Master CORO M1
Master Commande et Robotique
Master in Control and Robotics

EMARO+ M1
European Master on Advanced Robotics

Project Report
24.06.2020

Active Exposure Control For Robust Visual Odometry

Abigail POP
Gowri UMESH

Supervisor

Vincent Frémont
Professor , Ecole Centrale de Nantes
Member, ARMEN team @ LS2N Lab,UMR CNRS 6004

ABSTRACT

Cameras when used as a perception sensor in robotics to perform for example Visual Odometry,they remain very sensible to illumination conditions. However, modern cameras have a specific feature called High Dynamic Range (HDR) Imaging using which it is possible to dynamically control the camera parameters such as the exposure, the shutter time and the gain to improve the image quality in adverse illumination conditions. The objective of the project is implement the exposure control by maximizing a robust gradient-based image quality metric to improve robustness of visual odometry in challenging illumination conditions using HDR Images.

CONTENTS

I. INTRODUCTION

HDR environments challenge the robustness of VO algorithms. A camera's dynamic range is narrower than the real world environment's and the exposure time has to be adjusted to obtain the best image, failing which, it is very easy for the image to be either underexposed or overexposed. HDR environments have a wide range of luminosity, standard imaging techniques have a certain brightness limit after which there is no differentiation in the brightness level, it either appears as completely white (overexposed) or completely black (underexposed) . Therefore a proper exposure control algorithm has to be implemented which results in properly exposed images which later can be used by an active VO algorithm. However, change in exposure time between two frames breaks the brightness constancy assumption of the VO algorithms and has to be compensated.

HDR imaging widens the range of luminosity and helps in improving the quality of the images and thereby increasing the information that could be captured from those Images. In the project , this quality of images are evaluated using gradient based image quality metrics.

Active exposure control is adapted by implementing the algorithm that maximizes the gradient information of the image. Optimisation is achieved by exploiting the camera's photometric response function which is detailed later in this report. During the implementation different gradient based metrics are analysed and the best one among them is used to optimise and update the exposure time. We use Percentile Metrics (M_{perc}) and soft percentile metrics ($M_{softperc}$) for the analysis.

HDR Environment and Images

Dynamic range refers to the difference in brightness between the brightest and darkest point in the photo. The higher this value, the more information the photo can store. Both HDR (High Dynamic Range) and HDRI (High Dynamic Range Imaging) are well-known expressions that describe this value of photographs. When taking a simple photo, it will most likely use SDR (Standard Dynamic Range) or LDR (Low Dynamic Range). By comparison, photos taken with HDR are much richer in detail due to technical specifications. However, nowadays, there are a lot of performant cameras who use the HDR function and we will only focus on data sets made with it.

In each pixel of an image, the irradiance E describes the amount of energy that hits the pixel per time unit, and the exposure X is the total amount of energy received by the pixel during the exposure time Δt. For an HDR image, Δt has a wider range which changes the exposure of an image. Along with exposure time Δt the other factor that the exposure of an image depends on is the gain or ISO. These two factors can be used to control the exposure.

An example of the terms described above can be seen in Figure 1, which represent a data set of HDR images at different time exposure. In this example we can see a few of the possibilities of Δt (they also depend on the camera which we use in order to create our data sets):

Figure 1 : Scene with very wide dynamic range
Image source (wikipedia)

Visual Odometry and VSLAM

Visual odometry is one of the best methods for estimating movement camera in the context of image processing applications, using images from one or more cameras. The major advantages that these methods bring are the accuracy of the estimate movement, computational efficiency and reducing the number of odometry devices that must be synchronized, necessary for an accurate estimate. This algorithms are widely used in application for mobile robots, that require advanced localization techniques.

Similar, SLAM is a process in which a robot is required to localize itself in an unknown environment and build a map of this environment at the same time without any information previously known and only with the help of external sensors.

The main difference between VO and SLAM is that VO mainly focuses on local consistency and aims to incrementally estimate the path of the camera/robot pose after pose, and possibly performing local optimization. Whereas SLAM aims to obtain a globally consistent estimate of the camera/robot trajectory and map.[5]

Our application we will not include the development of a VO or SLAM algorithm part, but we shall consider them as known.

II. DETAILED DESIGN

The proposed algorithm follows the flowchart shown in Figure 2.

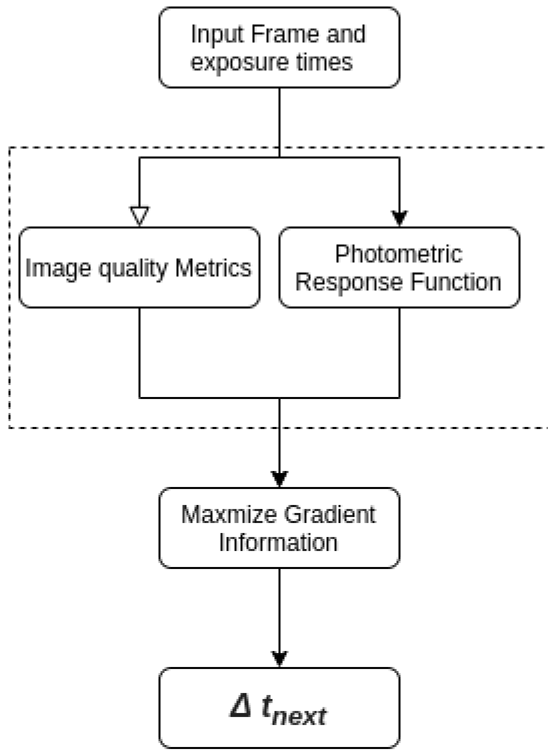

Figure 2 : Flowchart of the Algorithm

Input Image Dataset

This will be the first step of our project. In our case the data set must follow specific requirements such as: data set must be formed by HDR images, we should explore multiple exposure times and different angles of certain objects.

From the data set we will need to extract the exposure time (an information which should come with the data set as known, for each image as it was setted from the camera) and the intensity of pixels or average intensity of pixels which can be computed. This information will be the input of our algorithm. We consider 16 images from the same scene with different exposure times of 1 stop increment from 30sec to 1/1000secs.

Image Quality Metrics

In literature there has been usage of different quality metrics to evaluate the image quality and it depends on the application that images will be used in, in our case we build exposure control algorithm for robustness of VO and hence we decided to adopt gradient based Image quality metrics which informs us of the amount of information obtained from images

In our case we will compute and compare 2 quality metrics:

1. Percentile Metric (M_{perc})
2. Soft percentile Metric ($M_{softperc}$)

In computer vision a lot of applications take into consideration the gradient of an image, as they show us where major changes of pixel intensities appear in an image. However, this option it is not always the best in making quality assumptions about processed images; we will use its formula to further compute other quality metrics:

$$G(I, u, \Delta t)=|\nabla I(u, \Delta t)|^{2} \tag{1}$$

One of the metrics used for image processing is simply computing the sum of all gradients in the image. However, as previously said, this option might not be the best one in result when it comes to comparing the quality, and as it was tested in other papers, it does not give a robust estimator.

Instead, for our application we will use the following metrics:

1. Percentile metric is a quality metric that uses the gradient introduced above and it is meant to indicate the location of a score in a distribution. It can be described as a value below which a given percentage of measurements in a set falls. The formula we have used in our code can be seen below:

$$M_{perc}\ (p) = percentile(\{G(u_i\ )\}\ u_i \in I\ ,\ p) \qquad (2)$$

where p is the percentage of pixels whose gradient magnitudes are smaller than Mperc.

2. Soft percentile metric is the second quality metric we will compute for our application, and just like the percentile metric will include the computation of the gradients in the pictures. The difference between the two metrics is that here we will arrange the gradients in ascending order and the formula will imply weighting each of these, as we can see in the formulas below:

$$M_{softperc}\ (p) = \sum_{(o,S)} W_{\ ith}\ (p) \cdot G\ ith \qquad \textbf{(3)}$$

where S is total number of pixels in each image and weights are computed with the following formula:

$$W_{\text{ith}} = \begin{cases} \frac{1}{N} \sin\left(\frac{\pi}{2\lfloor p \cdot S \rfloor} i\right)^k, & i \leq \lfloor p \cdot S \rfloor \\ \frac{1}{N} \sin\left(\frac{\pi}{2} - \frac{\pi}{2} \frac{i - \lfloor p \cdot S \rfloor}{S - \lfloor p \cdot S \rfloor}\right)^k, & i > \lfloor p \cdot S \rfloor \end{cases} \qquad \textbf{(4)}$$

The results we get from the 2 metrics will be presented further in the report.

Photometric Response Function

In this work, we use the photometric response function proposed below . The image acquisition process is illustrated in Fig. 3. The photometric response function f maps the exposure X to the intensity I in the image:

$$I\ = f(X)\ = \ f(E\Delta t) \qquad (5)$$

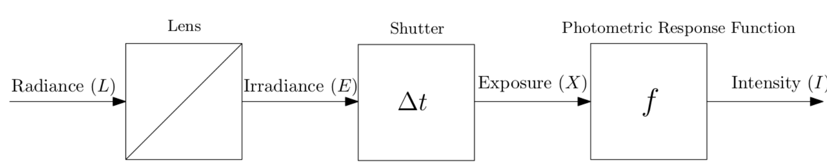

Figure 3 . Image Acquisition Process

Note that f (·) is invertible because the intensity should increase monotonically with the exposure. Then, for convenience, we can define the inverse response function

$$g = \ln f^{-1} \tag{6}$$

$$g(I) = \ln E + \ln \Delta t \tag{7}$$

Derivative of Gradient Magnitude

Since soft percentile proves to be the best among the considered gradient metrics we proceed with the use of the same. Because our metric is based on the image gradient magnitude, the first step is to calculate the derivative of the squared gradient magnitude G(·) with respect to the exposure time Δt.

$$2\nabla \mathbf{I}(\mathbf{u}, \Delta t)^{\top} \frac{\partial}{\partial \Delta t}[\nabla \mathbf{I}(\mathbf{u}, \Delta t)]. \tag{8}$$

The first term of the above equation is the Image gradient itself and the second term can be resolved using Schwarz's theorem as below:

$$\frac{\partial}{\partial \Delta t}[\nabla \mathbf{I}(\mathbf{u}, \Delta t)] = \nabla[\frac{\partial}{\partial \Delta t}\mathbf{I}(\mathbf{u}, \Delta t)]. \tag{9}$$

Results are obtained by further solving the equations the below:

$$\frac{\partial I}{\partial \Delta t} \overset{(1)}{=} f'[f^{-1}(I)]E(\mathbf{u}) = \frac{E(\mathbf{u})}{[f^{-1}]'(I)} \overset{(2)}{=} \frac{1}{g'(I)\Delta t} \tag{10}$$

$$\frac{\partial G(\cdot)}{\partial \Delta t} = 2[\nabla \mathbf{I}(\cdot)]^{\top} \nabla \left[\frac{1}{g'(\mathbf{I}(\cdot))\Delta t}\right] \tag{11}$$

In the subsequent steps the derivative of Metrics is used which is nothing but the weighted sum of the derivative of the gradient magnitude that is mentioned in the below equation.

$$\frac{\partial M_{\text{softperc}}}{\partial \Delta t} = \sum_{i \in [0,S]} W_{i\text{th}} \frac{\partial G_{i\text{th}}}{\partial \Delta t} \tag{12}$$

Exposure Control

We have shown that the soft percentile metric $M_{softperc}$ is a robust indicator of the image quality.Therefore, the goal of our exposure control is to maximize $M_{softperc}$ for future images. To achieve this goal, the exposure time is updated based on the latest image from the camera driver in a gradient ascent manner. In particular, given an image I and the corresponding exposure time Δt, the desired exposure time for the next image is calculated as

$$\Delta t_{\text{next}} = \Delta t + \gamma \frac{\partial M_{\text{softperc}}}{\partial \Delta t} \tag{14}$$

III. IMPLEMENTATION

Image Acquisition

Data set

The first step of our project, as mentioned above, will be the choosing of a data set. This was chosen according to the requirements it needed to fulfill for our application:

- Being a HDR data set
- Containing multiple (known) exposure times
- Having multiple perspectives

We have searched for a data set with known exposure times, as we will need these values to compute the photometric response function, as it will be explained below. The Data set has an frame of same scene with 16 different exposure time with 1 increment from 30sec to 1/1000 sec:

vector<double> exposure_times= { 30, 15, 8, 4, 2, 1, 1/2, 1/4. 1/8, 1/15, 1/30, 1/60, 1/152, 1/250, 1/500, 1/1000};

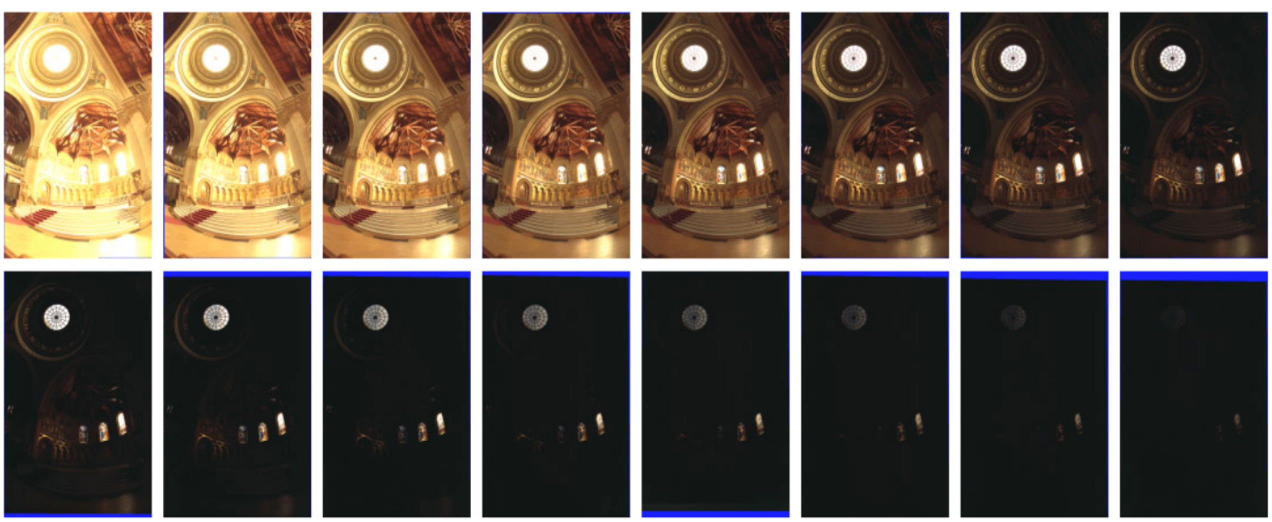

Figure 4: Sixteen photographs of a same scene taken at 1-stop increments from 30 sec to 1/1000sec

Exposure time

In photography theory we have three exposure control tools: exposure time, aperture and ISO sensitivity. Of these, the easiest to understand and most intuitive is the exposure time, which we will also be the only one we will modify for our exposure control. This parameter simply refers to the length of time the shutter is open and the sensor / film is exposed. This is usually expressed in fractions of a second, as occasions where times of more than one second are required are very rare. The exposure time is usually automatically or manually set from inside the camera, and our application proposes an algorithm to modify this value in order to improve the gradient-based quality metrics.

Pixel intensity

Another parameter we will need to use in the computation of the photometric response is the pixel intensity. This value is the primary information stored within pixels, it is the most popular and important feature used for classification in image processing techniques. The intensity for each pixel will have take a value from the gray-scale shown in below figure 5:

Figure 5 : Gray-Scale

Computing the intensity of pixels in our case is done by accessing the size of each image in our dataset and turning each one of them into a grayscale image:

```
cvtColor(frame,Gray_frame,CV_BGR2GRAY)
```

By using the below openCV function, we can directly compute the intensity of every pixel in an image and sequently, the average intensity for every picture:

```
float intensity = (float)Gray_frame.at<uchar>(i,j)
```

Photometric Response Function

We will make use of the inverse photometric function to check the optimization, as it represents the graph of the exposure times. This is directly computed with openCV class, where we will need to introduce the img_list (data set) and the time exposure for every picture:

```
Mat response;
    Ptr<CalibrateDebevec> calibrate = createCalibrateDebevec();
    calibrate->process(images, response, times);
```

Inverse camera response function is extracted for each brightness value by minimizing an objective function as a linear system. Objective function is constructed using pixel values on the same position in all images, extra term is added to make the result smoother.

Image quality Metrics Analysis

The 2 gradient based metrics require image gradient magnitude which is obtained by openCV's inbuilt *sobel* derivative function and *cv::magnitude()* respectively. Each frame of the same scene with different exposure times Gradient based metrics is calculated and the frame with highest gradient magnitude is selected as the best image and its exposure time is retained for further optimisation.

```
Sobel(src_gray, grad_x, ddepth, 1, 0);
Sobel(src_gray, grad_y, ddepth, 0, 1)
```

To verify the performance of the Metrics Fast Features are extracted from every frame and then it was verified that the frame with highest metrics value has the most number of features . The results of one such scene is shown in Table 1.

Frame no	$\mathbf{M_{perc}}$	$\mathbf{M_{softperc}}$	**Fast Features**
1	30	57	3932
2	**38**	**66**	**4534**
3	34	60	4141
4	27	51	3386
5	20	41	2612
6	13	31	1833
7	8	22	1245
8	4	16	914
9	3	12	717
10	3	9	599
11	2	6	509
12	1	5	395
13	2	4	308
14	1	2	202
15	1	2	96
16	1	2	29

Table 1 : Metrics and Fast features of a frame

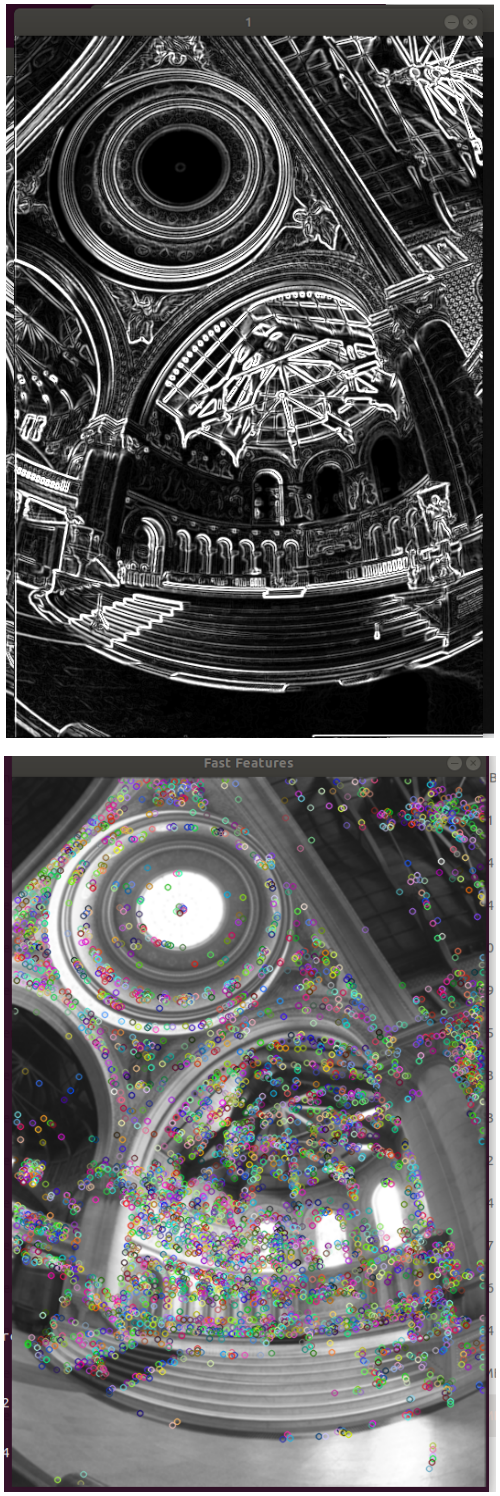

Figure 6: Gradient of the Best image and its corresponding Features

Derivative of Gradient magnitude of an Image

The best image is selected from the image quality metrics used to calculate the derivative of the gradient magnitude and the corresponding exposure time. Equation (8) and (9) are calculated using inbuilt openCV's operation on arrays . Once the gradient magnitude's derivative is obtained the result is used to furter calculate the derivative of the Gradient metrics $M_{softperc}$ mentioned in equation 10.

Updating the exposure time of next frame

Since we have gradient based metrics the optimisation is used in a gradient accent manner which maximizes the Metrics of the next frame and provides corresponding exposure time that has to be set for the next frame to get the best image possible.

This result is verified by checking the Metrics of the next frame for all possible exposure times and it was noted that the updated exposure time provides the best image.

IV. RESULTS AND CONCLUSION

In an HDR environment exposure time control plays a major role in obtaining the properly exposed image without losing much information, exposure control is a highly researched topic but when it comes to using a camera as a perception sensor in robotics, the application specific research are very few and the algorithm we have implemented can further be utilised in Visual odometry with a slight change in the algorithm's brightness constancy assumption.

By testing different approaches we have reached multiple conclusions in choosing the photometric function or quality metrics as explained in the chapters below, but the main idea of the application will be that active exposure algorithms will improve the quality of the our data sets.

Results obtained at each stage were verified with an alternative method proving the robustness of the algorithm. However, there can be further extension of the algorithm by implementing it on a camera and using it in visual odometry which will make the proposed algorithm more application specific and also improves the robustness of the VO algorithm.

V. References

1. “Active Exposure Control for Robust Visual Odometry in HDR Environments ”, Zichao Zhang, Christian Forster, Davide Scaramuzza
2. P. E. Debevec and J. Malik, “Recovering high dynamic range radiance maps from photographs,” in ACM SIGGRAPH. ACM, 2008, p. 31.
3. COVIS course material , Vincent Fremont
4. https://opencv.org/
5. https://www.ceva-dsp.com/ourblog/how-are-visual-slam-and-lidar-used-in-robotic-navigation/
6. http://www.pauldebevec.com/Research/HDR/ - link for dataset used