IMAGE PROCESSING TECHNIQUES TO IMPROVE APRIL TAG POSE ESTIMATION IN LOW-COST SYSTEMS

Connor Guzikowski, Carson Rae, Derrick Eliasi

Department of Electrical and Computer Engineering, University of California, San Diego

ABSTRACT

Accurate and efficient pose estimation of fiducial markers is essential for robust localization and mapping in robotic systems. This project aims to improve the accuracy and computational performance of AprilTag pose estimation in both single-image and SLAM-based use cases. In the first phase, we employ an open source AprilTag pose estimation library developed to perform controlled single image tests under varying conditions of viewing angle, distance, fog, lighting, and motion blur. In these tests, we will perform various image processing techniques such as gamma correction, histogram equalization, Retinex-based illumination correction, Gaussian and bilateral filtering for denoising, and deblurring through Wiener and blind deconvolution methods. Additionally, we plan to apply contrast-limited adaptive histogram equalization (CLAHE) and color balance adjustments. These experiments will quantify how such different image processing algorithms influence both pose accuracy and processing speed. In the second phase, these techniques are integrated into the TagSLAM [1] framework to evaluate their effects in multi-frame SLAM scenarios. To evaluate performance changes with image processing techniques, we will measure improvements in trajectory accuracy, drift, and computation time compared to the baseline system. By jointly analyzing accuracy-speed trade-offs across both single-image and SLAM contexts, this work aims to find improvements to an already useful system without sacrificing processing speed.

***Index Terms*—** AprilTag, fiducial markers, pose estimation, TagSLAM, image processing, low light, fog, motion blur, robotics

1. INTRODUCTION

AprilTags are one of the most common tools used for pose estimation in robotics because they provide a fast and reliable way to calculate a camera's position and orientation. They work well in controlled environments, but in real-world conditions the accuracy can drop significantly when lighting or visibility changes. Low-light scenes, fog, glare, and motion blur all reduce the contrast between the black and white tag regions, making it harder for detectors to locate corners accuratly. These small detection errors can accumulate in systems like TagSLAM [1], where pose estimation depends on consistent tag observations over multiple frames.

The TagSLAM research combines tag detections across many frames using a factor graph optimization approach in order to extend the AprilTag framework. It provides accurate pose estimation in normal imaging conditions, but it assumes that the tag detections are always reliable. Moreover, low-cost cameras used in robotics, such as Raspberry Pi cameras, often operate and struggle under poor lighting or environmental interference. In these settings, TagSLAM's accuracy is negatively affected. These problems are even more noticeable in low-power systems that do not have extra sensors or built-in lighting to help with detection.

To address these challenges, our project aims to improve AprilTag pose estimation by integrating image processing before tag detection. The focus is to evaluate how image enhancement methods affect both single-frame accuracy and multi-frame consistency within the TagSLAM framework. We plan to test the system under three main visual conditions that typically reduce performance: low-light and harsh shadow environments, motion blur, and foggy or hazy scenes. Each of these conditions introduces unique image distortions that interfere with accurate tag detection and reliable pose estimation.

In this project, we will test and compare different image enhancement methods, such as histogram equalization, Retinex-based correction, dehazing, and contrast normalization, to see which help the most to improve tag visibility and stability. After enhancing the images, we will add these versions to the TagSLAM pipeline and check how they affect pose accuracy, drift, and runtime. Ultimately, our goal is to make tag detection more reliable in real-world conditions while keeping the system fast and lightweight for low-cost robots.

2. RELATED WORKS

2.1. AprilTag Estimation

Several prior studies have sought to improve aspects of the April Tag based pose estimation, including accuracy, robustness, and computational latency. One of particular interest is the work done by Prfommer and Daniiilidis [1], who introduced *TagSLAM*, a system that integrates fiducial markers into a full Simultaneous Localization and Mapping framework. Unlike normal AprilTag detection pipelines that operate on individual frames, TagSLAM creates a factor graph that merges observations from multiple cameras and markers over time. This formula allows for global optimization of both camera and tag poses, significantly decreasing drift and performing with greater spatial consistency across larger environments. Their results demonstrated that incorporating fiducial markers in a graph-based optimization pipeline leads to more reliable localization, specifically in indoor and spaces without GPS.

While TagSLAM focuses on high-level pose graph optimization and mapping accuracy, its performance builds heavily upon the quality of the raw tag detections feeding into it's system. Previous works, such as Olson's foundational AprilTag system [2], emphasize detection speed and tag design for geometric strength, but they typically assume ideal image conditions. As a result, both frameworks are at-risk to degradation in quality under challenging visual scenarios such as fog, motion-blur, or low light. All of which, are conditions frequently encountered by UAV and robotic systems. This project builds upon the TagSLAM framework by addressing these limitations at the image-processing level, introducing preprocessing methods aimed at increasing the reliability of AprilTag detections prior to pose estimation.

2.2. Condition 1: Foggy Conditions

2.2.1. Enhancement Based Defogging Techniques

Enhancement-based defogging approaches aim to improve image visibility in foggy conditions through classical enhancement operations such as contrast stretching, histogram equalization, Retinex filtering, multi-scale fusion, and edge-preserving filtering. These methods focus on increasing local contrast and restoring suppressed color and texture features rather than modeling the physical scattering process that causes fog [3].

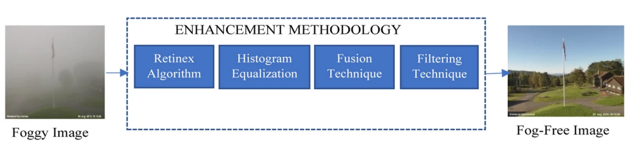

Fig. 1. Enhancement Based Image Processing Pipeline [3]

While these methods are computationally efficient and simple to implement, they often over-enhance or distort natural image statistics due to the absence of physical modeling. Consequently, enhancement-based techniques may produce oversaturated or geometrically inconsistent results, which can negatively impact downstream vision tasks such as corner detection and pose estimation in fiducial-based systems.

2.2.2. Prior-Based Defogging Techniques

Prior-based methods rely on the atmospheric scattering model and incorporate statistical or structural assumptions—known as priors—to estimate the transmission map and atmospheric light parameters [3]. These techniques balance physical interpretability and computational practicality by enforcing constraints derived from empirical observations of natural images.

Common priors include the *Dark Channel Prior (DCP)* [4], which assumes at least one color channel has low intensity in fog-free patches, and the *Color Attenuation Prior (CAP)* [5], which models the relationship between scene depth and color differences in brightness and saturation. The *Haze Line Prior (HLP)* [6] captures pixel distributions along haze lines in RGB space, while the *Color Ellipsoid Prior (CEP)* and *Difference Structure Preservation Prior* [3] extend these concepts to better model pixel distributions and maintain structural consistency.

These physically grounded techniques typically produce more accurate and artifact-free results than enhancement-based methods, especially in scenes that follow the assumed priors. However, they may fail under conditions such as non-uniform fog, bright sky regions, or non-Lambertian surfaces, leading to halo effects or incorrect depth estimations. For real-time applications like pose estimation and SLAM, prior-based methods offer improved geometric fidelity but at a higher computational cost.

2.3. Learning Based Techniques

Deep learning approaches have recently become prominent in image defogging due to their ability to model complex nonlinear relationships between foggy and clear images [3]. These models use convolutional neural networks (CNNs) or generative adversarial networks (GANs) trained on large datasets to directly predict clear images or intermediate representations such as transmission maps.

Notable examples include *DehazeNet* [7], which estimates transmission maps in an end-to-end manner, and *Ranking-CNN* [8], which introduces ranking layers to preserve structural relationships between foggy and clear features. Multi-resolution and cascaded CNN architectures combine global and local features to improve robustness, while models like *FAMED-Net* [9] emphasize lightweight, real-time inference.

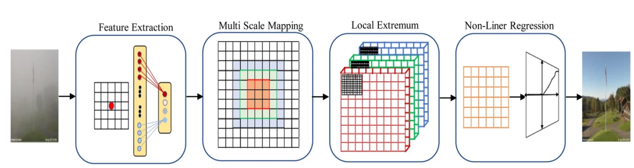

Fig. 2. DehazeNet Block Diagram [7]

Learning-based methods achieve high visual quality and generalization across fog densities but depend heavily on large, representative datasets. They are also prone to domain shift when trained on synthetic data and may introduce latency that limits real-time deployment. In fiducial-based vision systems, these techniques can significantly enhance tag visibility, though their effect on geometric accuracy must be carefully validated.

2.4. Condition 2: Motion Blur

2.4.1. Improved Handling of Motion Blur in Online Object Detection

Sayed *et al.* [10] investigated the effect of motion blur on object-detection accuracy and proposed an adaptive framework that works to condition detection models on the specific type and degree of motion blur. Instead of attempting to simply clear all blur, their method incorporates blur information into the detection process itself. The authors worked to develop synthetic motion-blur augmentation and blur conscious label generation to train convolutional networks to be capable of distinguishing blur patterns and compensating for them during inference. Experiments ran on the COCO dataset found that blur-aware models achieve a higher mean Average Precision (mAP) than standard detectors when tested on real blurred imagery, proving that blur-specific training can successfully mitigate accuracy loss.

This study is valuable because it looks at motion blur as a learning problem rather than a correctional problem. However, the approach is quite costly as it relies heavily on synthetic blur models and large-scale training, which is impractical for lightweight and on-the-go systems. While effective for object detection, it does not impact the geometric consistency needed for pose estimation. In contrast, this paper's proposed work focuses on pre-processing to restore image sharpness before detection, ensuring continued compatibility with existing AprilTag pose estimation pipelines without retraining detection networks.

2.4.2. Ghost-DeblurGAN for Fiducial Marker Detection

Liu *et al.* [11] introduced Ghost-DeblurGAN, a deep generative model designed to correct motion blur in images containing fiducial markers such as ArUco or AprilTag. The authors constructed a specialized dataset, YorkTag, that includes synthetic and real blurred images of markers under differing camera and object motions. The network uses a residual generator in combination with a perceptual loss that preserves the high-frequency texture details critical for reliable marker decoding. Experimental results showed vast improvement in marker detection rates and pose-estimation accuracy after deblurring, performing better than classical methods like Wiener and Richardson-Lucy filtering in both sharpness and downstream detection metrics.

This work is particularly relevant because it directly looks at deblurring with fiducial markers as a performance metric, rather than overall image quality. Its use of GAN-based restoration indicates that even subtle motion blur can drastically reduce detection reliability. However, the model's computational complexity and dependence on a large training set limits its portability to mobile robots, such as a UAV. The present project builds on this concept by looking to explore more lightweight image-processing techniques that can mitigate motion blur without retraining large networks, aiming for real-time improvement of AprilTag pose estimation in dynamic robot environments.

2.5. Condition 3: Low Light/Harsh Shadows

2.5.1. Low-Light Enhancement for Object Detection

One relevant study by Guo *et al.* [12] proposed a dynamic end-to-end framework for object detection in low-light environments. Instead of following a standard process of enhancing the environment and then detecting objects, their method combines both steps into a single network. This allows the system to automatically choose which enhancement path to apply based on the lighting conditions and noise level of each image. Their experiments on the ExDark dataset showed improvements in detection accuracy, especially for medium-sized objects, while keeping the overall computation time reasonable.

This approach works well because it adjusts the enhancement on a per-image basis instead of using one fixed model. That makes it more reliable in uneven or harsh shadow lighting scenes. However, the model is more complex and takes longer to train, making it less ideal for low-power or real-time systems. Since the design is focused on object detection and not geometric localization, it would also need modification before being applied to tasks like AprilTag pose estimation.

2.5.2. Low-Light Enhancement for License Plate Recognition

Another relevant study addressing visibility degradation in dark scenes is by Saputra *et al.* [13], who proposed a low-light image enhancement pipeline for license-plate recognition using URetinx-Net and TRBA. Their approach demonstrates a Retinex-based deep neural network to separate illumination and reflectance components, followed by a text-recognition network to improve character clarity under illumination levels below 20 lux.

The paper displays how the URetinx-Net processing method improves image contrast and readability compared to simple enhancement methods such as histogram equalization and CLAHE. Using this method, they were able to reach higher recognition accuracy while maintaining stable color balance, making it suitable for embedded or mobile systems. Furthermore, using this approach worked well in different light settings, illustrating how Retinex-based deep networks can adapt to diverse nighttime scenes.

Although the method performs well for improving text visibility, it is mainly designed for 2-D license plate recognition rather than pose estimation. The approach also requires heavy computation and supervised training data, making it less practical for real-time or low-cost robotic systems.

3. PROPOSED METHODS

3.1. AprilTag Estimation

The AprilTag system is the starting point for our work, as it provides a simple and reliable way to estimate a camera's position and orientation from a known marker. Our setup uses the open-source AprilTag 3 library, which is also used in the TagSLAM framework. TagSLAM builds on the normal AprilTag detector by combining tag observations from many frames using a factor graph, which helps reduce drift and improve overall consistency during motion.

In the basic AprilTag detection pipeline, the input image is first converted to grayscale and then thresholded to separate the black and white tag regions. Contours are found and grouped into quadrilateral shapes that match the tag layout. Each possible tag is then decoded using its unique binary ID pattern, which includes built-in error correction to avoid false detections even when light or noise affects the image.

Once the corners are detected, the 3D pose of the tag is calculated using a Perspective-n-Point (PnP) solver. With the known tag size and the intrinsic parameters of the camera, the system computes the transformation between the tag and the camera as:

$$\mathbf{T}_{cw} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, \tag{1}$$

where $\mathbf{R}$ represents the rotation and $\mathbf{t}$ represents the translation of the camera relative to the tag.

TagSLAM takes these individual AprilTag poses and refines them over multiple frames by optimizing all tag and camera positions together. This allows for smoother and more accurate trajectories, especially when the same tags are seen repeatedly. In our project, we use this setup as the baseline and then add different image enhancement techniques before the tag detection step. This lets us test how improved image quality can help TagSLAM detect corners more accurately and produce better overall pose estimates.

3.2. Condition 1: Defogging

In this project we will explore the impact of three defogging strategies: enhancement-based, prior-based, and learning-based techniques for AprilTag pose estimation accuracy and processing efficiency. Our goal is to evaluate how each approach affects fiducial detection reliability and runtime in both standalone and SLAM-integrated image sequences using TagSLAM.

3.2.1. Enhancement-Based Defogging via Retinex

For the enhancement based pipeline we will use a Python implementation of Single Scale Retinex (SSR) and Multi Scale Retinex (MSR) from the open-source project "Retinex-Image-Enhancement". [14] Retinex theory is based on decomposing the observed image into illumination and reflectance components:

$$I(x) = R(x)\,L(x), \tag{2}$$

where $I(x)$ is the observed intensity at pixel x, $R(x)$ is the reflectance (scene content) and $L(x)$ is the illumination. In MSR, the enhanced reflectance estimate can be expressed as

$$R_{\mathrm{MSR}}(x) = \sum_{n=1}^{N} w_n \left[\log I(x) - \log\big(I(x) * F_n(x)\big)\right], \tag{3}$$

where $F_n(x)$ are Gaussian kernels of varying scales, and w_n are weights summing to one. We will apply this implementation to foggy images of fiducial tags to boost contrast, reveal hidden edges and corners, and then apply our tag-detection pipeline. The key metrics will be pose error and frame processing time.

3.2.2. Prior-Based Defogging: Dark Channel Prior (DCP)

We will use an implemented Dark Channel Prior method from [15] following the classical atmospheric scattering model:

$$I(x) = J(x)\,t(x) + A\big(1 - t(x)\big), \tag{4}$$

where $I(x)$ is the observed foggy image, $J(x)$ is the scene radiance, A is global atmospheric light, and $t(x)$ is the transmission map. The dark channel of a clear image is defined as:

$$J_{\mathrm{dark}}(x) = \min_{y\in\Omega(x)} \Big(\min_{c\in\{r,g,b\}} J^c(y)\Big), \tag{5}$$

and the transmission estimate:

$$t(x) = 1 - \omega \min_{y\in\Omega(x)} \Big(\min_{c} \frac{I^c(y)}{A^c}\Big), \tag{6}$$

with $\Omega(x)$ the local patch around pixel x. The recovered radiance is then:

$$J(x) = \frac{I(x) - A}{\max\big(t(x),\, t_0\big)} + A, \tag{7}$$

where t_0 is a lower bound on transmission to avoid division artifacts. Guided filtering is added for edge-aware smoothing of $t(x)$. The output will serve as input to the tag detection and pose solver.

3.2.3. Learning-Based Defogging: DehazeNet Implementation

We will adopt the DehazeNet framework [16] as a learning-based defogging method. DehazeNet accepts a single foggy image as input and predicts a medium transmission map via a deep convolutional architecture engineered to embed haze-relevant features such as scattering and attenuation.

Once the transmission map $t(x)$ is estimated by the network, the haze-free image $J(x)$ is recovered using the standard atmospheric scattering model:

$$I(x) = J(x)\,t(x) + A\,(1 - t(x))$$

where $I(x)$ is the observed hazy image, A the global atmospheric light, and $t(x)$ the predicted transmission map. We will use the pretrained model provided in the DehazeNet repository and integrate it into our processing pipeline.

We will compare the performance of DehazeNet with our enhancement-based (Retinex) and prior-based (DCP) pipelines across both single-image tests and SLAM experiments to evaluate the trade-offs between accuracy, robustness, and runtime between the three.

3.3. Condition 2: Motion Blur

This section focuses on restoring motion-blurred images prior to AprilTag pose estimation. Two classical image restoration techniques, Inverse Filtering and Wiener Filtering are implemented and evaluated for their ability to correct linear motion blur. Both methods work in the frequency domain using a mathematically modeled degradation function $H(u, v)$ that represents the spread function (PSF) of motion blur aligned with the direction of the camera movement. See the following image for an example of what this project aims to accomplish via these filters:

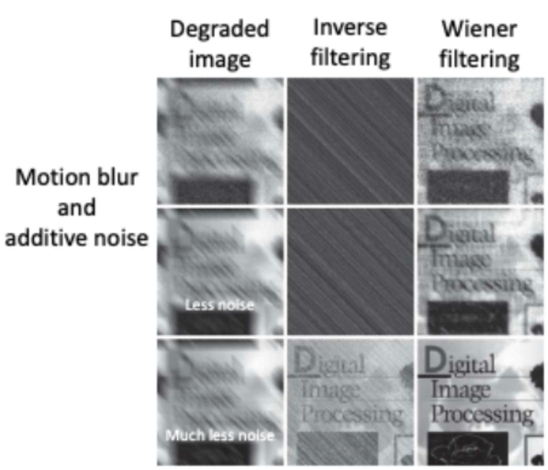

Fig. 3. Adapted from ECE 253 Lecture 7 Slides [17]

3.3.1. Inverse Filtering

Inverse filtering utilizes a direct approach to deconvolution by assuming that the degradation function $H(u, v)$ is well known and invertible.

$$G(u, v) = H(u, v)F(u, v) \tag{8}$$

where $F(u, v)$ is the original image, and the restored estimate is obtained by division:

$$\hat{F}(u, v) = \frac{G(u, v)}{H(u, v)} \tag{9}$$

To avoid division by zero where $H(u, v)$ is small, a small constant or a low-pass filter is introduced:

$$\hat{F}(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 + \epsilon}\,G(u, v), \tag{10}$$

where ϵ is a stabilization term. The PSF $h(x, y)$ is constructed according to the motion length L and angle θ of the blur, and its Fourier Transform $H(u, v)$ is used for restoration. Inverse filtering can recover high-frequency detail effectively when noise is negligible, but it is highly sensitive to additive noise, often amplifying high frequency artifacts and ringing. Despite some limitations, this method works as a strong baseline for evaluating more robust approaches.

3.3.2. Wiener Filtering

The Wiener filter extends inverse filtering by incorporating a statistical model of noise and signal power, minimizing the mean-square error between the estimated and original images. It's frequency domain formula is given by:

$$\hat{F}(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 + \frac{S_n(u,v)}{S_f(u,v)}}\,G(u, v), \tag{11}$$

where $S_n(u, v)$ and $S_f(u, v)$ are the power of the noise and unblurred image, respectively. In practice, the ratio $\frac{S_n}{S_f}$ is approximated by a constant K, getting the simplified expression

$$\hat{F}(u, v) = \frac{H^*(u, v)}{|H(u, v)|^2 + K}\,G(u, v). \tag{12}$$

This formula suppresses noise amplification while maintaining edge correctness. When it is applied to motion blurred images, the Wiener fileter achieves clearer and more visually correct reconstructions that inverse filtering. However, its effectiveness also depends on accurate estimation of the PSF and noise level. Overly large K values can lead to over smoothed results. In the context of AprilTag pose estimation, Wiener filtering can be expected to produce more stable tag detection under realistic noisy image conditions, providing a balance between sharpness recovery and noise suppression.

3.4. Condition 3: Low Light/Harsh Shadows

Low-light and shadowed scenes make it harder for AprilTag detectors to find tag corners accurately because the contrast between black and white regions becomes weaker. When this happens, TagSLAM's pose optimization also becomes less reliable, since it depends on the tag detections from each frame. To fix this, we plan to add an adaptive image enhancement preprocessing step before the AprilTag detection stage so that darker or unevenly lit frames are automatically corrected before being processed.

3.4.1. Adaptive Enhancement Processing

During the enhancement step, we plan to first look at the average brightness and contrast of each image. Moreover, we will compute the mean intensity μ_I and standard deviation σ_I of the grayscale image $I(x, y)$ as:

$$\mu_I = \frac{1}{N}\sum_{x,y} I(x,y), \qquad \sigma_I = \sqrt{\frac{1}{N}\sum_{x,y}(I(x,y)-\mu_I)^2}. \tag{13}$$

If μ_I is below a certain brightness threshold τ_{low}, that means the image is too dark. In that case, we plan to try applying a gamma correction and histogram equalization to boost the dark areas, using a similar approach to the method proposed by Vallabhaneni and Sriram [18], which combines transform-based gamma correction with histogram equalization for contrast enhancement.

$$I'(x,y) = \text{HE}\big(I(x,y)^{\gamma}\big), \tag{14}$$

where HE$(\cdot)$ stands for histogram equalization and γ is between 0.6–0.8 to brighten shadows.

Another method we plan to try for complex lighting, like strong shadows or uneven brightness, is to use a simple Retinex-based correction inspired by Jobson et al. [19], which models image intensity as the product of illumination and reflectance:

$$R(x,y) = \log(I(x,y)) - \log\big(G_\sigma * I(x,y)\big), \tag{15}$$

where G_σ is a Gaussian blur that estimates background illumination. The final enhanced image will then combine global and local adjustments as:

$$I_{\text{enh}}(x,y) = \alpha\, R(x,y) + (1-\alpha)\, I'(x,y), \tag{16}$$

with α (around 0.7) balancing between detail recovery and overall brightness. By comparing both algorithms, we will ultimately decide which is the best enhancement preproessing method of the two. Furthermore, we will apply this method into our AprilTag Detection system.

3.5. Integration into TagSLAM

After the image is enhanced, it will go directly into the AprilTag detector used in TagSLAM. For each frame t, the detector gives the corner positions of the tag $\mathbf{z}_t$ and an estimated pose of the camera $\mathbf{T}_{cw,t} \in SE(3)$. The backend will then minimize the total reprojection error:

$$E = \sum_t \sum_i \|\, \pi(\mathbf{T}_{cw,t}\, \mathbf{P}_i) - \mathbf{z}_{i,t} \,\|^2, \tag{17}$$

where $\pi(\cdot)$ is the projection model and $\mathbf{P}_i$ are the tag corner points. By improving image contrast before detection, we expect smaller pixel errors and better pose accuracy during optimization. In addition to this, by automatically adapting the brightness and contrast of the image before TagSLAM runs, this method should make pose estimation more consistent in various environments without slowing down the system. To measure improvements, we will compare the modified system with the original TagSLAM in multiple test scenes using pose accuracy, average corner pixel deviation, and computation time.

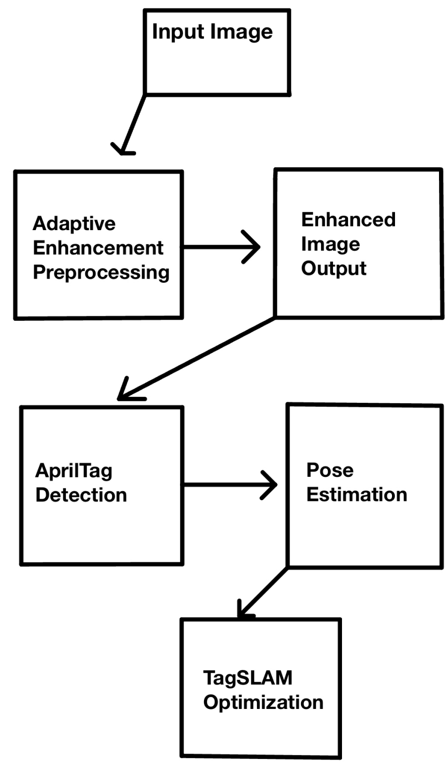

Fig. 4. Proposed pipeline for improving AprilTag pose using advanced preprocessing methods.

4. EXPERIMENT SETUP/DATASET

To evaluate our proposed enhancements to AprilTag pose estimation, we will test how they affect the speed and accuracy in both single image tests of pose estimation, and in trajectory tracking using the TagSLAM method. For the single image tests, we will collect our own data, taking single images of AprilTags under varying conditions and from varying positions. We will capture images in fog, under motion, and different lighting conditions to test the effectiveness of our methods. The images will be captured with a low-cost Raspberry Pi camera. To get the ground truths of the poses of the images, we will use an Oak-D Lite depth camera to measure how far away the april tags are. The Oak-D Lite and Raspberry Pi camera will be in a fixed position relative to each other using a 3D-printed jig.

To test the effectiveness in a multi-image use case, the existing TagSLAM dataset provides a reliable benchmark for testing pose estimation algorithms in structured environments with known tag arrangements and ground-truth trajectories. This dataset will be used as a baseline for accuracy and performance comparisons before and after applying our image processing enhancements.

While pose estimation accuracy is important, we will also be measuring the overall computation timing for our image processing pipelines. In real world robotics applications, not only is accuracy important, but speed matters as well, because of this we will be try to find the method that best improves the accuracy of the system while still maintaining a relatively low computational cost.

Using the dataset collected by TagSLAM, alongside our own collected images, we will be able to evaluate performance in both ideal and non-ideal conditions. This hybrid approach will help determine how much our preprocessing pipeline improves pose accuracy, stability, and keeps low computation time in realistic deployment settings.

5. CONCLUSION

This project aims to make AprilTag-based pose estimation more reliable while maintaining efficiency when dealing with challenging visual conditions. We plan to test several image processing methods on both single-image and SLAM-based setups to see which ones improve accuracy without slowing things down too much. By adding these techniques to the TagSLAM framework, we hope to better understand how image quality and environmental factors affect SLAM performance. Overall, our goal is to create a more robust fiducial-based localization system that can maintain good accuracy in different real-world settings.

6. REFERENCES

[1] Bernd Pfrommer and Kostas Daniilidis, "Tagslam: Robust slam with fiducial markers," *arXiv preprint arXiv:1910.00679*, 2019.

[2] Edwin Olson, "Apriltag: A robust and flexible visual fiducial system," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. 2011, pp. 3400–3407, IEEE.

[3] Pooja Pandey, Rashmi Gupta, and Nidhi Goel, "Comprehensive review of single image defogging techniques: enhancement, prior, and learning based approaches," *Artificial Intelligence Review*, vol. 58, no. 4, pp. 116, 2025.

[4] Kaiming He, Jian Sun, and Xiaoou Tang, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341–2353, 2011.

[5] Yingying Zhu, Gaoyang Tang, Xiaoyan Zhang, Jianmin Jiang, and Qi Tian, "Haze removal method for natural restoration of images with sky," *Neurocomputing*, vol. 275, pp. 499–510, 2018.

[6] Dana Berman, Tali Treibitz, and Shai Avidan, "Single image dehazing using haze-lines," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 3, pp. 720–734, 2020.

[7] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao, "Dehazenet: An end-to-end system for single image haze removal," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5187–5198, 2016.

[8] Yafei Song, Jia Li, Xiaogang Wang, and Xiaowu Chen, "Single image dehazing using ranking convolutional neural network," *IEEE Transactions on Multimedia*, vol. 20, no. 6, pp. 1548–1560, 2018.

[9] Jing Zhang and Dacheng Tao, "Famed-net: A fast and accurate multi-scale end-to-end dehazing network," *IEEE Transactions on Image Processing*, vol. 29, pp. 72–84, 2020.

[10] Mohamed Sayed, Gregory Shakhnarovich, and Ranjitha Kumar, "Improved handling of motion blur in online object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021, pp. 10953–10962, IEEE.

[11] Yijun Liu, Rong Zhang, and Jiayi Ma, "Ghost-deblurgan: Motion blur correction for fiducial marker detection," *arXiv preprint arXiv:2109.03379*, 2021.

[12] Haifeng Guo, Tong Lu, and Yirui Wu, "Dynamic low-light image enhancement for object detection via end-to-end training," in *Proceedings of the 25th International Conference on Pattern Recognition (ICPR)*, Milan, Italy, 2021, pp. 5611–5618, IEEE.

[13] Vriza Wahyu Saputra, Nanik Suciati, and Chastine Fatichah, "Low light image enhancement in license plate recognition using uretinx-net and trba," *Procedia Computer Science*, vol. 234, pp. 404–411, 2024.

[14] AKRISH22, "Retinex-image-enhancement: Practical implementation of single scale retinex and multiscale retinex algorithm in python3," .

[15] He-Zhang, "Single image haze removal using dark channel prior," .

[16] Caibolun, "Caibolun/dehazenet: Dehazenet: An end-to-end system for single image haze removal," .

[17] Ben Ochoa, "Ece 253 lecture 7: Image restoration," Lecture slides, Department of Electrical and Computer Engineering, University of California San Diego, 2025, Accessed from course materials on Canvas for ECE 253: Digital Image Processing.

[18] Sai Kumar Vallabhaneni and Varun Sriram, "Contrast enhancement of colour images using transform based gamma correction and histogram equalization," M.S. thesis, Blekinge Institute of Technology, 2017.

[19] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for color image enhancement," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 1997, vol. 3, pp. 451–455.