SLAM in Reduced Vision using Dehazing Techniques

Ankit Pensia [1], Kartik Gupta [2] and Ramavtar Malav [3]

***Abstract*— In this project, we use image processing technique of Single Image Dehazing to improve the accuracy of monocular SLAM in the presence of smoke and fog. For evaluation of proposed pipeline, synthetic dataset was generated using fog model.**

I. INTRODUCTION

Presently SLAM techniques which make use of low-cost sensors do not perform well in harsh environments. Harsh environments usually have reduced vision due to smoke and fog. Although Thermal Imaging Cameras (TIC) are robust to smoke but they are rather expensive. We believe that data from other low-cost sensors can be processed to increase the performance of SLAM in such environments. We wish to develop a low-cost method to perform SLAM in indoor and harsh environment, using the already-developed techniques in the field of image processing and SLAM.

Autonomous bots can be of great assistance in monitoring the area in hostile environments. Presently, SLAM techniques are generally developed to work in controlled environment. In controlled environment, GPS and Lasers can be used to get the information about environment without much error. GPS cannot be used indoors and lasers perform poorly in smoky environments. Monocular SLAM has been greatly studied but cameras images are also affected in smoke. Most of the recent work in such environments has been in the field of data fusion. However, By using image processing, we can increase the usability of the image(by dehazing) and thus can improve the performance of monocular SLAM.

II. PREVIOUS WORK

People have mostly tackled the problem of SLAM in harsh environments by fusing the data from multiple sensors. One of the first such work which we came across was by Sales [1]. Although, Sales [1] was not implementing SLAM, data from two complimentary sensors were used. Santos[2] then implemented data fusion layer for SLAM in ROS. Brunner[3] used Thermal Imaging Devices(TICs) to increase the robustness in smoky environment but TICs are expensive and thus dont satisfy our objective.
For dehazing techniques, Fattal[4] established a standard and benchmark in corresponding domain. Later a new approach using dark channel prior by He [5] was found to have improved results. An interesting approach is given in using a dehazing pipeline by Agarwal[6].
A brief summary of these relevant works is provided below.

1) SLAM Related:

- Multi-Sensor Person Following in Low-Visibility Scenarios[1], although not used for SLAM, proposes methods to perform a person following in low visibility conditions. Readings from Laser Range finders and sensors are weighted by presence of smoke in environment. Visual images are used to quantify the presence of smoke but it requires the clear image of same environment. The authors also conducted a series of experiment to measure and compare performance of Laser Range Finders(LRF) and SONAR in different smoke densities as shown in Figure 1.
  As can be seen from the Figure 1, while the measurements obtained form the laser sensor are accurate for low density, it becomes highly inaccurate for high density smoke. On the other hand, the measurements obtained using the sonar sensor are noisy for all smoke density levels.
- In SmokeNavSimultaneous Localization and Mapping in Reduced Visibility Scenarios[2] ,a 2D SLAM technique for low visibility scenarios was proposed. It made use of data from complementary sensors which improved the performance of SLAM in harsh environments. As for implementation, the authors used Gmapping algorithm (ROS) to read data from both Laser and Sonar. Their goal was to implement an 'intelligent layer', called SmokeNav by them, to build a probabilistic model to assign weights to sensors online. Still, the quality of map was affected by smoke as both sensors were affected by smoke.
- In Automatic Selection of Sensor Modality for Resilient Localisation in Low Visibility Conditions[3], the authors attempted to apply SLAM technique by combining information obtained using Thermal Imaging

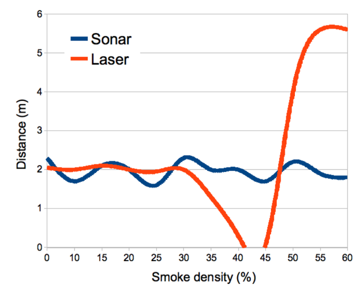

Fig. 1: Sonar and LRF performance w.r.t smoke density [1]

*This is a course project for EE698G

[1]Ankit Pensia is in department of Electrical Engineering, IIT Kanpur ankitp@iitk.ac.in
[2] Kartik Gupta is in department of Civil Engineering, IIT Kanpur kartikg@iitk.ac.in
[2] Ramavtar Malav is in department of Electrical Engineering, IIT Kanpur rmalav@iitk.ac.in

Cameras(TIC) and Visual Cameras(RGB camera) with the performance of TIC being the critical parameter. However in the presence of smoke the accuracy decreased even if one sensor remained unaffected. One drawback of the methodology was that the weights were not being assigned online with respect to the environment.

2) Dehazing Related:

- In his paper Single Image Dehazing [4], Fattal gives an approach for removing the haze layer and obtaining a reliable transmission estimate which may be used for image refocusing. This method also gives us a depth map of the image. A refined image formulation model is obtained that provides for transmission function as well as surface shading. This new approach is physically sound and relies only on the assumption that the transmission and surface shading are locally uncorrelated.
- Then a new technique, Single Image Haze Removal Using Dark Channel Prior [5] was given using dark channel prior that promised better dehazing results. he authors note that most local patches in haze-free outdoor images contain some pixels which have very low intensities in at least one color channel. This low intensity color channel, called dark channels are used for haze transmission estimation. Colluding this technique with a soft matting interpolation method a good depth map may also be produced. Experimental results show this approach to have better results compared with Fattals[5] approach in dense haze conditions.
- In the experimental study ,Visual Odometry in Smoke Occluded Environments [6] focus was put on comparing different techniques of Visual Odometry and Dehazing methods to be used together in smoke occluded environments. The study proposed using a dehazing pipeline with the pipeline component techniques being chosen after comparison studies. It uses a combination of contrast enhancement and depth based image enhancement for improving the image.

III. Proposed Pipeline

We propose using a dehazing process pipeline to tackle our problem. The pipeline takes foggy images as input, then selected dehazing techniques are applied on the image dataset followed by application of monocular SLAM techniques. Pipeline is shown in Figure 2 An advantage of adding dehazing to the pipeline is that no extra hardwares are required. We wish to use existing dehazing and monocular SLAM techniques, selected after literature survey and experiments. The selected pipeline component techniques are -

- Dehazing Technique :- Haze Removal using Dark Channel Prior [5]
- Monocular SLAM Used :- ORB SLAM [7]

The pipeline components are explained in detail.

A. Images with fog

The environments containing fog and smoke, when captured through normal camera will have smoke/fog. This will be the input to our pipeline. To evaluate our pipeline, we generate the datast by synthetically adding fog. It is described in detail later.

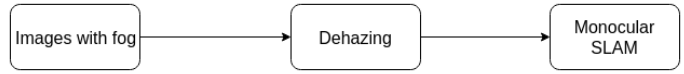

Fig. 2: Proposed Pipeline

B. Haze Removal using Dark Channel Prior

Images of outdoor scenes are usually degraded by the turbid medium (e.g ., particles, water-droplets) in the atmosphere. Haze, fog, and smoke are such phenomena due to atmospheric absorption and scattering. The irradiance received by the camera from the scene point is attenuated along the line of sight. Furthermore, the incoming light is blended with the airlight (ambient light reflected into the line of sight by atmospheric particles). The degraded images lose the contrast and color fidelity, as shown in Figure1(a). Since the amount of scattering depends on the distances of the scene points from the camera, the degradation is spatial-variant.

Hazing Model

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

where

- I stands for the observed image
- t is the portion of the light that is not scattered
- J(x)*t(x) is called direct attenuation
- A*(1-t(x)) is called airlight.

However, haze removal is a challenging problem because the haze is dependent on the unknown depth information. The problem is under-constrained if the input is only a single haze image. Therefore, many methods have been proposed by using multiple images or additional information.
The dark channel prior is based on the statistics of haze-free outdoor images. We find that, in most of the local regions which do not cover the sky, it is very often that some pixels (called dark pixels) have very low intensity in at least one color (rgb) channel. In the haze image, the intensity of these dark pixels in that channel is mainly contributed by the airlight. Therefore, these dark pixels can directly provide accurate estimation of the hazes transmission. Combining a haze imaging model and a soft matting interpolation method, we can recover a hi-quality haze-free image and produce a good depth map (up to a scale).

C. ORB Monocular SLAM

We have chosen to work with ORB-SLAM technique given by Mur-Artal et al. It is a feature-based monocular SLAM system that operates in real time, in small and large, indoor and outdoor environments. An important advantage of using this technique is that it allows loop closing and relocalization as well as it provides full automatic initialization.

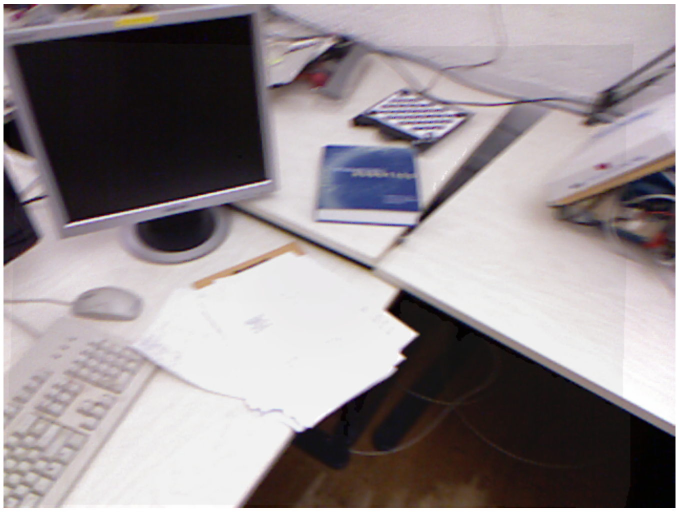

Fig. 3: Image obtained after hazing with fog coefficient, k = 0.005

Fig. 4: Image obtained after hazing with fog coefficient, k = 0.02

IV. Experimental Results

As mentioned in III we have generated synthetic fog and performed dehazing followed by application of SLAM. The following sections contain the results obtained at each step.

To evaluate the results of our pipeline, we required dataset of the same environment at different level of fog smoke, and with the ground truth. To perform experiment in controlled manner, we considered it better to add fog synthetically.

- Dataset : - TUM RGB-D [8] fr1desk
- Type of data :- RGB-D maps.
- Environment :- Indoor scene(office). It contains 4 desks and loop closures.
- No. of Images - 600 images.

A. Dataset Generation

By Koschmieders law,
$L(u, v) = L_0(u, v) * e^{-k*d(u,v)} + L_s * (1 - e^{-k*d(u,v)})$
where

L(u,v) = apparent luminance, k = fog coefficient, d(u,v) = depth of image, $L_s(u, v)$ = luminance of the sky

So, if depth information is present, the fog can be artificially added to images. Based on value of k and L_s at different pixel locations, different types of fogs can be added to the image. Both homogeneous and heterogeneous fog were added along the lines described in [9].

1) Homogeneous Fog: Assumption is that fog Coefficient(k) and Luminance(L_s) of sky is independent of pixel location. As k increases, the fog density increases.

5

2) Heterogeneous Fog: To add more realistic fog in images, the assumption of constant k and L_s needs to be broken. In reality, K is not uniform but is a function of pixel location but it doesn't directly depend on depth of pixel. Thus, Fog Coefficient and Luminance of sky depend on pixel location. As the fog coefficient can't be completely arbitrary, sampling from a uniform distribution do not lead to realistic images. Perlin's Noise [10] is used to create spatially

Fig. 5: Image hazed with Perlin's Noise

Fig. 6: Image obtained after applying Dehazing with Dark Channel Prior on Figure 5

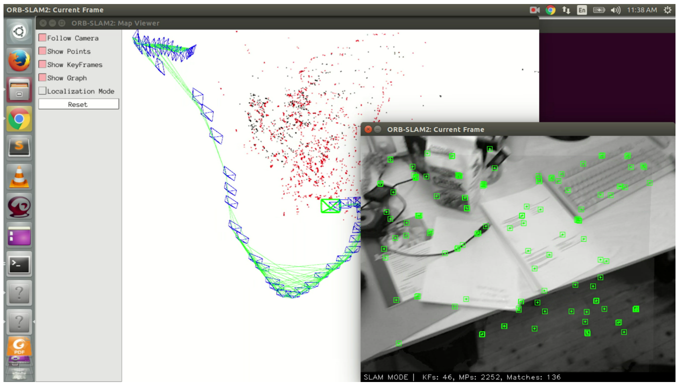

Fig. 7: Result of ORB SLAM on the dataset hazed with fog coefficient, k = 0.009

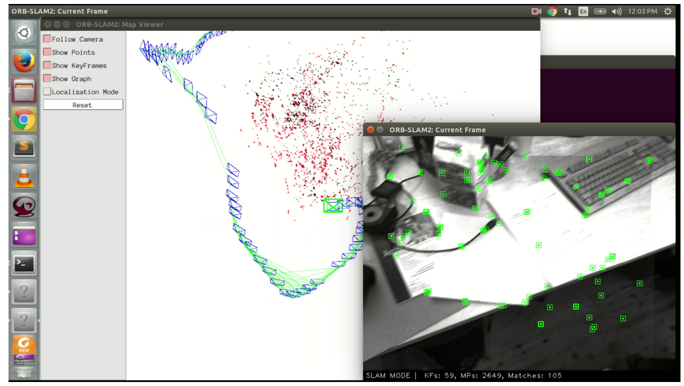

Fig. 8: Result using our pipeline on the dataset hazed with fog coefficient,k = 0.009

correlated noise. Perlin's noise has been used extensively to render fog and smoke in CGI applications. The idea is to add several noise functions with increasing frequency (spatial frequency) and decreasing amplitudes.

$Turbulence(pixel) = \sum_{i=1}^{N-1} noise_i(pixel)$

The result of application of Perlin's Noise can be seen in Figure

B. Dehazing results

The Figure 5 and Figure 6 show the results obtained on applying dark channel prior dehazing technique.

As we can see, after dehazing, the edges and features are more prominent in the image.

C. Final ORB SLAM results

On applying ORB SLAM technique on both the fogged as well as the dehazed images for different fog density levels, we obtained the 3D point cloud of the concerned region as well as the pose map of the camera as it moved through the region. We compared our slam results using the number of keyframes selected for the SLAM technique. The data corresponding to the number of features obtained along with their error estimates were not collected. The results are summarised in Figure 7,8,9,10,11 and Table I. We can see that with the increase in fog coefficient, the quality of SLAM(and no. of keyframes) on Hazed images decreases. With the application of proposed pipeline, the quality increases compared to using direct SLAM.

TABLE I: Comparison of Number of Keyframes

Value of K (Fog Coefficient)	Hazed	Dehazed
Original	66	.
0.09	55	68
0.012	20	71
0.015	16	73

V. Conclusion and Discussions

As observed in the results we were able to improve our SLAM results considerably using the proposed pipeline. However the pipeline failed at very high smoke density levels. As future scope of the work we propose the following:

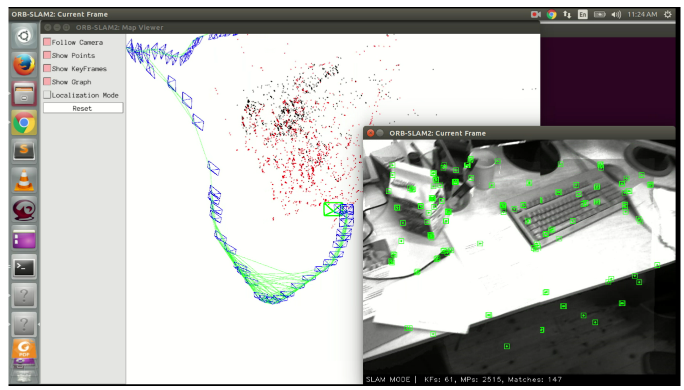

Fig. 9: Result using our pipeline on the dataset hazed with fog coefficient,k = 0.012

- To test on real fog dataset.
- Effects of Sensor fusion, in addition to dehazing pipeline, should be experimented.
- To test on outdoor dataset.

The problem is one of the biggest unsolved problems and promises much scope for improvement.

References

[1] Jorge Sales et. al (2010) Multi-Sensor Person Following in Low-Visibility Scenarios. Sensors
[2] JPM dos Santos (2013) SmokeNav-simultaneous localization and mapping in reduced visibility scenarios.
[3] Christopher Brunner et. al(2012) Automatic Selection of Sensor Modality for Resilient Localisation in Low Visibility Conditions
[4] R. Fattal(2008) Single Image Dehazing.
[5] He et. al(2010) Single image haze removal using dark channel prior
[6] Aditya Agarwal et. al(2014) Visual Odometry in Smoke Occluded Environments.
[7] Raul Mur-Artal et. al(2015) ORB-SLAM: a Versatile and Accurate Monocular SLAM System
[8] J. Sturm et. al (2012) A Benchmark for the Evaluation of RGB-D SLAM Systems
[9] J. P. Tarel et. al (2010) Improved Visibility of Road Scene Images under Heterogeneous Fog
[10] Ken Perline (1985) An Image Synthesizer

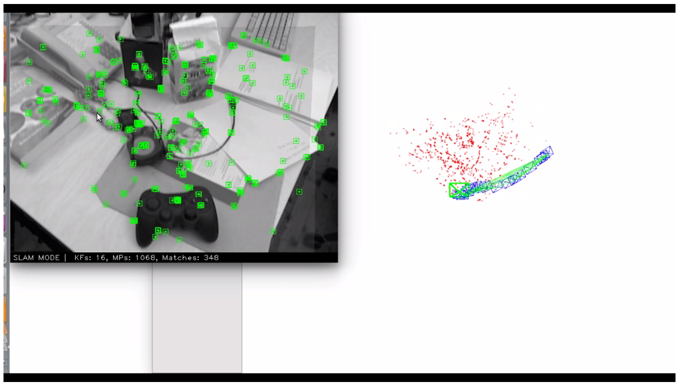

Fig. 10: Result of ORB SLAM on the dataset hazed with fog coefficient, k = 0.015

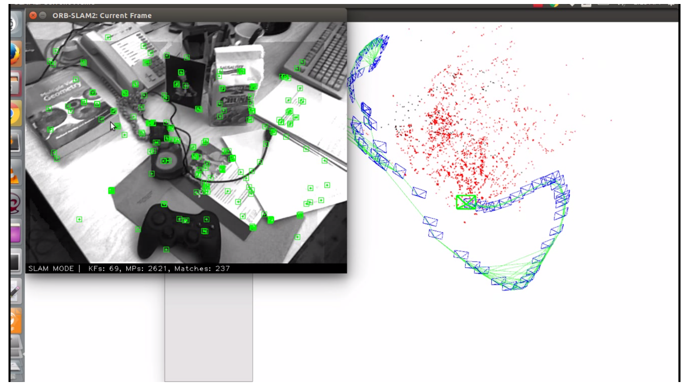

Fig. 11: Result using our pipeline on the dataset hazed with fog coefficient,k = 0.015