

# A MATLAB-Based Ghost Detection Algorithm Development in the Context of Stray Light Test

Zhu Wang<sup>1</sup>, Hongyun Liu<sup>2</sup> *Corresponding author*, Siyi Xiong<sup>2</sup>, Jiaxin Dai<sup>3</sup>, Man Li<sup>1</sup>

1.School of Geography & Environmental Science Guizhou Normal University, Guiyang 550001, Guizhou China

2.Institute of Surveying and Mapping Guizhou Geology and Mineral Exploration Bureau, Guiyang 550001, Guizhou China

3.School of Business Macau University of Science and Technology, Macau 999078 China

---

**Abstract:** High dynamic range (HDR) video cameras with modern, non-linear CMOS image sensors can provide not only better visual expression, but also a new and powerful driving force for the development of advanced driver assistance systems (ADAS). However, mechanical defects, impurities, and imperfections in image sensors or lenses are inevitable in actual volume production. As a result, light reflections can cause disturbing ghost artifacts in received images. Effectively detecting these artifacts in dynamic scenes, camera jitter, inaccurate image registration, and large displacements of moving objects is a key problem. Therefore, a comprehensive understanding of the currently unknown characterization of ghost artifacts must be improved. Currently, there is no automated method based on image processing tools to identify ghost artifacts in HDR images obtained from camera systems. In this paper, an algorithm to detect ghost artifacts has been developed and evaluated for future automated manufacturing test systems using Matlab. The result of the first module of the algorithm indicates that not only the existence of ghost artifacts can be detected, but also the type of ghost can be determined in single testing images with an acceptable reliability of image processing methods. Furthermore, a typical feature of ghost artifacts can be extracted by identifying the significant portion of the ghost as a “Ghost Ring” from the image with normal ghost within a reasonable level of accuracy and success rate.

**Keywords:** Ghost detection ; Matlab ; Straight light test

---

## 1. Introduction

The distribution and intensity of stray light artifacts in acquired images can cause advanced driver assistance systems (ADAS) to mistakenly recognize and understand objects, potentially leading to risks or unpredictable consequences. As such, ghost detection in images and videos is a crucial step in stray light testing for ADAS. In the framework of developing an automated test system, extended image processing algorithms for classifying stray light artifacts must be developed and evaluated. While algorithms for Halo and Beam artifacts have already been developed, this study focuses on the detection of ghost artifacts with unknown features.

Several ghost detection algorithms have been proposed in the literature to automatically identify and remove ghosting artifacts in images and videos. One approach to ghost detection involves using Fourier analysis to identify and remove the ghost image. In an early study by Learner et al. (1996)<sup>[1]</sup>, a Fourier-based algorithm was proposed to detect ghosting artifacts in high dynamic range (HDR) images. The algorithm used the Fourier transform of the image to detect the ghosting artifact, and then removed it by selectively filtering the image. The results showed that the proposed algorithm was effective in removing ghosting artifacts from HDR images. Another approach to ghost detection involves using machine learning algorithms to automatically identify and classify ghost images in images and videos. In a study by Kyathanahally et al. (2018)<sup>[2]</sup>, a deep learning algorithm was proposed to detect ghosting artifacts in videos captured by a moving camera. The algorithm used a convolutional neural network (CNN) to classify the frames of the video as either containing ghosting artifacts or not. The results showed that the proposed algorithm achieved high accuracy in detecting ghosting artifacts in the video.

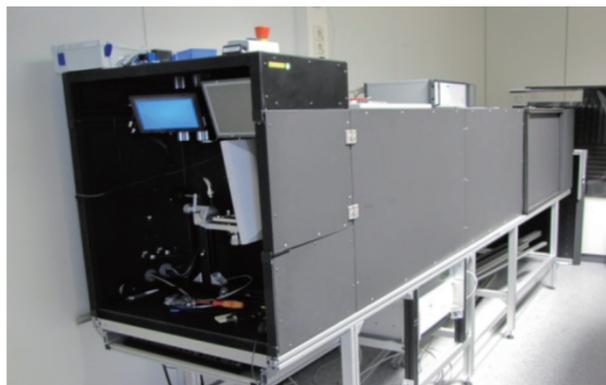
In addition to Fourier analysis and machine learning algorithms, MATLAB is a powerful software platform widely used to implement and evaluate various ghost detection algorithms. For example, in a study by H. F. An et al. (2014)<sup>[3]</sup> proposes an improved algo-

rithm to detect and remove ghosting artifacts in digital camera images. The algorithm is implemented in MATLAB and involves the following steps: 1) estimate the motion vector field using the hierarchical block matching algorithm, 2) identify the ghosting artifact by comparing the estimated motion vector field with a reference motion vector field, and 3) remove the ghosting artifact by replacing the pixels in the ghost region with the corresponding pixels from the previous or next frame.

Overall, ghost detection algorithms have shown great promise in detecting and removing ghosting artifacts in images and videos. While these algorithms have been shown to be effective, there is still ongoing research to further improve their performance and applicability in real-world scenarios. The main focus of this paper is to develop a simple algorithm for ghost extraction, calculation of ghost features and statistics, providing a basis for further classification of stray light artifacts. A classifier based on image processing is developed, which separates the test images into two categories - those with visible ghosts and those without. Secondly, for the category of images with visible ghosts, an algorithm primarily based on image processing is implemented and thoroughly tested to detect and extract the features of various types of ghost artifacts from images acquired by different camera samples. The higher brightness region of every ghost artifact is extracted as a part of a ring shape.

## 2. Data connection

The images used for the algorithm of ghost detection and feature extraction were acquired from testing HDR CMOS camera samples using special test device in laboratory. The HDR CMOS camera is one part of the 2<sup>nd</sup> Generation Video Multi Purpose Camera (MPC2) of Bosch GmbH. The appearance and the inside construction of the device used for stray light test in laboratory is shown in Figure 2.1. A cuboid box is about 3 meters long, and it is total dark inside. A halogen lamp with controllable switch is installed at the rear of the box to simulate the spectral characteristics of headlight. In the front face of the box a circle hole is left for the lens of the testing camera. The control device consists of a handle plat for device under test, which can turn the lens  $\pm 45$  degrees on a horizontal axis and on a vertical axis, two displays and a control center.



(a) test device



(b) internal structure

Figure 2.1 Test device (a) and internal structure (b)

The ghost artifacts inspection test was implemented as a subjective test. The inspector sat ahead of the tunnel, moved the device by holding a flexible holder manually and watched the images on a display at the same time. The images represented on the display were live images from the testing HDR CMOS camera with 25 fps, which are updated every 40ms, and the live images are 12bit, managed as unit16 data format. For visualization on LCD displays the collected data is converted to the 8bit range. In total, 118 acquired images were used in the algorithm designing procedure.

## 3. Algorithm Design

The complex task of image processing can be broken down into a series of steps. Generally, the basic hierarchical framework should consist of image acquisition, image preprocessing, feature extraction (which partly includes segmentation), and classification. In actual manufacturing test approaches, the stray light artifacts test is comprised of several testing contents aside from ghost testing. Based on the common features of ghost artifacts in the whole test images, the algorithm was designed in three main sections using

Matlab. First, the algorithm automatically detects whether a visible ghost artifact exists in the image. If it does, the ghost region is segmented from the whole image as a sub-image used in the next step. Secondly, the algorithm extracts the higher intensity pixels of the ghost artifact in the shape of a part of a ring, which is called the “Ghost Ring”. Finally, the algorithm calculates the mean values of the “Ghost Ring” as ghost features, which are then integrated into the final stray light features set.

### 3.1 Ghost region detection

The brightness and size of ghost artifacts change depending on the position of the spot. Ghosts were classified into visible or invisible types, where the invisible ghosts are not significant in ghost feature extraction. Therefore, a crucial step before extracting the feature of the ghost is to detect if a visible ghost appears in the image.

#### 3.1.1 Basic Flowchart

The first step of the algorithm can be divided into three parts: preprocessing, ghost type judgment, and ghost region segmentation. In the preprocessing part, a round mask is built to cover the spot in the test image, followed by applying noise reduction, image smoothing, edge detection, and morphology processing on the masked image to acquire several connected components. In the second part, the ghost type is determined based on the distance between the spot and centroids of the connected components and the number of connected components larger than the defined area. Once the type is identified, the region with a normal type ghost is segmented as a sub-image.

#### 3.1.2 Image preprocessing

In an original image taken from the test camera, the only bright area besides the spot region is the ghost artifact. Based on this brightness distribution feature, the proposed algorithm adopts a spot-covering method as the first step of image preprocessing. It should be noted that the mask radius size, as well as other thresholds and filter mask size, are determined after repeated tests on the available test images to ensure the desired processing effects. These parameters are applied to all acquired images from the four test lenses.

First, the raw image is converted to a binary image using the “im2bw” function in Matlab. The resulting image, called “BW”, replaces all pixels in the original image with luminance greater than a certain level with the value 1 (white), and all other pixels with the value 0 (black). The level is determined by computing a global threshold using Otsu’s (1979)<sup>[4]</sup> method, which chooses the threshold that minimizes the intraclass variance of the black and white pixels.

Next, the exterior boundary of the spot is extracted using the ‘bwboundaries’ function. The center of the spot can be calculated mathematically by treating the spot boundary as a circle. Then, a round mask with a radius of 85 pixels is created to cover the spot and the surrounding stray light in the raw image. The mask must avoid covering the ghost artifact, and the gray value of the mask is set to the mean gray value of the image’s background region.

Once an image containing only the higher intensity ghost artifact has been obtained, the next step is to reduce the inherent noise. In order to emphasize the ghost artifact with respect to the background noise, a group of filters are applied to the spot-masking image after several rounds of testing with different filter combinations and parameters.

There are several sources of noise in digital images, including acquisition and digitization (e.g., lens, A/D converter, and image sensor), image transmission, and computation. However, since the test images are acquired from a night vision video camera, only two main types of inherent image noise are considered: salt & pepper noise and Gaussian noise.

A group of filters are employed to reduce noise. First, a median filter is used to reduce salt and pepper noise. Second, Gaussian filter with a standard deviation of 3.0 are applied twice to clean Gaussian noise. The resulting image is then blurred. An average filter of size 9 by 9 is used three times repeatedly to obtain a smoother image. Finally, a clearer ghost artifact with a blurred background can be obtained using smallest order filtering of size 9 by 9.

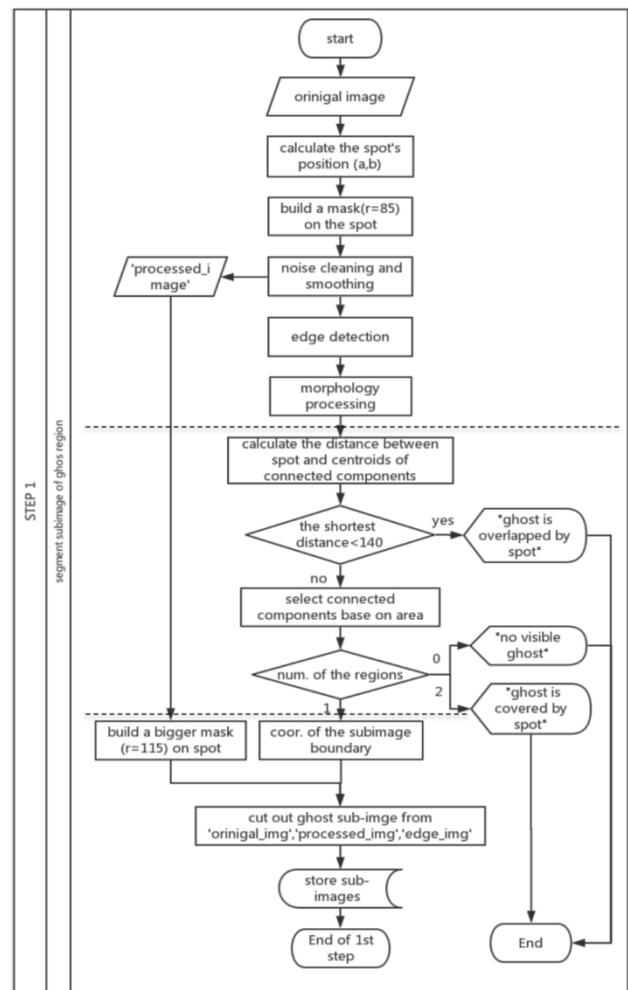


Figure 3.1 Flowchart of the ghost region detection

Here, an image with a typical ghost artifact of middle intensity is used as an example to illustrate the image processing steps (Figure 3.2).



Figure 3.2 An example image with typical ghost in middle intensity

The following Figure 3.3 express the image change processes through scaling image data to the gray value range of 800 and 1200 and displaying the image underlying Matlab function 'imagesc'.

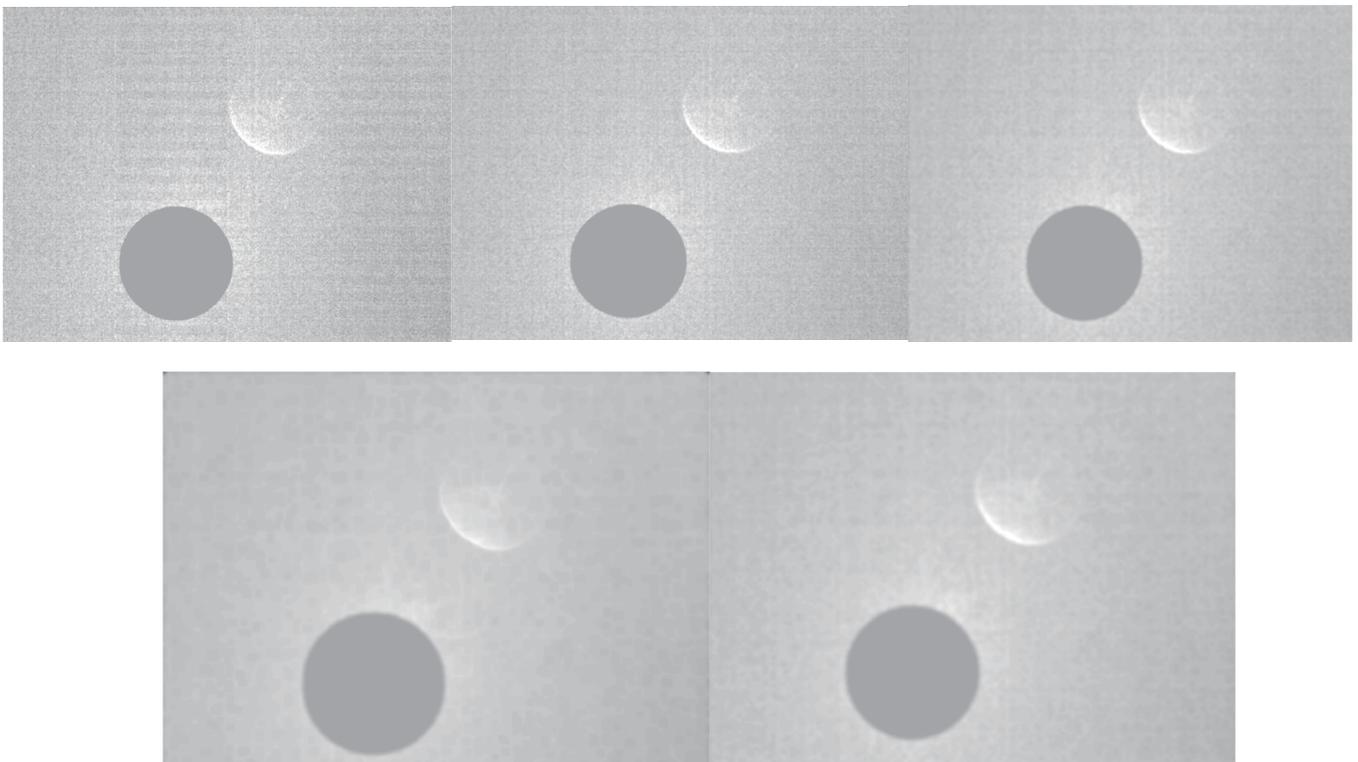


Figure 3.3 Image changes after applying the following filters in sequence: median filter, Gaussian filter 1 time, Gaussian filter 2 times, average filter 1 time, average filter 3 times

As there is no large gray value difference in the background region, a "Roberts" edge detection operator is used to detect the ghost

artifact. However, the simple “Roberts” operator cannot guarantee good continuity of edge detection. Therefore, a morphological processing technique called “majority” is used, which sets a pixel to 1 if five or more pixels in its 3-by-3 neighborhood are 1, to ensure completeness of the ghost edge (MathWorks, 2016). [5] Then, the individual white pixels (values are 1) near the edge and the mask circle are cleared. In the next step, morphological dilation is applied using a “disk” structure element with a radius of 6 pixels repeated 4 times. This is followed by a morphological close operator using the same structure element to acquire a connected ghost artifact region without holes. The result is shown in Figure 3.4.



Figure3.4 Image change after edge detection and morphology operations

### 3.1.3 Ghost Type Judgment

As shown in the above figure, besides the ghost component some other connected regions are generated simultaneously. The result of connected components is various under different ghost situations. Therefore, the ghost type judgment is based on the properties of connection components.

The type “ghost is overlapped by spot” is judged according to the distances between the spot center and the whole centroids of connected components. The preprocessing result of this kind of ghost is displayed in Figure 3.5. This ghost type is not obvious as the other independent ghost due to the impression of the stray light around the spot. So this case is not studied here, but it has to be detected by the program. Through multiple computation and analysis on testing images, the distance threshold is defined as 140 pixels. This means the ghost is identified as “overlap” type when the shortest distance is below 140 pixels. Otherwise, carry out following operations.

After removing the above situation, the next task is detecting the visibility of ghost. As the distribution and size of connected components of each image changes with different spot position. Based on repeated tests on part-covered ghosts and normal ghosts, the study on the relationship between spot position and connect component’s area of ghost provide the following results. According to the coordinates  $(a, b)$  of spot center  $P$ , two situations can be set:

$$P \begin{cases} \text{'middle', if } 550 \leq a \leq 750 \text{ and } 450 \leq b \leq 600; \\ \text{'around', otherwise;} \end{cases}$$

If the spot center belongs to the ‘middle’, connected regions with an area larger than 2500 pixels are detected; otherwise, connected regions with an area larger than 5000 pixels are detected. This method typically does not provide a suitable connected region for images with invisible ghost artifacts. For images with ghosts covered by spots, two connected regions will be detected. However, since the spot and other stray light cover the most significant region of the ghost, the rest of the detected ghost regions provide no useful information and are not the focus of this thesis. For the last case, images with normal ghosts, one satisfactory connected component will be detected. The following research focuses on this case.

For the first three cases shown in Figure 3.5, notice of “ghost is overlapped by spot” or “no visible ghost in this image” or “ghost is part covered by spot” will be displayed and the program will break at this point due to detecting if ghost artifact exists is the only task for these images. For the case of images within obvious and complete ghost, the features of the normal and typical ghosts, which are studied and processed in the next module of this algorithm.

### 3.1.4 Ghost Region Segmentation

The coordinate range of the connected component in the normal ghost image is detected. Next, the sub-image of the ghost region in rectangular shape is segmented from both the “original image” and the “processed image”. The “processed image” contains the spot mask with a radius of 115 pixels, which has been determined after multiple tests. During the preprocessing stage, the radius of the spot mask has to be carefully chosen to ensure that it covers as much spot stray light as possible while avoiding covering

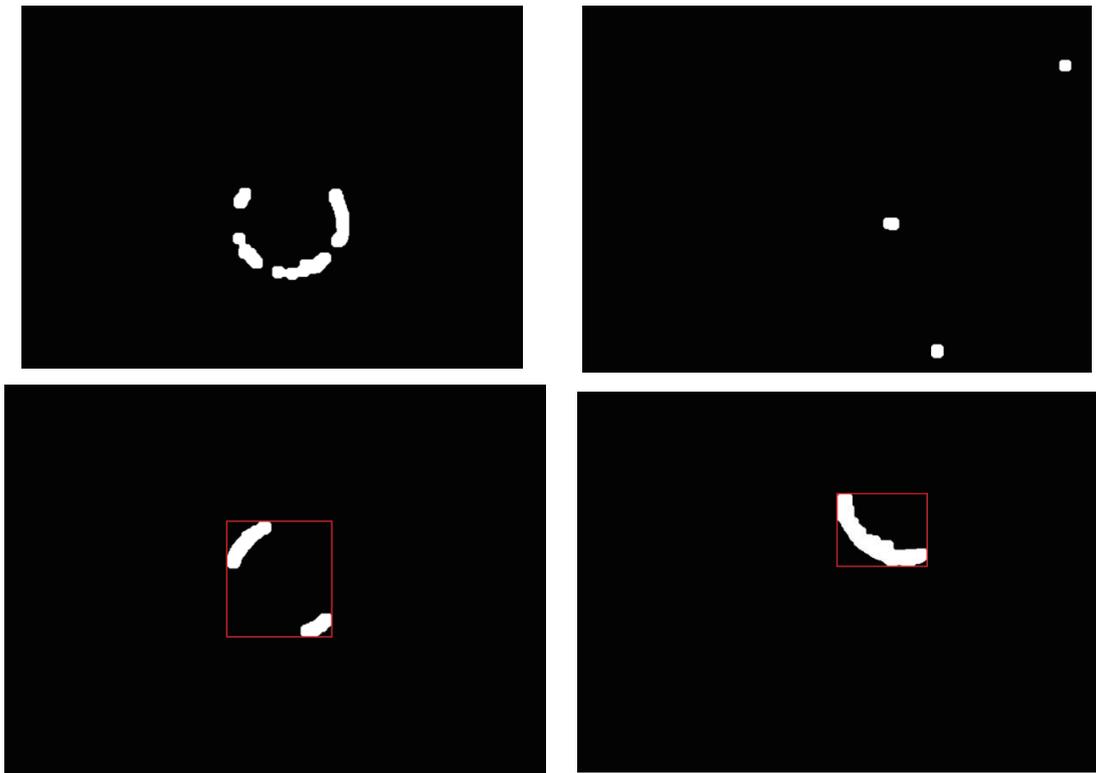


Figure 3.5 Connected components of ghost in different types: ghost overlapped by spot (top-left); invisible ghost situation (top-right); ghost part-covered by spot (bottom-left); normal ghost (bottom-right)

the ghost. The radius is set to 85 pixels to account for image smoothing and the detection of other ghost types. However, to ensure more accurate ghost curve fitting, a bigger spot mask has to be applied to cover the disturbed stray light around the spot that might affect the accuracy of the curve fitting. The width of the rectangular sub-image is the same as the coordinate range of the connected component in the horizontal direction, and the length is determined accordingly. The image segmentation of the ghost region simplifies the image content for “Ghost Ring” extraction and provides a basis for the detection of ghost features. The sub-image of the ghost region for the typical example is shown in Figure 3.6.

### 3.2 Ghost Ring Detection

In the first module, the ghost region has been segmented from the testing image. After analyzing and observing the ghost artifacts, a region named the “Ghost Ring,” which contains pixels with higher intensity, is chosen to represent the ghost and compute its features. This section explains the procedures for calculating and extracting the “Ghost Ring.”

As the shape of ghost can be treat as circle, the idea of this module is extracting the ghost feature according to the shape characteristic. Firstly, extract some points inside the high intensity part of ghost to roughly represent the shape of ghost. And then build a ring that contains the significant part of ghost based on circle fitting using these points. Because the significant part of ghost just takes a quarter of the entire ring approximately, this region has to be extracted to supply a basis for ghost feature computation. In order to detect the significant part of ghost, a special point that in the middle of the significant part is designed to extract first. With

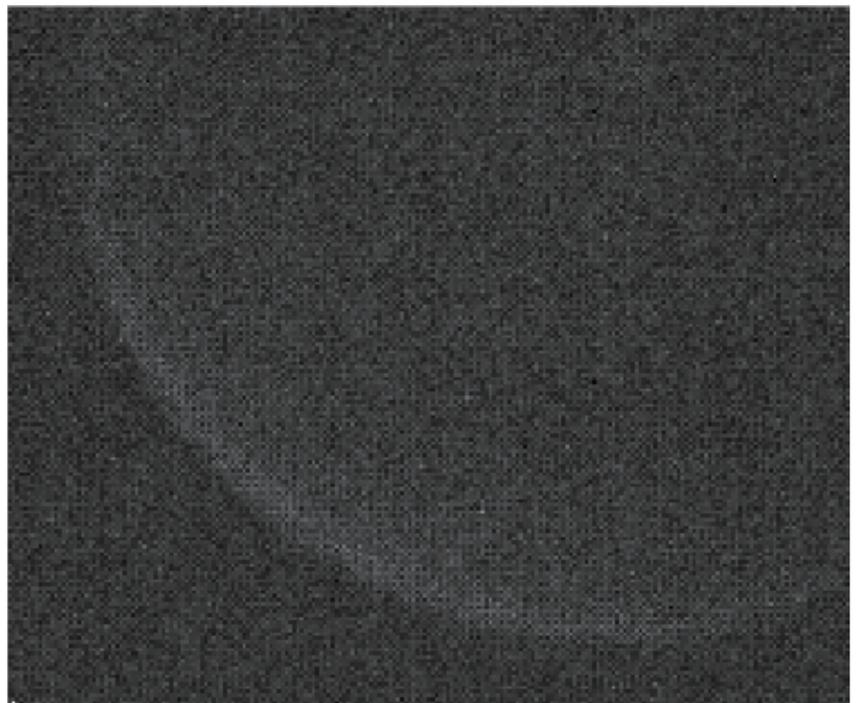


Figure 3.6 Sub-image of Ghost region after segmentation

the extension from the point towards both sides along the ring, the suitable quarter of ring is formed. This quarter part of ring is named “Ghost Ring” in this thesis.

The specific methods of this module will be elaborated in following three aspects (as shown in flowchart): ghost middle curve extraction and circle fitting; ghost feature point computation; building method of “Ghost Ring”.

### 3.2.1 Ghost Middle Curve Extraction and Circle Fitting by Least Square Method

In this module, the data used in the following steps are the “processed\_image” sub-image within a bigger spot mask and a sub-image from the original image. The sub-image of the ghost region from the original image does not contain the spot in very high light intensity, which requires 16bit to store information with a large range. In addition, 12bit is enough to store the information of the ghost artifact and the background. Therefore, it is better to translate this sub-image from 16bit to 12bit first in order to display the results according to human vision. The following processes are based on the sub-image from the “processed-image”.

To find the ghost curve inside the ghost artifact, we need to first find the lightest pixel of the ghost in each column (or row). Then, we sort the pixels of each column based on their gray values and take the first five pixels to calculate their mean position. The number of pixels considered, which is five, is determined through repeated testing. The pixel of the mean position can be considered as the approximate lightest pixel. Using the same method, we can find the lightest pixel of each column and compute the ghost curve. Circle fitting based on the middle part of the ghost curve is then applied.

However, the points located in the first and last parts of the curve may not be reliable and stable due to the influence of spot stray light, which is unavoidable in the sub-images. Therefore, only the middle three-fifths of the point data will be used in the next step of curve fitting. The shape of the ghost artifact is very similar to a circle, and even when the ghost is near the edges of the image, it will be closer to an ellipse, where the light part can also be considered as a part of the circle.

As the point position of ghost middle curve  $(x,y)$  distribute in one side of a circle-shape on the xy-plane, least squares regression is used to find the equation that best fits the data and to obtain more accurate “ghost ring” in further. The fitting method is given by following equations according to Chernov et al. (2005)<sup>[6]</sup> The general equation of circle is:

$$(x - a)^2 + (y - b)^2 = r^2 \quad (3.2)$$

Where  $r$  is the radius of the fitting circle, and  $(a,b)$  is the center. These three parameters have to be determined such that the curve given by (3.2) provides a good fit around the data points.

With least squares, “best fit” means that minimize the following equation by solving the system  $\partial F/\partial a = 0$ ,  $\partial F/\partial b = 0$ , and  $\partial F/\partial r = 0$ .

$$F(a,b,r) = \sum_{i=1}^m [(x_i - a)^2 + (y_i - b)^2 - r^2]^2 \quad (3.3)$$

Here,  $m$  is the number of points data used in the circle fitting computation. The equation of the circle can be linearized as follows:

$$x^2 + y^2 = 2ax + 2by + r^2 - a^2 + b^2 \quad (3.4)$$

$$x^2 + y^2 = Ax + By + C \quad (3.5)$$

This equation is now linear in the undetermined coefficients A, B, and C, which means the least squares problem can be solved

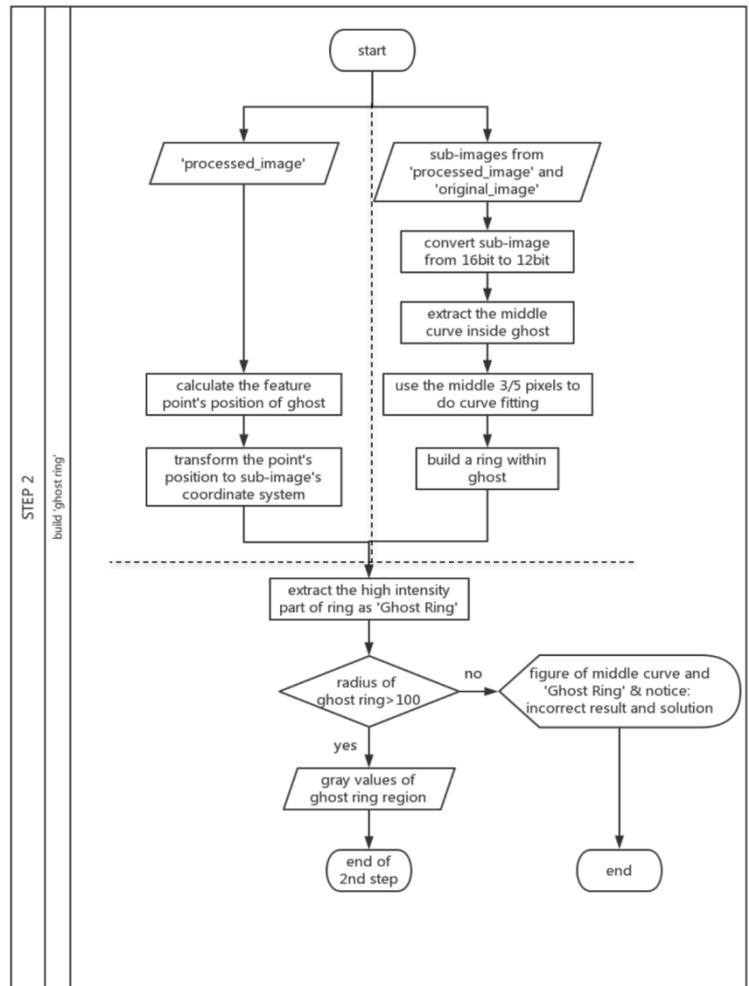


Figure 3.7 Flowchart of the 2nd part of algorithm: “Ghost Ring” extraction

with matrices. The value of a, b and r can be computed backward after getting A, B and C. Hence, the next step is finding A, B, and C with matrices.

The matrix equation for circular regression is given by:

$$\begin{bmatrix} \sum x_i^2 & \sum x_i y_i & \sum x_i \\ \sum x_i y_i & \sum y_i^2 & \sum y_i \\ \sum x_i & \sum y_i & n \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} \sum x_i(x_i^2 + y_i^2) \\ \sum y_i(x_i^2 + y_i^2) \\ \sum x_i^2 + y_i^2 \end{bmatrix} \quad (3.6)$$

If the 3-by-3 matrix on the left is invertible, then there is a unique set of values for A, B, and C that generates the circle of best fit. Once know the values for A, B, and C, a, b and r can be obtained with these transformations:

$$a = -A/2 \quad (3.7)$$

$$b = -B/2 \quad (3.8)$$

$$r = \frac{\sqrt{4C+A^2+B^2}}{2} \quad (3.9)$$

### 3.2.2 Ghost Feature Point Extraction

The basis of the build of “Ghost Ring” is finding the middle point of ghost artifact, which is also the middle of “Ghost Ring”, and this point can be also considered as the nearest pixel of the ghost to the spot center.

In Figure 3.8, the feature point is marked as a blue “\*“, the spot center is marked as a blue “o“, and the symmetry of the ghost artifact is shown by the red and yellow lines. The feature point is obtained by approximate calculation: 1, Sort the pixels in the ghost connected component by the distances between these pixels and spot center in ascending order. 2, Compute the coordinates of the middle pixel (feature point) of the ghost artifact from the first 1000 pixels by averaging their coordinates in column and in row separately. 3, Connect the middle pixel and the spot center, the yellow line in the figure can be obtained, and the red line is the perpendicular of the yellow line. Then the symmetry of the ghost artifact is obvious in the figure, and the feature point can serve as the middle point of the ghost.

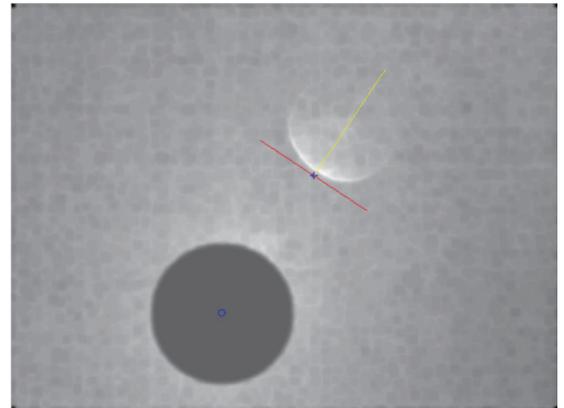


Figure 3.8 Ghost feature point and geometrical relationship in the image

### 3.3 Ghost feature Calculation

#### 3.3.1 Basic Flowchart

As illustrated in the flowchart (Figure 3.9), the primary function of this module is to calculate the features of the ghost artifact to provide data support for final analysis. It is an efficient method to understand the ghost stray light performance in the filter sub-sampling mode based on the special property of the “CRBC” filter pattern of the imager. Firstly, the pixels of the ghost sub-image, segmented from the “original image”, are separated into four single filter patterns according to their position relationship in each block. Then, the four images are displayed and their visual discrepancies and reasons are analyzed. Next, the gray value histograms of the “Ghost Ring” with respect to the four separated filter patterns are calculated, which explain the visual

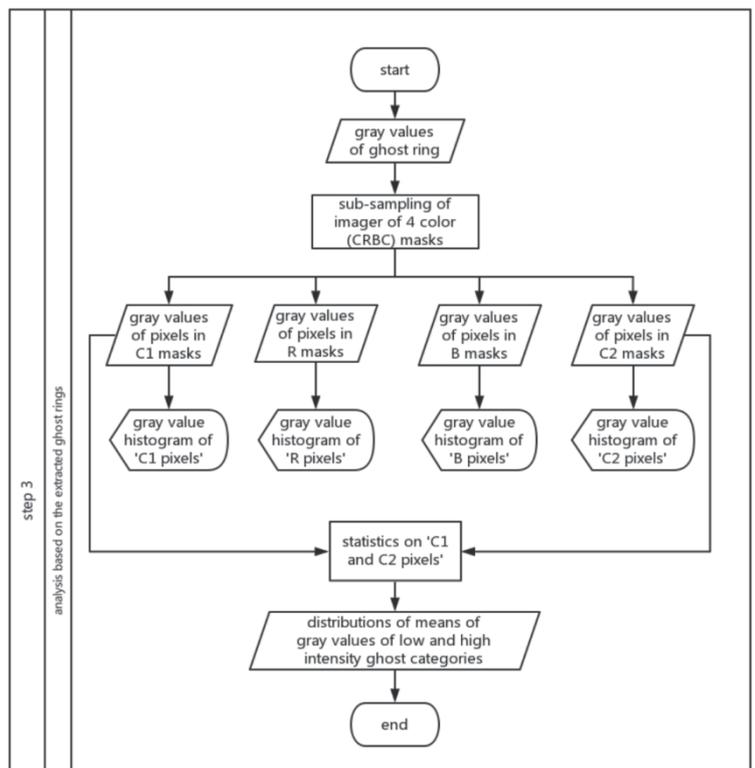


Figure 3.9 Flowchart of the 3rd part of algorithm: Ghost feature calculation

discrepancies in detailed gray value data and are helpful for an analysis from a physical perspective.

Since the clear pixels contain more information of the whole wavelength of visible light, it is more significant to use the statistics on the mean value of the “Ghost Ring” using only the clear pixels’ gray values. The analysis of the distributions of the mean values of ghost artifacts can be carried out between the low light intensity category and the high category.

### 3.3.2 Ghost Cases

With the support of the NV-Cruncher (12bit) function of software smaXview, the high-intensity ghost category and low-intensity ghost category were classified subjectively by observation of human eyes. The images in each category are selected from normal ghost type. Based on the ghost types, the ghost types can be divided into three cases as indicated in Figure 3.10. The ghost feature calculation and algorithm analysis are based on these categories and cases.

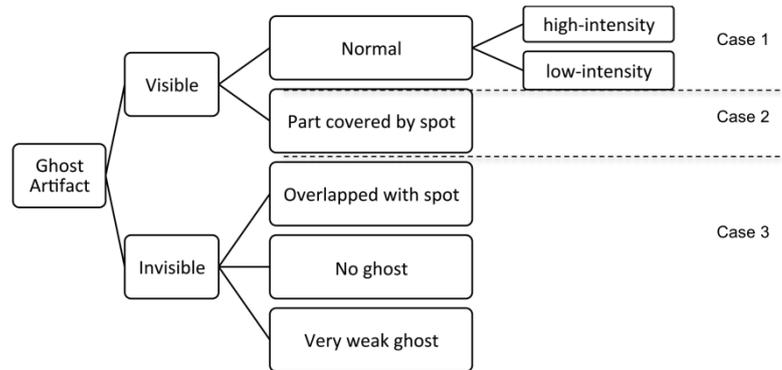


Figure 3.10 Ghost cases

As marked in the figure, case 1 is the main objects in this thesis. The “Ghost Ring” as well as

features of ghost are designed to extract from the images of case 1. For case 2 and case 3 (the entire OK-images), it has to be detected that if ghost exists and has to be judged which type it is.

## 4. Analysis and Discussion

### 4.1 Performance of Ghost in Separated Imager Filter Pattern

The pixels of a single color filter in the CRBC imager are picked out and used to rebuild four new images. The clear pixels are sensitive to all visible light wavelengths and contain more information than the blue and red pixels. The image effects are expected to be different for each of the four single color filter images. Here, a sub-image with a high-intensity ghost is used as an example to display the results (Figure 4.1).

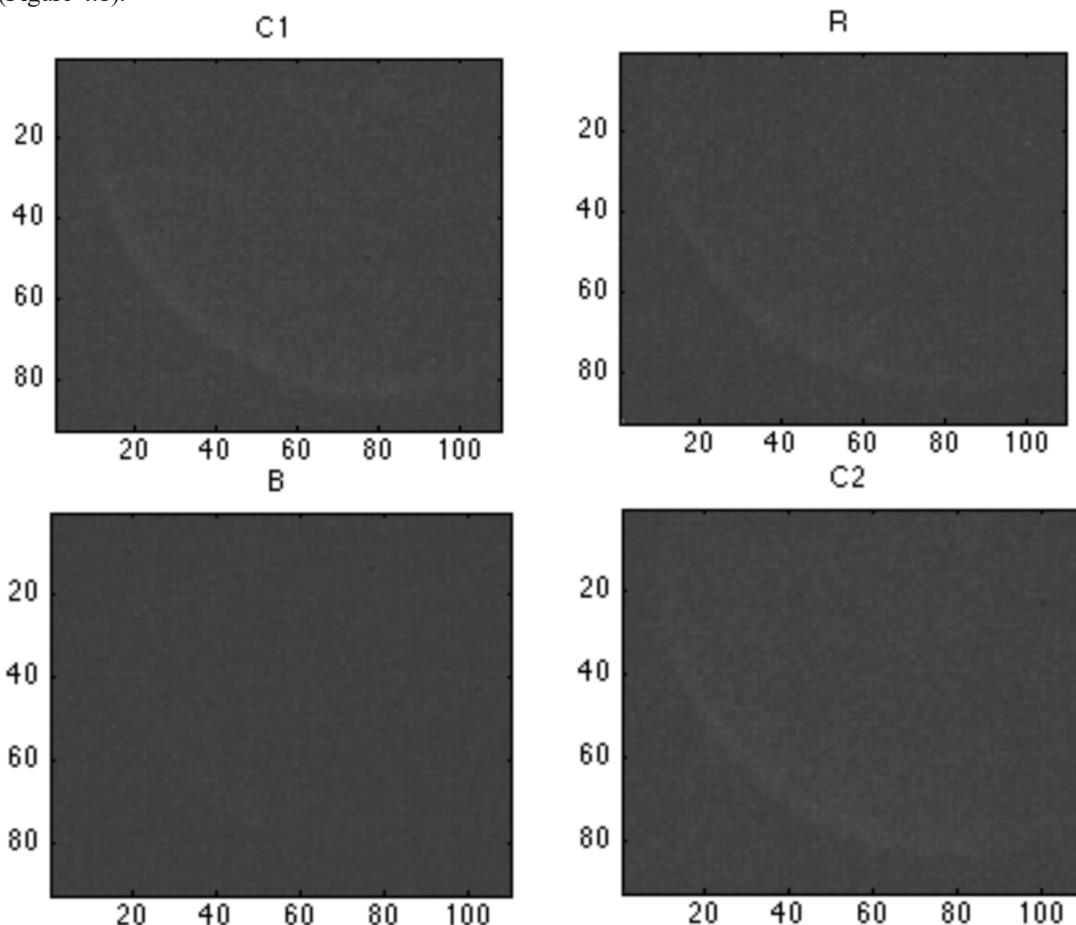


Figure 4.1 Ghost performance in separated imager filters

The ghost artifacts in clear-filter images, which are with title “C1” and “C2”, are in the same intensity level. The intensity of ghost in “R”(red) is a little bit lower than the ghosts in “C1” and “C2”. For the blue sub-image, there is almost no information of ghost artifact. From the gray value histograms (Figure 4.2) of “Ghost Ring” in each sub-images from single color filters, it can be found that gray values of blue filter image are lower than others as a whole.

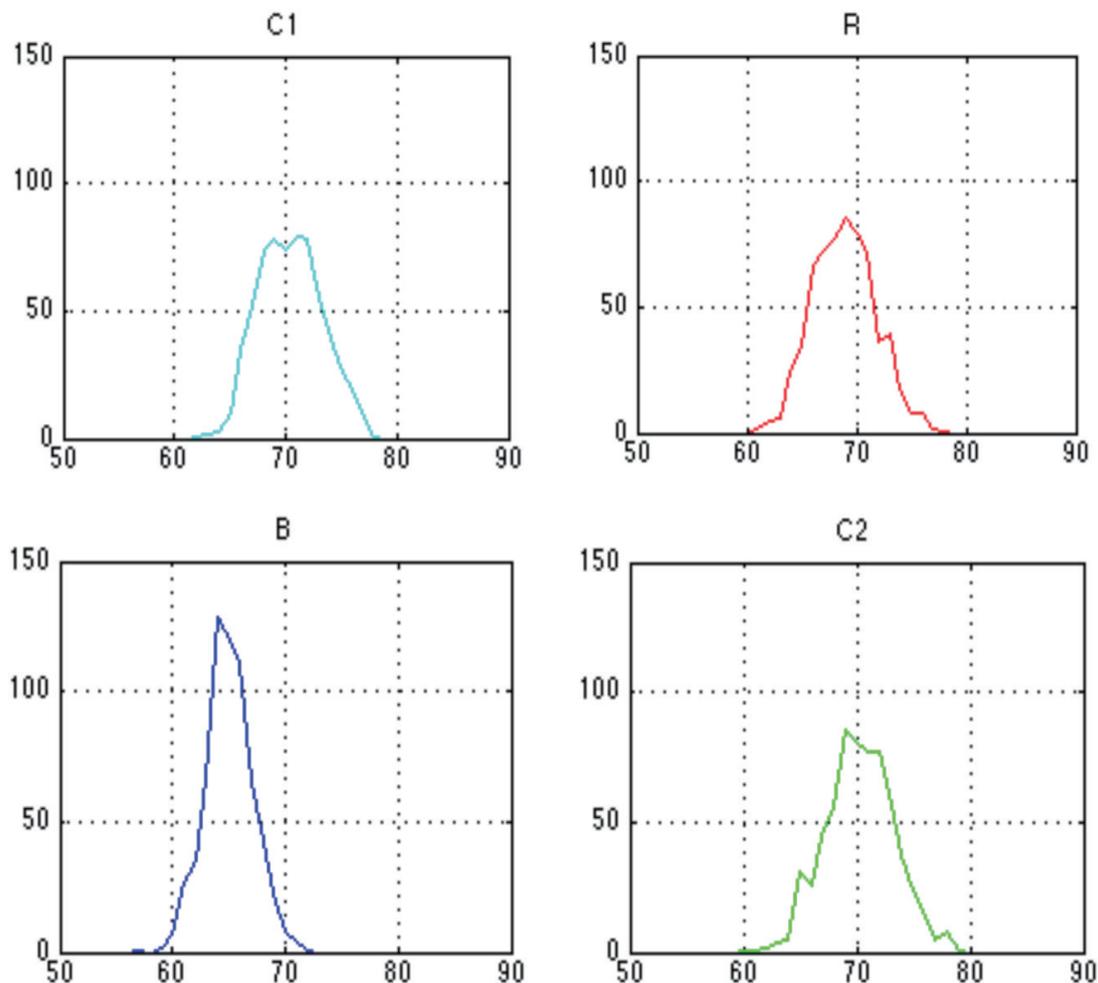


Figure 4.2 Gray value histograms of “Ghost Ring” in separated imager filters

The feature of the image sensor CMOS that used in the HDR camera causes the above phenomenon. The photodiodes can acquire the intensity of light only. Color filters are applied on the photodiodes for capturing color information. The CMOS applied in the testing cameras exhibits different photodiodes spectral sensitivity on different wavelength as shown in figure 4.3.

Spectral sensitivity is the relative efficiency of detection, of light or other signal, as a function of the frequency or wavelength of the signal. In photography, film and sensors are often described in terms of their spectral sensitivity, to supplement their characteristic curves that describe their responsivity (Langford 1998).<sup>[8]</sup> For the images used in this thesis, only visible spectrum (about 400nm-750nm) is acquired. The photodiode has low sensitivity in blue band (450nm-495 nm) according to the spectral sensitivity curve, which means a very small portion (10%-25%) is captured and translate into digital signal by the image sensor, that’s why the ghost stray light is in very low visibility in the sub-image with the pixels in blue filters. The photodiode shows a high sensitivity of red band (620nm-740 nm), so the ghost in the sub-image with the pixels in red filters displays little difference from the ghost in “clear” pixels both in vision and histogram.

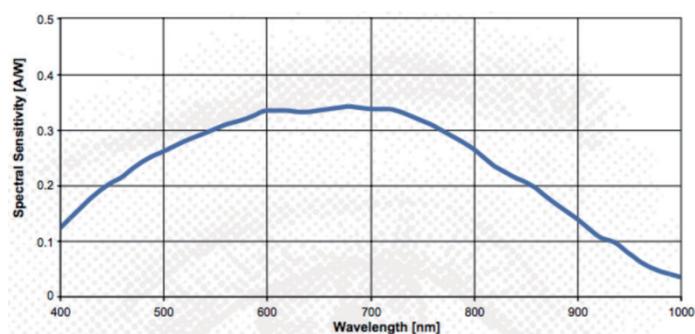


Figure 4.3 Photodiodes spectral sensitivity of HDR CMOS on different wavelength (Chips 2000)<sup>[7]</sup>

#### 4.2 Statistics analysis of “Ghost Ring”

Color feature is gray value feature for gray scale images, is the emphasis of following analysis. As analysis in above section 4.1, the ghost artifact information in blue pixels and red pixels are not complete, only clear pixels are used in the calculation of ghost artifact features. The Region of Interest (ROI) in this step is the region named ‘Ghost Ring’. Pixels in clear filters are selected to represent the ghost information and used in feature calculation. By computing the mean values of the whole clear filter pixels of the “Ghost Ring” in the images of high-intensity category and low-intensity category, the feature data of low intensity ghosts category and complete high intensity ghosts category were obtained. The statistics on the mean values of “Ghost Ring” have been executed in following three steps.

#### 4.2.1 Normal Distribution Test Using Q-Q Plot

Build histogram of each data set (Figure 4.4). The distributions of mean values in the three cases approximate to the normal distribution. To confirm this assumption, normal distribution test is adopted using Quantile-Quantile plot (Q-Q plot). Q-Q plot is used because it provides a visual comparison of the sample quantiles to the corresponding theoretical quantiles (normal distribution quantiles), and also for the reason that the sample numbers of the two sets are small.

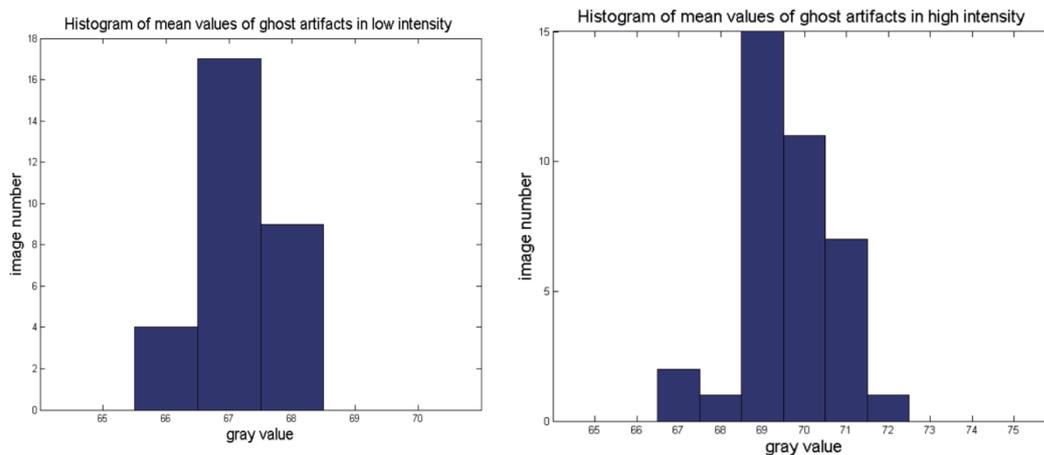


Figure 4.4 Histograms of gray value means of “Ghost Ring” in low-intensity and high-intensity

According Rice (2006) [9], the Q-Q plot is an exploratory graphical device used to check the validity of a distributional assumption for a data set. In general, the basic idea is to compute the theoretical expected value for each data point based on the distribution in question. If the data indeed follow the assumed distribution, then the points on the q-q plot will fall approximately on a straight line.

The first figure in Figure 4.5 displays a Q-Q plot of the sample quantiles of mean values of “Ghost Ring” in high intensity versus theoretical quantiles from a normal distribution, and the second figure is for the “Ghost Ring” in low intensity. According to Matlab, if the distribution of mean values is normal, the plot will be close to linear. The plot has the sample data displayed with the plot symbol ‘+’. Superimposed on the plot is a line joining the first and third quartiles of each distribution (this is a robust linear fit of the order statistics of the two samples). This line is extrapolated out to the ends of the sample to help evaluate the linearity of the data. As shown in the following figures, the two plots are all close to linear, which means both of the two distributions can be defined as normal distributions. Therefore, the statistics can be continued with respect to the properties of normal distribution.

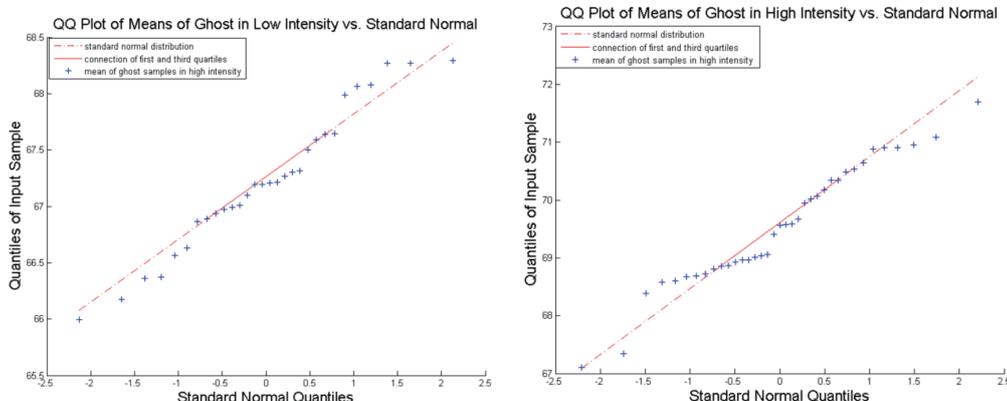


Figure 4.5 Q-Q plot of means of “Ghost Ring” in low-intensity versus standard normal(up) and Q-Q plot of means of “Ghost Ring” in high-intensity versus standard normal(down)

#### 4.2.2 Normal Distribution of Mean Values of “Ghost Ring” in Different Intensity

The probability density of the normal distribution is given by:

$$f(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4.1)$$

Where  $\mu$  is the expectation of the distribution,  $\sigma$  is the standard deviation, and  $\sigma^2$  is the variance. As the normal distribution of samples is confirmed, the normal distribution of each set is build using the variances and expectations of the two sets of mean values. From the Figure 4.6 the difference between the two normal distribution curves is obvious. The expectation of the distribution of means in low intensity is **67.2296**, and the standard deviation is 0.6246. For the distribution of means in high intensity, the expectation is 69.5508, and the standard deviation is 1.0467. The standard deviations of the distributions of means of “Ghost Ring” in high intensity and in low intensity are small, which means the dispersion of each set of mean values is low. It also can be found from the figure, the mean values of each set are in tiny range. The standard deviation of the data in high” set is higher than that in “low” set, it might be caused by the error in classification by eyes, the method of “Ghost Ring” extraction as well as the lack of samples.

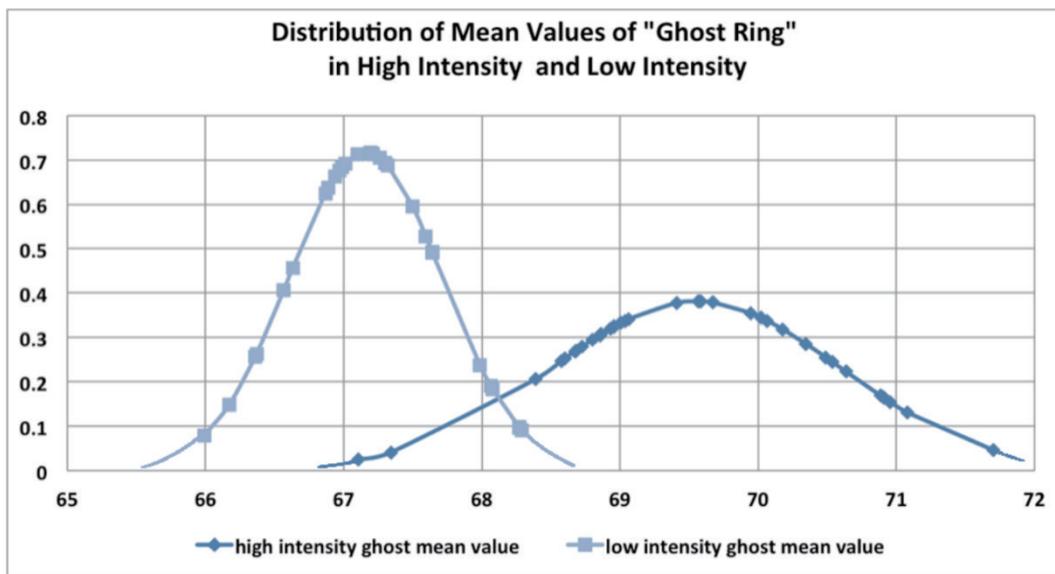


Figure 4.6 Distribution of mean values of “Ghost Ring” in high-intensity and low-intensity

While figure 4.6 expresses the difference of the mean value of ghost in different light intensity, if the mean of gray values of “Ghost Ring” is a feature of the ghost artifact has to be verified by t-test.

#### 4.2.3 T-Test of the Low Ghost Samples and High Ghost Samples

This test is used for comparing the means of two samples (or treatments), even if they have different numbers of replicates. According to online statistics education written by David Scott (), in simple terms, the t-test compares the actual difference between two means in relation to the variation in the data (expressed as the standard deviation of the difference between the means).

Apply t-test to the mean values of “Ghost Ring” in “high” set and “low” set using Matlab to check if the means of two groups are statistically different from each other. The following parameters are obtained.

$$h = 1$$

$$p = 5.8256e^{-16}$$

The null hypothesis for testing is that the data of two sets are independent random samples from normal distributions with equal means and equal but unknown variances. The test decision  $h$  is 1, which means the test rejects the null hypothesis at the **5%** significance level. Furthermore, as the  $p$  value is  $5.8256e^{-16}$ , there is a  $(1 - p) * 100\%$  chance of the means of two sets being significantly different. By convention, we say that a difference between means at the **95%** level is "significant", a difference at **99%** level is "highly significant" and a difference at **99.9%** level is "very highly significant"

As a conclusion, the different normal distributions shown in Figure 4.6 is not a accidental situation, there is a “very high significant” difference between the distributions of mean values of “Ghost Rings” in the “high” set and “low” set. Hence, the mean value of “Ghost Ring” is suitable to be used as the feature of ghost artifact.

#### 4.3 Stability of Ghost Detection and “Ghost Ring” Extraction Algorithm

The stability analysis is based on the image cases as description in section 3.3.2. For the module of ghost detection, the algorithm can correctly judge if a ghost exist in the image and which categories it belongs for all testing images. According to the ghost’s type,

the algorithm is able to process the images in corresponding methods. For the image of case 2, the ghost artifact can be detected and marked out with a red rectangular. A notice of “ghost is part covered by spot” is appeared. However, 2 complete ghosts in high intensity are detected as the “part-covered ghost” because of their short distance to the spot.

The similar response is happened to the first of the case 3 “ghost is overlapped by spot”, but without mark. As there is no visible ghost in the images of the rest of case 3, the response is the warning “no visible ghost”. The algorithm works successfully on the whole images of each category mentioned above. The results conform to the expectation and the design of the algorithm.

The “Ghost Ring” extraction algorithm only runs in case 1. It can extract the “Ghost Ring” in the images of low class with 100% accuracy. While for the normal ghost in high intensity, there are 3 ghosts cannot be extracted as a fit “Ghost Ring” among the 35 images due to the influence from the spot stray light which cannot be avoided to segment into the sub-image. In this case, two solution reminders are shown out, one is changing the number of pixels using in circle fitting, another is resizing the radius of the mask of spot. Meanwhile, For the statistic module, the algorithm can work successfully on every case with respect to the different requirements of statistics or analysis.

## 5. Conclusion

Within the scope of a feasibility study, this paper proposes a new algorithm for automatically detecting ghost artifacts and extracting ghost features in the context of stray light tests for the video camera MPC2 from Bosch. With respect to the results at different stages of this thesis, the major cores are summarized as follows:

- The ghost’s type was subjectively classified into four types: invisible ghost, normal ghost, part-covered ghost, and overlapped ghost. The result of the first module of the algorithm indicates that not only can the existence of ghost artifacts be detected, but also the type of ghost can be determined in single testing images with acceptable reliability in image processing methods.

- A typical feature of ghost artifacts can be obtained by extracting the significant portion of ghost as the “Ghost Ring” from the image with normal ghost within a reasonable accuracy and success rate. Since the algorithm designed in the second module of “Ghost Ring” extraction is sensitive to other stray light, an improved approach should be developed to achieve better robustness.

- The result demonstrates that the mean of gray values from clear filters in the “Ghost Ring” region significantly represents the ghost feature by means of statistics. However, due to the restriction of the number of camera samples, the data used in statistics cannot be used to define a feature value of ghost artifacts through the mean value distributions.

In short, the results from 118 images from 4 test cameras show that this study has achieved the priority goals of detecting ghosts, judging ghosts, and extracting ghost features, which replace the subjective ghost detection by eyes with an objective procedure using software. However, there are still several aspects that can be improved with respect to the accuracy of “Ghost Ring” extraction.

## References:

- [1] Learner, R. C. M., Thorne, A. P., & Brault, J. W. (1996). Ghosts and artifacts in Fourier-transform spectrometry. *Applied Optics*, 35, 2947-2954.
- [2] Kyathanahally, S.P., Döring, A. and Kreis, R. (2018), Deep learning approaches for detection and removal of ghosting artifacts in MR spectroscopy. *Magn. Reson. Med*, 80: 851-863.
- [3] An, J., Ha, S. J., & Cho, N. I. (2014). Probabilistic motion pixel detection for the reduction of ghost artifacts in high dynamic range images from multiple exposures. *Journal of Image and Video Processing*, 42.
- [4] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62-66.
- [5] MathWorks, Inc. (2016). Image Processing Toolbox. Retrieved from: <http://www.mathworks.com/help/images/>.
- [6] Chernov, N., & Lesort, C. (2005). Least squares fitting of circles. *Journal of Mathematical Imaging and Vision*, 23(3), 239-252.
- [7] Chips, I. M. S. (2000). HDRC VGA imager and camera data and features (Tech. Rep.). Inst. Microelectronics, Stuttgart, Germany.
- [8] Langford, M. (1998). *Advanced Photography*. Taylor & Francis.
- [9] Rice, John. (2006). *Mathematical Statistics and Data Analysis*. Cengage Learning.
- [10] Scott, D. (2014). Quantile-Quantile (q-q) Plots. Retrieved from: [http://onlinestatbook.com/2/advanced\\_graphs/q-q\\_plots.html](http://onlinestatbook.com/2/advanced_graphs/q-q_plots.html)