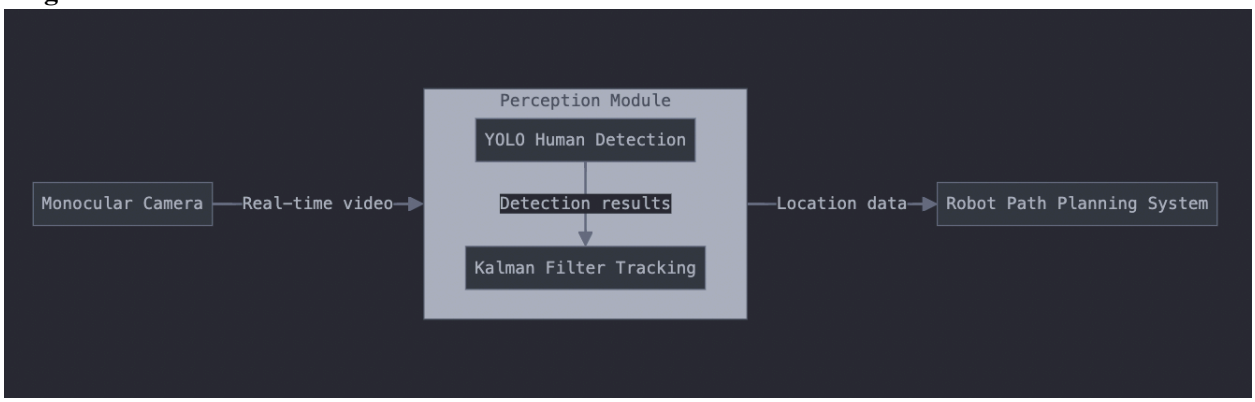
Multipurpose Perception Module for Acme

Audience: Acme Chief Technology Officer

1. Overview and Purpose of the Project

The **Multipurpose Perception Module** is a flexible system designed to improve the perception capabilities of Acme's robotic platforms by detecting and tracking one or more humans ($N \geq 1$) in real-time using a monocular camera. It employs a pre-trained YOLO model for human detection and a Kalman filter for smooth trajectory estimation, providing human location data directly usable in the robot's reference frame. This module enhances situational awareness, allowing robots to safely navigate around human obstacles, optimize path planning, and interact proactively in dynamic environments such as industrial inspection, healthcare, and collaborative assistance. The system integrates seamlessly with existing navigation and control modules, enabling effective path adjustments in real time.

Diagram:



3. Design and Development Process

The development process will involve iterative cycles to ensure quality and robustness at every stage. Key steps include:

- **Design Phase**: Create the system architecture, including class diagrams and sequence diagrams to depict interactions between components.
- **Development Phase**: Implement the module in C++ with a focus on modular design and maintainability.
- **Testing Phase**: Perform rigorous testing, including unit tests, integration tests, and end-to-end system validation.

Quality Assurance:

- **Unit Testing**: Developed using GoogleTest to validate each component.
- **Integration Testing**: Ensuring seamless data flow between the detection and tracking components.
- **Performance Benchmarks**: Evaluating detection and tracking latency to ensure real-time performance under varying conditions.

4. Technologies to be Used

- **Programming Language**: C++ for high efficiency and real-time performance.
- **Build System**: CMake to manage builds and dependencies.
- **External Libraries**:
    - **OpenCV**: For image processing and video stream handling (Apache 2 License).
    - **GoogleTest**: For unit testing and ensuring code quality.

5. Algorithms and Techniques

- **Human Detection**: **YOLOv4** will be used for accurate, real-time detection of humans in video frames, generating bounding boxes.
- **Human Tracking**: **Kalman Filters** will predict human positions and estimate trajectories for stable tracking.
- **Coordinate Transformation**: Detected positions will be transformed from the camera frame to the robot's reference frame for navigation.

6. Risks and Unknowns

Technical Risks:

- **Depth Estimation**: Challenges with depth estimation using a monocular camera will be mitigated by estimating human size and movement.
- **Real-Time Synchronization**: Latency issues in detection and tracking will be addressed through efficient threading and optimizations.
- **Hardware Variability**: Variations in platform capabilities may affect performance, mitigated by optimizing model size.

Project Risks:

- **Kalman Filter Limitations**: Ineffectiveness in highly dynamic environments will be mitigated by parameter tuning and thorough testing.

7. Final Deliverables to Acme

1. **Perception Module**: A robust and efficient module capable of detecting and tracking humans using monocular camera input.
2. **Documentation**: Detailed documentation, including installation steps, usage guides, and integration guidelines.
3. **Test Cases and Results**: Comprehensive unit and integration tests, with code coverage exceeding 90%.
4. **UML Diagrams**: Detailed diagrams illustrating the system architecture, interactions, and class hierarchies.

8. Team and Pair Programming Execution

- The software development cycle will use pair programming. Sarang Shibu will work on the human detection component of the perception system, Prathinav Karnala Venkata will work on the human tracking system, and Abhey Sharma will act as the design keeper, ensuring adherence to the design and OOP best practices.