Progress Journal for Meeting and Learning – Group 10

Group Members: Allan Cheboiwo #73661290, Tarek Alkabbani #84930239, Haoyu Wang #42343871, Vanessa Laurel Hariyanto #72484546

1. **Date:** January 14, 2025

Participant: Allan Cheboiwo, Tarek Alkabbani, Haoyu Wang, and Vanessa Laurel Hariyanto

Meeting content:

This is the first meeting of our team, and the main objective of this meeting is to select the project topic and develop a work plan. After full discussion, the group voted to identify Multi-Object Tracking (MOT) as the research topic of this project. After the subject was determined, we divided the literature research tasks in the early stage of the project. All members will individually research and summarize the following over the next two weeks:

• The development and research status of multi-object tracking;

• Typical application scenarios and actual requirements of MOT;

• Current mainstream technology paths, including deep learning methods and traditional methods;

• Challenges and technical bottlenecks of the current approach.

In addition, in order to safeguard the record of information during the project, we have designated Haoyu to maintain the Progress Journals required for this project. The log includes two parts, Topic Presentation and Project, which record meeting dates, stage progress, task assignment, technical decision and learning reflection, etc., in order to support subsequent report writing and project evaluation.

Reflection: Through the group discussion, we realized that multi-object tracking is a hot research direction at present, which has a wide range of applications in many fields, and involves complex and diverse technologies, which requires systematic understanding and division of labor of each of us.

Analyzation: We analyze the current research trends and challenges of multi-object tracking, and believe that to fully understand the field, we need to start from three aspects: application scenarios, mainstream methods (such as deep learning and traditional methods), and existing problems.

Conclusions: The team defined the research topic and made the preliminary work plan, which laid a good foundation for the following investigation and project progress.

2. **Date:** February 2, 2025

   Participant: Allan Cheboiwo, Tarek Alkabbani, Haoyu Wang, and Vanessa Laurel Hariyanto

   Meeting content:

   The main objective of this meeting was to unify the basic understanding of the team members in the field of multi-object tracking and to initially explore the possible application directions of the project.

   At the beginning of the meeting, members shared relevant background information that they had recently reviewed to enhance their overall understanding of the current state of MOT research. Through information exchange, each member has a preliminary grasp of the basic concepts, mainstream methods and key technologies in the field.

   In the sharing session:

   • Haoyu has provided and explained several review papers on multi-object tracking [1][2], detailing the basic definition of MOT problems, common application scenarios (such as video surveillance, autonomous driving, etc.), and the classification of traditional methods and deep learning methods;

• Tarek shared a MOT practice from the iWildcam 2021-FGVC8 project [3] and showed a summary page on multi-object tracking models and their corresponding Papers [4], a resource that helps us understand the features and scenarios of current mainstream models.

Subsequently, the group held in-depth discussions around the potential application areas of MOT technology. Considering that multi-object tracking has a high degree of task difference in different industries, such as traffic monitoring, pedestrian tracking, industrial detection, etc., we decided that each member would select and investigate some feasible application directions in the following time, and report and evaluate them in the next meeting, so as to determine the final target application scenarios of this project.

Reflection: By sharing information and explaining to each other, each of us has a preliminary and balanced understanding of the basic concept and research status of multi-object tracking. This also lays a good foundation for the progress of the next project.

Analyzation: We analyze the application requirements of multi-object tracking in different scenarios, and realize that the choice of model is highly dependent on the specific application background, so further research needs to focus on specific areas.

Conclusions: We decided that each member would investigate some specific application areas and report the research results in the next meeting to determine the final application direction of the project.

3. **Date:** February 15, 2025

Participant: Allan Cheboiwo, Tarek Alkabbani, and Haoyu Wang

Meeting content:

In this meeting, we reported and discussed the application scenarios of multi-object tracking in our respective investigations, and finally determined the specific research direction of the project through voting: pedestrian tracking using Kalman Filter.

We chose this direction for the following reasons:

• Pedestrian tracking has a wide range of practical applications, such as video surveillance, security protection and crowd statistics;

• Classroom environment suitable for system demonstration;

• There is a wealth of research and publicly available data sets in this field that are easy to implement and compare.

Members then shared methods and resources related to pedestrian tracking using Kalman Filter:

• Allan introduced the paper Simple Online and Realtime Tracking (SORT) [5], which combines Kalman Filter with Hungarian Algorithm, It is one of the traditional pedestrian tracking methods widely used at present.

• Haoyu shared the paper "PKF: Probabilistic Data Association Kalman Filter for Multi-Object Tracking" [6] and proposed to improve this work. After discussion, we believe that its improvement space is limited, and considering the project time limit, it is difficult to be implemented in the short term.

Tarek shared a Kinect V2-based 3D pedestrian dataset [7] and recommended to integrate depth information into existing 2D detection frameworks to mitigate occlusion's impact on tracking accuracy.

In the end, we decided to adopt Tarek's suggestion and build a pedestrian tracking system based on Kalman Filter and Hungarian Algorithm, incorporating depth information to improve robustness of tracking.

In addition, based on our existing research content and research direction, the team divided the work of writing the Part I report:

• Haoyu is responsible for the Introduction section;

• Allan is responsible for the Related Works segment;

• Tarek is responsible for the Proposed Method section;

• Vanessa is responsible for the Abstract and Conclusion and Future Work sections.

Reflection: Through the exchange of research results and method discussion in this meeting, our team reached a unified understanding on the direction of pedestrian tracking using Kalman Filter, and realized that choosing a direction with strong realizability, rich technical resources and suitable for demonstration among many application scenarios is the key to efficiently complete the project. This meeting not only clarified the research direction of the project, but also enhanced the awareness of technical exchange and collaboration among team members. Our active participation in literature analysis and program selection has laid a good foundation for the subsequent efficient promotion of the project.

Analyzation: We comprehensively analyze the advantages and disadvantages of different schemes from four dimensions: application value, realizability, richness of existing achievements and demonstration conditions. Although complex models such as PKF have a certain frontier, the realization threshold is high and the risk is large. In contrast, the SORT scheme that integrates deep information has more advantages in performance and balance, which meets the actual needs of the current stage.

Conclusions: The team finally decided on "Pedestrian tracking system based on depth information fusion of Kalman Filter and Hungarian Algorithm" as the project theme. In addition, we have completed the writing division of the Part I report, and will carry out specific code development and experimental work in the next stage.

4. **Date:** February 21, 2025

Participant: Allan Cheboiwo, Tarek Alkabbani, Haoyu Wang, and Vanessa Laurel

Hariyanto

Meeting content:

The main objective of this meeting is to integrate the content of Part I report, and further clarify and refine the system architecture design of this project on this basis.

First, members upload the Part I report they are responsible for into the report file. Based on the content of the integrated report, the team discussed the core technical modules of the system implementation and formally determined the final system structure. As shown in Figure 1, our pedestrian tracking system is preliminarily determined to consist of the following modules:

1. 2D Image Data Input: The system input is RGB image data of continuous frames;

2. Object Detection: Detect pedestrian boundary boxes in images by deep learning method;

3. Depth Estimation: Deep learning models are used to obtain the depth information of pedestrian, which is used to enhance the object description;

4. Object Descriptors: Extract the space and appearance features by combining the boundary box area and depth information;

5. Hungarian Algorithm: Using Hungarian algorithm to match targets based on prediction and detection results;

6. Trajectory prediction module (Kalman Filter with Depth): Kalman filter integrating depth information is used for state prediction and update;

7. Feature Matching of unmatched objects (SIFT Matching) (Optional): Image local features (such as SIFT) are used for further matching of unmatched objects;

8. Bounding Box Labels: Update matching labels, assign new object ids to unmatched targets, and maintain the trajectory continuity of targets.

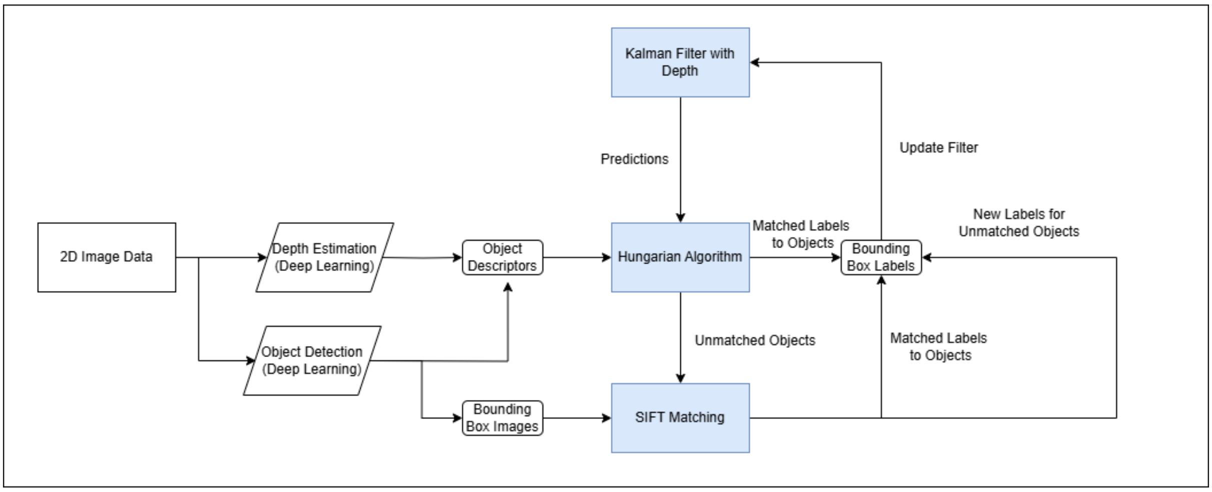

Figure 1. System Design.

Reflection: Through the content integration and system architecture review, we deeply realized that a complete pedestrian tracking system involves a number of key modules from data preprocessing, detection, depth estimation, state prediction and data association. Through collaborative reporting and graphical construction, members have enhanced their understanding of the overall process of the system and also improved their understanding of the interaction logic between modules.

Analyzation: During the system architecture design process, we fully considered the common challenges in pedestrian tracking, such as occlusion, target confusion, and scale variation. In order to improve the robustness and matching accuracy of the system, we adopt a multi-level feature fusion strategy (using depth information and introducing SIFT feature matching mechanism). The modular design of the system is also convenient for subsequent parallel development and independent debugging.

Conclusions: In this meeting, our team completed the integration of Part I report content, and based on the discussion results, officially determined the overall structure and data flow of the project system. In the future, tasks will be divided according to the system modules, and the implementation and debugging work will be gradually carried out.

5. **Date:** February 23, 2025

Participant: Allan Cheboiwo, Tarek Alkabbani, Haoyu Wang, and Vanessa Laurel Hariyanto

Meeting content:

The main task of this meeting is to conduct the final review of the Part I project report, and on this basis, clarify the specific work arrangement and task division after the project enters the code implementation stage.

In this meeting, all members jointly reviewed the report document paragraph by paragraph, focusing on the consistency of content, formatting and citation integrity. After the report is reviewed, the team submits the final version on time after the meeting.

After the report was submitted, the team officially entered the code implementation phase of the project. In order to ensure the efficient implementation of the task, we have carried out the following division of work:

• Allan and Haoyu are responsible for implementing the SORT module, including the prediction and update flow of Kalman Filter and the data association logic of the Hungarian algorithm. At the same time, they will jointly investigate and screen deep learning models (such as YOLO + Monodepth) that can be used for pedestrian border detection and depth information acquisition.

• Tarek and Vanessa investigate and screen suitable data sets, focusing on open data sources, ensuring that the data sets meet the typical challenges of multi-target scenes, occlusion, scale variation, etc., to test system performance and compare methods.

Reflection:

After this meeting, the project officially entered the code development stage from the literature research and system design stage. Through collective review of the report, members' understanding of the system structure and technical details was further consolidated.

Analyzation: We sorted out and analyzed the main modules of the system development stage, and divided the task into two main technical lines: basic algorithm realization and data preparation. This division of labor reflects the overall control of the project development process and the awareness of module collaboration.

Conclusions: The team has completed the final review and submission of the Part I report and officially entered the project implementation phase. The division of work of each member has been clarified, and the development and testing tasks will be carried out in parallel by module, and the progress and preliminary results will be reported in the next meeting.

6. **Date:** March 4, 2025

Participant: Allan Cheboiwo, Tarek Alkabbani, Haoyu Wang, and Vanessa Laurel Hariyanto

Meeting content:

During the meeting, members reported on the recent work progress and shared the resources and implementation related to the project:

• Allan shared the implementation code of SORT algorithm [8] and proposed to use YOLOv8 [9] to detect pedestrian boundary boxes and MiDaS [10] to obtain corresponding depth information.

• Haoyu shared another implementation for the kalman filter based multiple object Tracker [11], which can be used as a supplement or reference to the existing tracking process.

• Tarek and Vanessa recommended a database resource website [12] and pointed out that it contained datasets suitable for testing in this project and applicable to our current assessment needs.

After evaluating all the shared content and discussing the practical feasibility, our team decided to adopt the “SORT + YOLOv8 + MiDaS” implementation code provided by Allan as the main framework of the current stage, and use the data set recommended by Tarek as the test platform to carry out follow-up experiments.

The next phase was defined: Allan and Tarek were responsible for integrating in-depth information into the existing system framework and conducting preliminary testing and analysis of the system's performance. The results will be discussed at the next meeting.

Reflection: Through this meeting, our group has made substantial progress in resource integration and technical route confirmation. Each member not only completed the content of the preliminary task assignment, but also actively explored the realizable code framework and available data resources.

Analyzation: The meeting assessed available resources and established a unified starting point for development on this basis. Through comparative analysis of different implementation versions, we believe that the code structure based on YOLOv8 and MiDaS provided by Allan is clearer and easier to expand the subsequent modules. In terms of data set selection, the data source provided by Tarek can meet our experimental requirements. This stage also validates the rationality of the previous system design and makes the connection between theory and practice closer.

Conclusions: The meeting decided to use Allan's implementation code as the basis for system development, and selected the data set provided by Tarek as the test platform. Allan and Tarek will then be responsible for integrating the depth information into the SORT framework and evaluating its effectiveness in pedestrian tracking.

7. **Date:** March 25, 2025

Participant: Allan Cheboiwo, Tarek Alkabbani, Haoyu Wang, and Vanessa Laurel Hariyanto

Meeting content:

In this meeting, Allan and Tarek successfully integrated the depth information into the existing system and presented the results in the meeting (see Figure 2). The demonstration shows that the system works when processing the pedestrian tracking task with integrated depth features, and verifies the correctness of the code and the cooperative logic between modules.

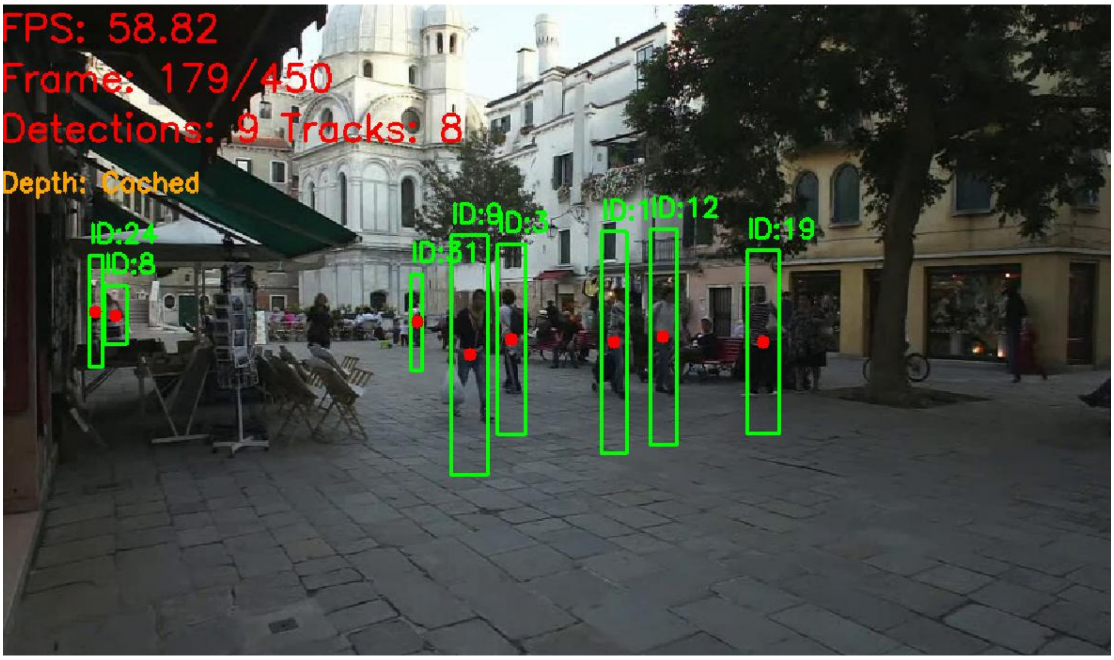

Figure 2. Pedestrian Tracking System Including Depth Information.

We then evaluated the system's performance on selected test datasets, focusing on the following metrics: HOTA, MOTA, IDF1 and IDs. The test results showed that the system performed poorly on these metrics. In response to this problem, the team members analyzed the possible reasons, including detector false detection, depth estimation error caused interference to the tracking module, and the data set itself scene complexity and the training model is not well matched. In order to further verify the robustness and generalization of the system, the team decided to continue to collect and test more representative data sets in the next few days to compare performance changes under different data conditions and provide a basis for subsequent optimization.

On the basis of the completion of the system, our team has divided the task of writing the Part II report and Topic Presentation, and the specific arrangement is as follows:

Part II Report:

• Haoyu: Introduction and Literature Review

• Allan and Tarek: System Design and Results

• Vanessa: Conclusion and Future Work

Topic Presentation:

Each member will be responsible for the presentation and slides preparation of the same part of the report to ensure consistency between written and oral presentation.

Reflection: We are pleased to see that the depth information is integrated into the system framework and can be successfully applied to the actual pedestrian tracking process.

Analyzation: Our system did not perform well on some metrics, we thought about the possible reasons and will continue to address this issue.

Conclusions: This meeting completed the system function verification, experimental evaluation and follow-up task arrangement. The team will prepare the Part II report and Topic Presentation based on the determined division of labor to ensure consistency of content.

8. **Date:** March 30, 2025

Participant: Allan Cheboiwo, Tarek Alkabbani, Haoyu Wang, and Vanessa Laurel Hariyanto

Meeting content:

Based on the testing and analysis in the past few days, through the in-depth analysis of the test results, we found that the performance of the system in several metrics is unstable, mainly because the depth estimation model currently used is not suitable for the target scene, resulting in large fluctuations in the depth prediction value. We have written this finding into our report and explained our results in detail.

In view of the tight schedule of the project, we also conducted a centralized review of all existing submissions during the meeting to ensure that the content of the documents is complete and the format is standardized. All members reviewed the completed Part II report paragraph by paragraph, focusing on the completeness of the content, the clarity of the technical logic, and whether the format specification met the submission requirements.

Reflection: This meeting allowed the team to clearly identify a key limitation in the current system. Collaborative efforts in reviewing both documentation and system implementation helped reinforce the importance of early model validation and iterative testing.

Analyzation: Through careful evaluation of performance metrics, we found that fluctuating depth estimates were the key cause of tracking inaccuracies. The parallel review of deliverables also ensured that both written and visual materials meet the required standards for final submission.

Conclusions: The team identified a technical bottleneck affecting system reliability. Final report was reviewed, and work will continue on improving system robustness through targeted testing and debugging.

References:

[1] Luo, W., Xing, J., Milan, A., Zhang, X., Liu, W., & Kim, T. K. (2021). Multiple object tracking: A literature review. *Artificial intelligence*, *293*, 103448.

[2] Hassan, S., Mujtaba, G., Rajput, A., & Fatima, N. (2024). Multi-object tracking: a systematic literature review. *Multimedia Tools and Applications*, *83*(14), 43439-43492.

[3] Elijah Cole, Maggie, Phil Culliton, sbeery, and Ștefan Istrate. iWildcam 2021 - FGVC8. https://kaggle.com/competitions/iwildcam2021-fgvc8, 2021. Kaggle.

[4] https://paperswithcode.com/methods/category/multi-object-tracking-models

[5] Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016, September). Simple online and realtime tracking. In *2016 IEEE international conference on image processing (ICIP)* (pp. 3464-3468). Ieee.

[6] Cao, H., Pappas, G. J., & Atanasov, N. (2024). PKF: Probabilistic Data Association Kalman Filter for Multi-Object Tracking. *arXiv preprint arXiv:2411.06378*.

[7] https://www.kaggle.com/datasets/ahmedfawzyelaraby/kinect-v2-multi-objects-with-3d-positions/data

[8] https://github.com/abewley/sort

[9] https://github.com/ultralytics/ultralytics

[10] https://github.com/isl-org/MiDaS

[11] https://github.com/JSaunders97/kalman-multi-object-tracker

[12] https://motchallenge.net/