov, and Serge Egelman. 2018. Contextualizing privacy decisions for better prediction (and protection). In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

[18] Shuning Zhang, Lyumanshan Ye, Xin Yi, Jingyu Tang, Bo Shui, Haobin Xing, Pengfei Liu, and Hewu Li. 2024. " Ghost of the past": identifying and resolving privacy leakage from LLM's memory through proactive user interaction. *arXiv preprint arXiv:2410.14931* (2024).

[19] Shuning Zhang, Xin Yi, Haobin Xing, Lyumanshan Ye, Yongquan Hu, and Hewu Li. 2024. Adanonymizer: Interactively Navigating and Balancing the Duality of Privacy and Output Performance in Human-LLM Interaction. *arXiv preprint arXiv:2410.15044* (2024).