Chapter 7

Adaptation decision criteria

In order to perform adaptation we need some metric to identify the need for adaptation and to make choices about suitable adaptation goals. This chapter will cover the metrics and the rules for adaptation regarding localization. **Add more parts as intro for this chapter**

7-1 Image Quality Metrics (IQM)

Capturing properly exposed images with low noise and high dynamic range is essential for vision based robotic systems. However cameras hardware limitations such as low dynamic range, low sensor sensitivity, built-in auto exposure are the reasons of failures for the vision based algorithms. Changing illumination is considered as a crucial problem in robot vision [36]. In uncontrolled environment scene radiance is unpredictable due to illumination change over time and space. Motion blur can occur in low illuminated environment due to long exposure and can result in noisy images due to high gain. Various methods are available in literature to overcome varying illumination of environment such as [43, 41, 42, 26]. However [43, 41, 42] rely on exposure control for varying illumination and [26] creates multiple maps for different illumination level and then detects the current illumination level along appropriate map and camera exposure time to overcome low lighting conditions. All these methods gives good results than built-in auto exposure of camera but it needs varying exposure camera which puts hardware limitations on these methods. Software methods can also be used for illumination change such as Histogram Equalization, Contrast Limited Adaptive Histogram Equalization, [29], [36] Add other methods from [29] such as PSO etc .

In order to improve the quality of visual localization it is necessary to quantify the accuracy of localization which at times is not possible because unavailability of ground truth or co-variance matrix. But image quality can be used to analyze the accuracy of the visual localization algorithms. The famous Image Quality Assessment (IQA) approaches available in the literature such as BLIND [40] and BRISQUE [32] gives scores for image quality but does not give any insights about the individual attributes of the image such as noise, gradient, brightness, etc. The most basic matrix used is intensity histogram [48, 34] which is very fast approach but are dependent on environment illumination. Other matrices such as image gradients [42], entropy are robust against illumination changes but are sensitive to noise due to inappropriate gain. Hence level of noise is also need to quantify the image quality. Now we will define the image quality metric which can be used to estimate the accuracy of the visual localization algorithms in combination with co-variance matrices if available.

7-1-1 Histogram based metrics

Intensity histogram provides the information about various characteristics of the image such as brightness, contrast. If the histogram is concentrated in narrow region means the image has low contrast and broad histogram reflect image with significant contrast. The mean and standard deviation of the histogram gives the idea about the overall histogram distribution. Hence we will use intensity histogram for brightness and contrast metric.

The RGB image can be converted to Gray scale image(Y) which denotes the intensity of the pixels of the image using the conversion shown below.

$$Y = 0.299 * R + 0.587 * G + 0.114 * B \tag{7-1}$$

Gray image consists of intensity pixels Y and histogram of the Gray image give overall intensity distribution of an image. Using the gray image histogram we can define the brightness and contrast metric of the image.

Definition 1: The brightness metric L_{bright} of the image is the mean values of the pixel distribution of gray scales which is given by

$$L_{bright} = \mu = \sum_{i=1}^{N} i * P(i) \tag{7-2}$$

where *P(i)* is the probability of gray value and N is the number of possible gray values in the histogram (256 for 8 bit image)

Definition 2: The contrast metric $L_{contrast}$ of the image is the standard distribution of the intensity pixel distribution of the gray scale image which is given by

$$L_{contrast} = \sigma = \sqrt{\sum_{i=1}^{N} (i - \mu)^2 P(i)}$$

where *P(i)* is the probability of gray value and N is the number of possible gray values in the histogram (256 for 8 bit image), μ is the mean intensity of gray scale image

The intensity histogram varies with the change in the illumination of the surrounding which is captured by brightness and contrast metric.

7-1-2 Gradient based metric

The image gradient is the attribute related to the fine details, texture and edges of the image and can be used to extract information from images. Various image features such as

Histogram of oriented gradients (HOG) [13], Scale-invariant feature transform (SIFT) [31], edge detection, object detection are computed using image gradient information. The purpose of the gradient metric is to quantify the gradient information available in the image. The nonlinear gradient mapping function shown in Eq.(7-7) adopted from the Shim et al. [41] relates the gradient magnitude and the amount of gradient information in the image. The activation threshold δ prevents the affect of noise on the mapping function as the mapping function is only activated if the gradient value is above activation threshold. The logarithmic mapping function adjusts the difference between strong and weak gradients and λ is the control parameter to adjust the mapping information. The image gradient of intensity image (Y) is calculated using Sobel operator [45]. The operator uses *3x3* kernel which is convoluted over the image to calculate the approximate derivative. The horizontal and vertical changes are calculated separately and are merged.

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * I \qquad G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * I \tag{7-3}$$

where * denotes 2-dimensional convolution operation, G_x, G_y represent the image gradient in horizontal and vertical direction respectively.

The overall gradient magnitude of the intensity image Y is given by:

$$G = \sqrt{G_x^2 + G_y^2} \tag{7-4}$$

The mapping function relates the gradient magnitude to the gradient information.

$$\tilde{g_i} = \begin{cases} \frac{1}{N_g}(\lambda(g_i - \delta) + 1), & \text{for} g_i \geq \delta \\ 0, & \text{for} g_i < \delta \end{cases} \tag{7-5}$$

$$\text{s.t. } N_g = log(\lambda(1 - \delta) + 1),$$

where $g_i \in [0, 1]^1$ which is the gradient magnitude at pixel i, δ is the activation threshold which is used to filter out noise, λ is the control parameter to adjust the mapping behaviour, N_g is the normalization factor and $\tilde{g_i}$ is the of gradient magnitude of pixel i. However this mapping function is strongly biased to the gradients caused by the high exposure [43]. In order to remove the bias the image is divided into grids and the gradient information in each grid is aggregated. This modification in the mapping function reduces the bias caused due to non-uniform distribution. The gradient quality metric is given by

$$L_{gradient} = E(G) \tag{7-6}$$

where $E(G)$ is the mean value of the distribution of the gradient information in each grid cells G_j which is given by:

$$G_j = \sum_{i \in C_j} \frac{\tilde{g_i}}{Wj * Hj}, \qquad j = 1, 2, \ldots, N_c \tag{7-7}$$

where G_j is the j^{th} grid cell, N_c is the total number of grid cells, Wj, Hj are the width and height of the j^{th} grid and $\tilde{g_i}$ is the amount of gradient information. Large $L_{gradient}$ indicates that the gradient information is strong and uniformly distributed and low $L_{gradient}$ indicates weak and biased gradient.

7-1-3 Entropy based metric

The gradient based metric is calculated from the gradient of the intensity image, however global image entropy represents the amount of information in a image. Using global image entropy as a quality metric in combination with gradient based metric gives the overall amount of information contained in the image and its gradient. The entropy based image quality metric is given by

$$L_{entropy} = -K_e \sum_{i=0}^{N-1} P(i) log_2 P(i) \tag{7-8}$$

where $P(i)$ is the probability of pixel value i in the gray scale image, K_e is the normalization factor and N is the total number of gray scale values (255 for 8 bit image)

7-1-4 Noise based metric

As the none of the above metric captures explicitly the noise in the image we will use two noise based image quality metric. There are various methods to estimate the noise in the image such as Peak signal-to-noise ratio (PSNR), Feature similarity index (FSIM) which compares the distorted images with original image. However in our case there is no reference image to compare with hence we will use the noise metric without using reference image. Also the methods proposed in [11] based on eigen value analysis and [39] based on principal component analysis give more accurate results they are computational expensive. Hence we use method proposed in [46] which used a filter based approach which is computational fast. One metric estimates the noise using laplacian operator proposed by Immerkaer in [24] with some modifications of adaptive edge detection proposed in [46] to prevent over-estimation of the noise. Donoho [16] also proposed a estimation algorithm which is based on the median absolute deviation of the wavelet transform coefficients. We will take weighted sum of two methods for our noise based metric. For noise estimation we assume that the image has additive zero-mean Gaussian noise.

Immerkaer [24] proposed a noise estimation operator which mask the image using 3x3 kernel given by

$$M = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} \tag{7-9}$$

However the above laplician kernel is sensitive to image objects which will result of inaccurate noise estimates. To improve the accuracy the laplician filter should be applied only to unsaturated homogeneous regions. Hence we uses two masks one for unsaturated regions which is a simple threshold mask given by:

$$U(i) = \begin{cases} 1, & \text{for} \quad \tau_l \leq Y(i) \leq \tau_h \\ 0, & \text{for} \quad otherwise \end{cases}$$

where τ_l, τ_h are the lower limit and upper limit for unsaturated pixel, Y(i) is the intensity image.

The homogeneous region mask H which is a adaptive threshold mask is given by

$$H(i) = \begin{cases} 1, & \text{for} \quad g_i \leq \delta_p \\ 0, & \text{for} \quad g_i > \delta_p \end{cases}$$

where the adaptive threshold δ_p is the p-th percentile of gradients in the image, and g(i) is the gradient magnitude of the pixel i.

Therefore the fast noise variance estimator is

$$\sigma_{fastnoise} = \sqrt{\frac{\pi}{2 * Ns}} \sum_i H(i).U(i).|I * M|(i) \tag{7-10}$$

where N_S denotes the number of valid pixels in the mask H·U, * denotes the convolution operator and $|\cdot|$ denotes the absolute operator.

The robust wavelet-based estimator $\sigma_{wavelet}$ of the (Gaussian) noise standard deviation is obtained using $estimate\_sigma(image)$ function of skimage restoration library [5].

The image gradient is given by combination of two noise metric

$$L_{noise} = 0.5 * \sigma_{wavelet} + 0.5 * \sigma_{fastnoise} \tag{7-11}$$

The overall image quality metric is given by

$$\boxed{IQM = w1 * L_{brightness} + w2 * L_{contrast} + w3 * L_{gradient} + w4 * L_{entropy} - w5 * L_{noise}} \tag{7-12}$$

where w1,w2,w3,w4,w5 are the weights that adjusts the effects of each term which can be changed according to the requirements. In the IQM metric the noise metric is subtracted as it degrades the quality of the image.

The metric will be used to perform adaptations which can either be using hardware such as changing the exposure time of the camera, or turning on any light if available on the robot in case of low illumination or can be software methods such as improving quality by image processing techniques such as gamma correction, histogram equalization. And if none of technique is able to improve the image quality then the configuration can be changed to perform laser based localization.