Computer Vision for Indoor Multimedia Geolocation

Kanwal Aftab and Mark Scanlon

School of Computer Science, University College Dublin, Ireland.
kanwal.aftab@ucdconnect.ie, mark.scanlon@ucd.ie

MOTIVATION

THE WHY: Indoor multimedia geolocation is becoming an increasingly essential component of the digital forensics toolkit.

Assisting Law Enforcement: Helps combat crimes such as human trafficking and child exploitation [1].

Hotel Identification: Victims are often photographed in hotel rooms; timely identifying the location can enable rescues and potentially save lives.

Misinformation Detection: During emergencies and natural disasters, e.g., earthquakes and floods, AI-driven geolocation can help counter social media disinformation, allowing responders to provide timely, targeted aid.
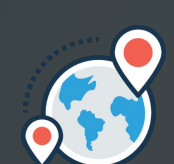
Missing Persons Cases: Pinpoints the last known location of missing persons by analysing digital images and videos for clues.
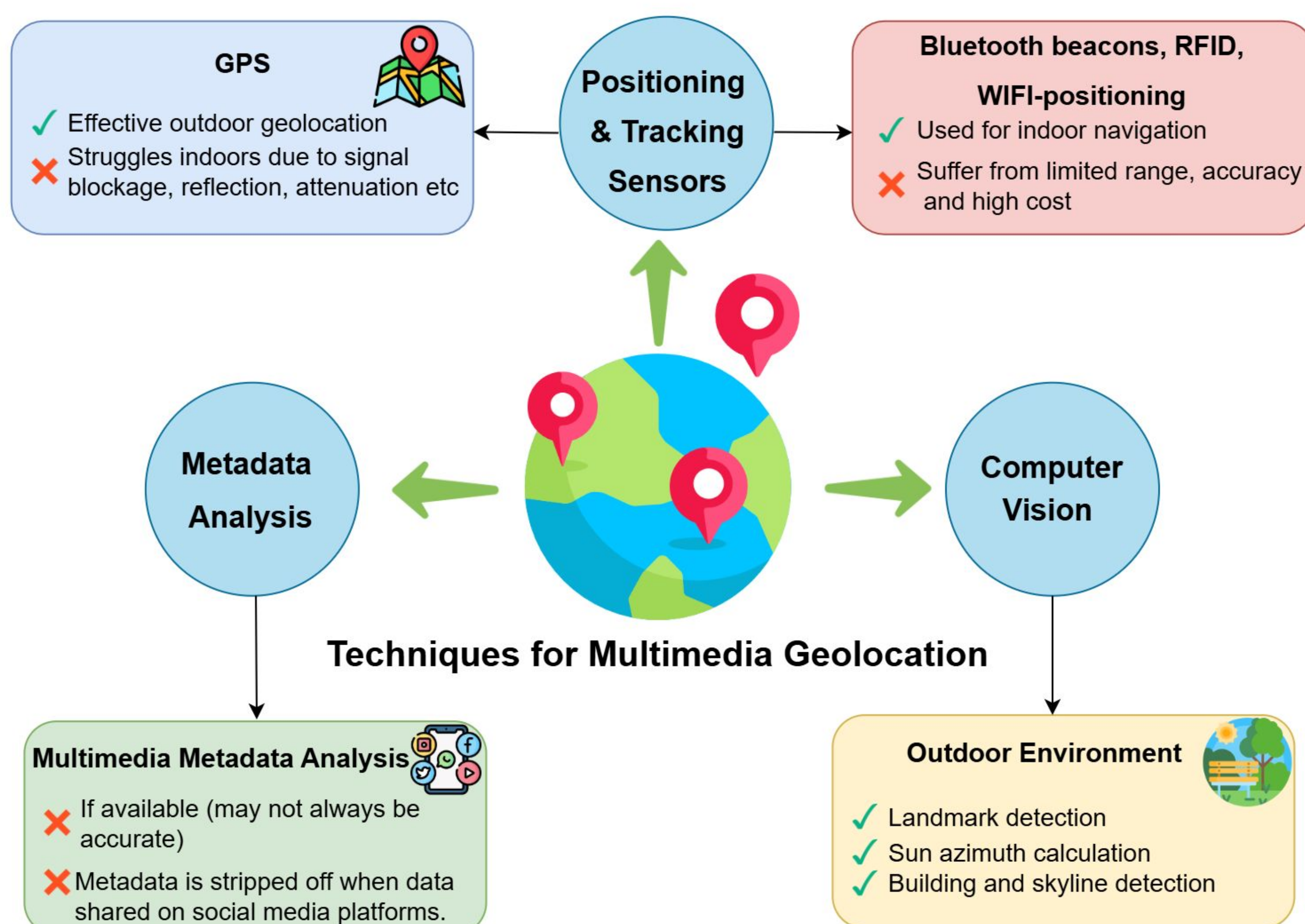

Figure 1: Comparison of multimedia geolocation techniques with strengths and limitations.

OUR RESEARCH FOCUS

THE WHAT: While multimedia geolocation using computer vision has gained significant attention outdoors, its potential indoors remain unexplored. Our research focuses on using computer vision techniques for indoor multimedia geolocation to reduce the search space for hotel room identification.

CHALLENGES OF INDOOR GEOLOCATION

1. Lack of Unique Features – Unlike outdoor spaces with landmarks and road signs, hotel rooms have similar layouts, furniture, and decor, making identification difficult.
2. Frequent Changes & Varying Angles – Renovations and different photo angles make old images unreliable.
3. Visual Ambiguity – White beds, wooden furniture, and plain surfaces provide minimal distinguishing details, making it harder to differentiate rooms.
4. Lighting Variability – Indoor lighting shifts affect image consistency and recognition.
5. No Reliable GPS – Weak or no signals make GPS ineffective indoors.
6. Limited Data – Computer vision relies on vast datasets, but collecting data, especially in sensitive areas like human trafficking cases, is challenging.

HOTELS-50K DATASET [2]

- Scope: Over 1 million images from 50,000 hotels worldwide.
- Metadata includes hotel name, location, and chain.

Figure 2: Data sources include images from travel websites and the TraffickCam mobile app.

METHODOLOGY

THE HOW: Different computer vision techniques can be used to address the problem of multimedia geolocation.

Key approaches include:

- CBIR: Searches and retrieves images based on visual content such as colour, texture, and shape.
- Image Embedding & Distance Matrices: Measures similarity between images.
- Neural Networks based Feature Extraction: Used for scene classification and object-based similarity.
- Semantic Scene Understanding: Goes beyond appearance to analyse relationships between elements.

Methods for Evaluating Results:

- Computer Vision Evaluation: Use a mean average precision (mAP) of 0.5 for automated assessment.
- Human Expert Evaluation: When step one generates 5 to 20 images from 50,000, a human domain expert evaluates the results.

Figure 3: Given the hotel room image as input, the top colours are extracted by Herrmann et al.[4].

FUTURE DIRECTIONS

To enhance accuracy and robustness, future research directions can focus on: combining HSV colour model with shape or texture, brand identification (i.e shopping bag, beverage cans), language detection (OCR), magazine analysis, unique pattern recognition (curtain, carpets), electric plug socket classification, close look at furniture.

REFERENCES

1. Bamigbade, O., Sheppard, J., & Scanlon, M. (2024). Computer Vision for Multimedia Geolocation in Human Trafficking Investigation: A Systematic Literature Review. *arXiv preprint arXiv:2402.15448*
2. Stylianou, A., Xuan, H., Shende, M., Brandt, J., Souvenir, R., & Pless, R. (2019, July). Hotels-50k: A global hotel recognition dataset. In *Proceedings of the AAAI Conference on Artificial Intelligence* Vol. 33, No. 01, pp. 726-733.
3. Herrmann, J., Bamigbade, O., Sheppard, J., & Scanlon, M. (2024, November). Perceptual Colour-based Geolocation of Human Trafficking Images for Digital Forensic Investigation. In *2024 Cyber Research Conference-Ireland (Cyber-RCI),* IEEE.