# DIGITAL IMAGE PROCESSING AND SPATIAL FREQUENCY ANALYSIS OF TEXAS ROADWAY ENVIRONMENT

by

ZHEN TANG, B.S.

A THESIS

IN

ELECTRICAL ENGINEERING

Submitted to the Graduate Faculty
of Texas Tech University in
Partial Fulfillment of
the Requirements for
the Degree of

MASTER OF SCIENCE

IN

ELECTRICAL ENGINEERING

Approved

December, 1999

# ACKNOWLEDGEMENTS

The author wishes to thank Professor Bobby Green, Dr. Harlan Thorvilson and Dr. Douglas Gransberg for their guidance and extra work that lay the foundation of this thesis. Special thanks and recognition should be given to Dr. Jon Bredeson and Dr. Michael Giesselmann. The author expresses his appreciation to Dr. Olkan Cuvalci and Dr. Roman Stemprok for their generous and valuable advice. It has been a pleasure for the author to work with Mr. Jokirk Newborogh and Mr. Ilker Karaca. Special thanks are given to Ms. Mary Nichols for her constant support.

TABLE OF CONTENTS

# ABSTRACT

A report presented on the acquisition, storage, processing and analysis of digital images of both fireants activities and small target visibility, beginning with the general introduction of background knowledge in digital image representation, covering acquisition, storage, enhancement and finally development of methods to extract information of interest from the digital images. Fast Fourier transform and digital image processing techniques are reviewed and utilized.

# LIST OF TABLES

LIST OF FIGURES

x

# CHAPTER I
# INTRODUCTION TO DIGITAL IMAGE

Visual images are central to human existence, and the recording and communication of images have stirred interest through the ages. A talented artist could create a representation of a natural scene with simple drawing tools. From ancient cave murals to oil paintings, people have used their eyes to observe the surroundings and their brains to subjectively process the materials and their hands to perform the re-manifestation. Obviously, the human artist as an image-recording device is limited and to some extent untrustworthy because the artist will almost certainly add some element of personal interpretation to the image. To reduce these limitations, a device for recording images was sought. With the invention of photography, people started recording images objectively for the first time. The first practical method of photography was demonstrated in 1839. Since that time, classical photography has advanced tremendously in capabilities, and several other types of image storage technologies have been developed, including video, motion pictures, facsimile, and photocopying. The iconoscope, the first practical camera tube, was invented by Zworykin in 1932 and used in television.

As more inventions become mature products in video business, we now have many effective tools to record and store interesting images to the best extent. For example, for the research described in this paper, the author utilizes Sony DCR-VX1000 Digital Handycam and its Sony DV™ format cassette to record the video and DPS SPARK™ Direct DV Editing System to down load the video into '.avi' format files on a computer hard drive.

An image is simply considered two dimensional within this thesis. It can have representations in both spatial domain and frequency domain although in our day-to-day conversations we usually refer an image to the former. Thus, in this chapter, it is very reasonable to introduce the spatial domain, first.

## 1.1 Spatial Domain Representation

In computer imaging, a two-dimensional image can be represented by a two-dimensional function $f(x, y)$ where $x$ and $y$ are the spatial coordinates in the 2-D image plane and $f$ is either a scalar describing the gray level or a vector containing three components describing the intensities of the three primary colors at the location represented by coordinates $x$ and $y$.

## 1.2 Monochrome Digital Image

Achromatic light has only one attribute, its *intensity*. The term *gray level* refers to a scalar measure of intensity that ranges from black, to grays, and finally to white.

A monochrome analog image is an array of pixels represented by scalars indicating light intensities at various points. To say it in another way, the value of $f$ of coordinates (x, y) is proportional to the light intensity recorded on the corresponding point on the image plane by a two-dimensional photodetector array.

A monochrome digital image is an array of pixels represented by discrete scalars indicating light intensities at various points. For example, Figure 1.1 shows a monochrome digital image of planet Mercury (the picture is rotated 90 degree clockwise for the sake of the page layout). In this case, a telescope collects the light from a certain solid angle and records the light intensities from different directions onto different locations (pixels) on the image plane according to the optical path of the telescope (Figure 1.2). The intensity value is then converted into digital values, or quantized.

## 1.3 Color Digital Image

Color theories state that nearly every arbitrary color can be composed of three primary colors of appropriate intensity according to the structure of human eye, i.e., any color can be split into these three basic colors. In choosing these

primary colors, it is essential that none of them can be generated from the other two.

The CIE (the International Commission on Illumination) designated in the year 1931, the following specific wavelength values to the three primary colors (Table 1.1).

Table 1.1: Three primary colors of CIE

| Color | Wavelength, $10^{-9}$ m |
|---|---|
| Red | 700 |
| Green | 546.1 |
| Blue | 435.8 |

For color TV (and color cathode-ray tube monitor), the following three primary colors have been determined [1] (Table 1.2).

Table 1.2: Three primary colors of color CRT

| Color | Wavelength, $10^{-9}$ m |
|---|---|
| Red | 615 |
| Green | 532 |
| Blue | 470 |

The color picture to be transmitted is scanned with three optical devices, each with a different color filter in front of it. The three channels (R, G, and B) are adjusted such that if a plain white area is scanned, all three outputs have equivalent voltages [1]. The outputs are then separately quantized to be stored in various digital formats.

In computer displays, all color variations for the pixels are derived from the three primary colors indicated above. For today's 24-bit images, we use three bytes per pixel to represent a color value. Each primary color is represented by

one byte (8 bits). For example, a white color in a 24-bit image is represented by FFFFFF (in hexadecimal) meaning 100% red (FF), 100% green (FF) and 100% blue (FF). Figure 1.3 is copied from a bitmap image file. The original file has 64 x 64 pixels and 'True Color' (24 bits per pixel). It contains 16384 bytes of data.

## 1.4 Conversion from Color to Monochrome

Since the calculations in this paper usually include the transformation of the image data from 'True Color' (24 bit) to gray-level, it is necessary to describe the principles of the conversion.

The method RGB2GRAY provided in MATLAB® (The MathWorks, Inc, Natick, Massachusetts) converts RGB images to grayscale by eliminating the hue and saturation information while retaining the luminance. This method has been broadly accepted and used in the color TV and black-and-white TV compatibility.

Let us look at a pair of pixels (29, 29) and (32, 32) of Figure 1.3. Before the conversion, the values of the first pixel are 48/255 red, 62/255 green and 177/255 blue, those of the second 152/255, 149/255 and 211/255. The brightness part of the black-and-white signal is generated by adding the intensities of these primary colors. Because the human eye exhibits different sensitivity for colors of the same intensity, it must be a weighted addition. The *brightness* –or *luminance* (Y)– signal is given by [1]:

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

After performing the 'rgb2gray' conversion, the single value of the first pixel becomes 70.9/255, that of the second 157.0/255. (From now on all the images involved in the calculations are implicitly gray level images.)

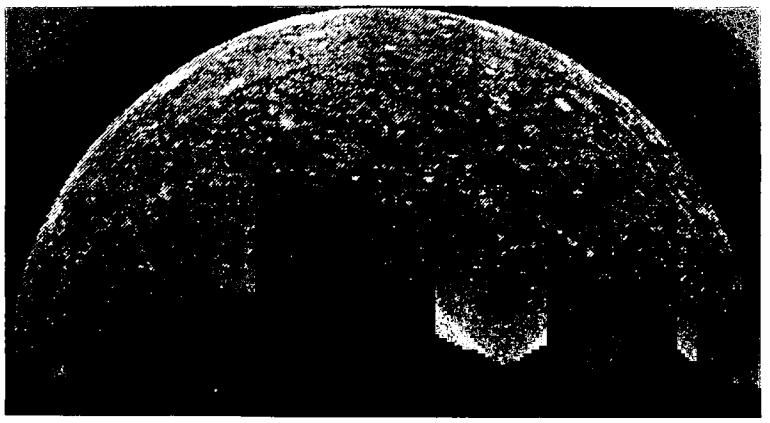


Figure 1.1: A picture of Mercury.

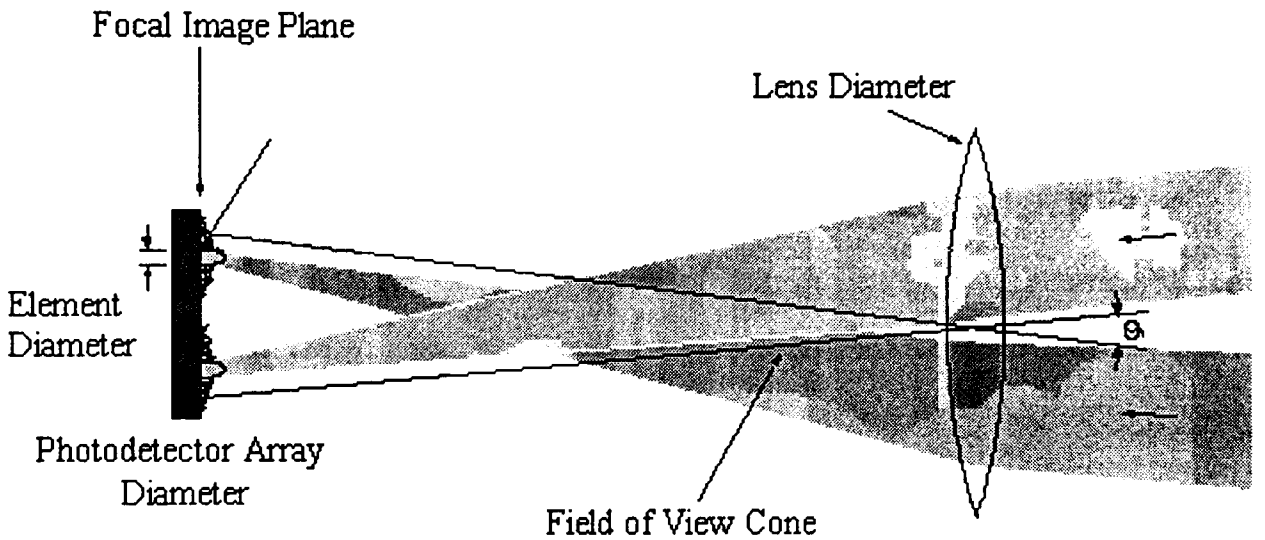



Figure 1.2: Illustration of telescope imaging system.

Figure 1.3: Earth (64 x 64, True Color).

Figure 1.4: Earth (64 x 64, Gray Level).

# CHAPTER II
# ACQUISITION, COMPRESSION AND STORAGE

This chapter will mainly introduce the procedures of the digital videotaping, converting the video format signals into image file format acceptable in computers. The various technologies involved will only be discussed to allow the readers have a general understanding. More of the technologies should be available through contacts to companies that developed them.

## 2.1 Charge Coupled Device Imaging

The author has been using digital video camcorders. Just like the telescope mentioned in Chapter I, the optical system of our digital video camcorder collects the light from a certain solid angle and records the light intensities from different directions onto different locations (pixels) on the image plane according to the optical path of the telescope. In this case, installed on the image plane is the most featured part of this kind of equipment, the charge-coupled device (CCD).

The CCD has revolutionized the imaging industry and is now the detector of choice in high-performance imaging systems. The fundamental process that occurs in CCD imaging is the conversion of photonic input to electronic output. As illustrated in Figure 2.1, photons incident on the CCD will be converted to electron/hole pairs, and the electrons will be captured under the polysilicon gate electrodes of the CCD. These electrons are then transferred in a "bucket brigade" fashion to the output amplifier where the charge is converted to a voltage output signal. An analog processing chain further amplifies this signal, and finally it is digitized before being transferred to the storage part of the camcorder or a host computer for display, image processing, and/or storage.

## 2.2 Digital Video Format (Sony DV™)

All the images that the author deals with are captured in motion on SONY $^{Mini}$ DV tapes using Sony Digital Handycam (a model of Sony Digital Video Camera Recorder).

Sony DV™ is designed as an acquisition format—its reason for existence is for people to take pictures using it. People will almost always copy a SONY $^{Mini}$ DV tape to some other format to finish their tasks.

For this reason, Sony DV™ picks a different compression system than MPEG-2 (the Motion Picture Expert Group) that Digital Video Disk (DVD) uses. MPEG-2 is similar to the popular JPEG (the Joint Picture Expert Group) format - you can do moderate compression and have a perfect image, or you can do a large amount of compression and have a barely viewable one. The main difference between this and the compression scheme used for DV tape is that the DV tape scheme is simpler and compresses a great deal less. This is because it has to be done in real time, as pictures are streaming into the camera at incredibly high speeds.

DV video is sampled at the rate of 720 pixels per scan-line although the color information is sampled at 4:1:1 in 525-line (NTSC) and 4:2:0 in 625-line (PAL) formats.

DV compression is a hybrid — it is compressed using a Discrete Cosine Transform (DCT), the same sort of compression used in motion-JPEG. However, DV's DCT allows for more optimization within the frame than do JPEG compressors, allowing for higher quality at the nominal 5:1 compression factor than motion-JPEG.

DV also uses intra-frame compression: Each compressed frame depends entirely on itself and not on any data from preceding or following frames. However, it also uses adaptive inter-field compression; if the compressor detects little difference between the two interlaced fields of a frame, it will compress them together, freeing up some of the "bit budget" to allow for higher overall quality

elsewhere in the image. In theory, this means that static areas of images will be more accurately represented than areas with a lot of motion; in practice, this can sometimes be observed as a slight degree of "blockiness" in the immediate vicinity of moving objects (an artifact).

DV video information is carried in a nominal 2.5 megabit per second (Mbps) data stream. Once you add in audio, subcode (including timecode), Insert and Track Information (ITI), and error correction, the total data stream comes to about 3.56 Mbps. One Gigabyte of storage is about 4 minutes 45 seconds of DV video. Two Gigabytes is about 9 minutes 30 seconds.

## 2.3 Discrete Cosine Transform

Following is a very simple introduction [5] to discrete cosine transform, the concept involved in last section.

The 1-D discrete cosine transform (DCT) is defined as

$$C(u) = \alpha(u)\sum_{x=0}^{N-1} f(x)\cos\left[\frac{(2x+1)u\pi}{2N}\right] \tag{2-1}$$

for $u = 0,1,2,\ldots N-1$. Similarly, the inverse DCT is defined as

$$f(x) = \sum_{u=0}^{N-1} \alpha(u)C(u)\cos\left[\frac{(2x+1)u\pi}{2N}\right] \tag{2-2}$$

for $x = 0,1,2,\ldots N-1$. In both Eqs. (2-1)and (2-2), $\alpha$ is

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & for \quad u = 0 \\ \sqrt{\frac{2}{N}} & for \quad u = 1, 2, \ldots, N-1. \end{cases} \tag{2-3}$$

The corresponding 2-D DCT pair is

$$C(u,v) = \alpha(u)\alpha(v)\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} f(x,y)\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cos\left[\frac{(2y+1)v\pi}{2N}\right] \tag{2-4}$$

for $u,v = 0,1,2,\ldots N-1$, and

$$f(x,y) = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\alpha(u)\alpha(v)C(u,v)\cos\left[\frac{(2x+1)u\pi}{2N}\right]\cos\left[\frac{(2x+1)v\pi}{2N}\right] \quad (2\text{-}5)$$

for $x = 0,1,2,\ldots N-1$, where $\alpha$ is given in Eq. (2-3).

In recent years, the discrete cosine transform has become the method of choice for image data compression because it provides a good compromise between information packing ability and computational complexity. It has become the international standard for transform coding systems. Compared to other input independent transforms, it has the advantages of having been implemented in a single integrated circuit, packing the most information into the fewest coefficients, and minimizing the block-like appearance (blocking artifact) when boundaries between sub-images become visible.

## 2.4 'AVI' and 'BMP' Format

The DV video format requires a sustained data rate of 3.56 Megabyte/sec in order to play real time video and stereo audio. This can only achieved if the disk drive used for recording and playback can consistently meet this specification. It is often recommended to select an SCSI A/V drive which is connected to a wide or ultra wide SCSI adapter. It is also necessary to make sure that the drive is not the host computer's primary system drive.

Given a video clip stored on DV tape, we have two choices. One choice is through a DPS Spark Direct DV™ Editing System plus a IEEE-1394 I/O ("FireWire" or "i.LINK" cable) to use the Spark Recorder program to transfer full resolution digital video and audio to the host computer's hard disk drive in the Microsoft Windows AVI file format. After the clip has been recorded, edit it into the Microsoft Windows BMP file format pictures using Adobe Premiere or any other nonlinear editing program which supports the Microsoft Windows AVI file format. The other choice is to use the Spark Recorder program's 'Grab BMP Still' function. The output will be a BMP file format picture, which saves a lot of work. The only

problem with this method is it is hard to pick the right moment if the objects in the video are moving rapidly.

The AVI (Audio Video Interleave) format was developed by Microsoft (Seattle, Washington). It is a special case of the RIFF (Resource Interchange File Format). It is the most common format for audio/video data on PC's, and comes with the Windows operating system as part of the "Video for Windows (VfW)" distribution. The format is interleaved such that video and audio data are stored consecutively in an AVI file (i.e., a segment of video data is immediately followed by a segment of audio data.)

BMP is the native bitmap file format of the Microsoft Windows® (Microsoft, Seattle, Washington) environment. There are many other file formats for the same purpose, including RLE, JPEG, TIFF, TGA, PCX, PNG, PCD and GIF. The term *bitmap* is a throwback from when monitors could only display one other color besides black. For two-color data files that store an image, each bit in the data file represents one pixel; a 1 meant the pixel was on, a 0 meant the pixel was off (Figure 2.2). Therefore, a two-color image is a map of bits. Windows bitmap files are stored in a device-independent bitmap (DIB) format that allows Windows to display the bitmap on any type of display device. The term "device independent" means that the bitmap specifies pixel color in a form independent of the method used by a display to represent color. The default filename extension of a Windows DIB file is .BMP.

The bitmaps studied in this section will be 256-color bitmaps, where eight bits represents one pixel. Windows' BMP format can be stored uncompressed, so reading BMP files is fairly simple; most other graphics formats are compressed, and some, like GIF, are difficult to decompress. There are a few different sub-types of the BMP file format. The one studied here is Windows' RGB-encoded BMP format. For 256-color bitmaps, it has a 54-byte h1er followed by a 1024-byte palette table. After that is the actual bitmap, which starts at the lower-left hand corner. It efficiently stores mapped or unmapped RGB graphics data with pixels 1-,

4-, 8-, or 24-bits in size. Data may be stored raw or compressed using a 4-bit or 8-bit RLE data compression algorithm.

BMP is an excellent choice for a simple bitmap format which supports a wide range of RGB image data. Most importantly, in our research, it is very convenient to use a high level software package, such as MATLAB® to read image files of this format.

## 2.5 DPS Spark Plus™ System

It will be important and convenient for us to find a way to transfer digital audio and video to the hard drive, without worrying about generation loss or tedious analog-to-digital conversions. With DPS Spark Plus™(Markham, Ontario, Canada), we are able to move or copy data from one storage device to another without alteration or degradation. No matter how many times we transfer your video and audio, the quality will remain equal to our original DV source. Spark Plus eliminates the analog video digitization process of current video capture hardware by transferring already digital DV data from DV tape to a hard drive. Then we can easily extract single images from the footage with Adobe Premier(tm) 4.2 (San Jose, California).

At the heart of this complete editing system is the AHA-8945(tm) SCSI and FireWire host adapter card. The FireWire connects digital camcorders, digital VCRs and computers; the SCSI host adapter controls our computer's SCSI hard drive or any other SCSI devices we may include in the future.

DPS Spark Plus offers a choice of display sizes. Choose from 90x60, 180x120, 360x240 of full DV resolution at 720x480. There's also a "Fit -to-Window" option. This representation of the full frames doesn't affect the resolution of the DV files which remain on our hard drive.

DPS Spark System includes:

DPS Spark Plus(tm) Software for 32-bit Windows,

Adaptec IEEE-1394 to PCI Host Adapter,

Adobe Premiere(tm) 4.2 Editing Software - Full Version*,

FireWire Cable: 4-pin to 6-pin.

Host Requirements are:

Windows NT (Version 3.51 or higher) or Windows 95,

Pentium 133 or faster CPU,

Available PCI Expansion Slot,

32 MB RAM,

Fast/Wide, Ultra or Ultra/Wide SCSI host adapter,

A/V Type SCSI Drive (dedicated for DV data),

DV Camcorder or DVCR with bi-directional IEEE-1394 Port,

SVGA Accelerated Graphics Adapter,

CD-ROM Drive,

Mouse.

## 2.6 Storing, Archiving and Frame Grabbing

The size of the '.avi' files generated by the conversion are usually so big that only a large hard drive can hold several of them. But when there are too many '.avi' files for a limited size hard drive, we have to use recordable CD's to archive the gigantic size files that are not in use for the moment.

For the '.avi' files that are currently in use, the author uses Adobe Premiere 4.2 to open them which allows us to 'export' any frame as a Bitmap picture. Originally in the video recorded on the tape, a timing system shows the time to the accuracy of $1/30^{th}$ second. It is very convenient for a user to track the time and decide which exact moment he wants to export a frame. This function provides us with a possibility of pulling out frames at regular time intervals that could be very helpful in our research. Details of timing and naming of the '.bmp' images that we use will be discussed later.

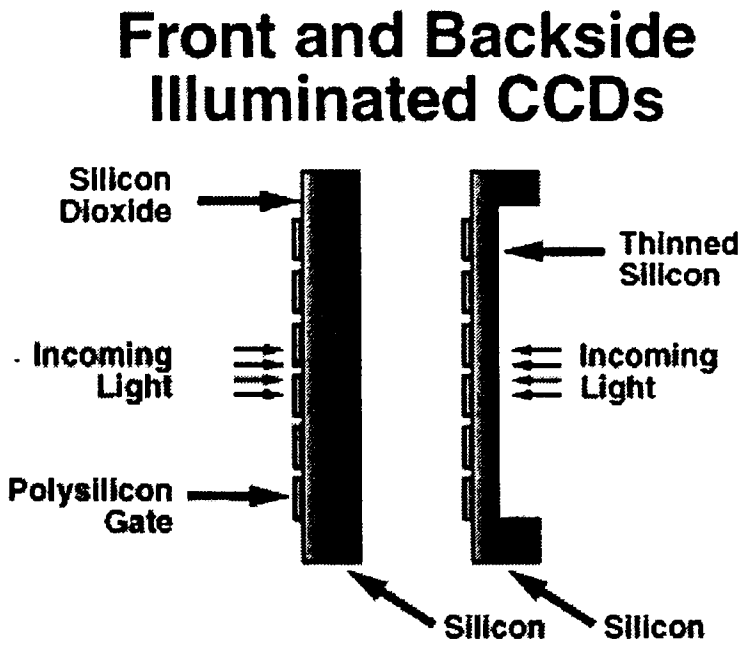



Figure 2.1: Front and Backside Illuminated CCDs.

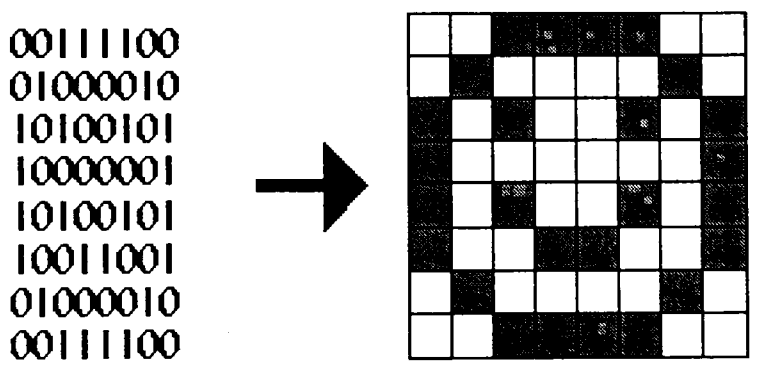



Figure 2.2: A black & white bitmap
in memory and on the screen.

# CHAPTER III
# IMAGE ENHANCEMENT

## 3.1 Enhancement by Point Processing

### 3.1.1 Compression of Dynamic Range

Compression of dynamic range is a simple intensity transformation for preparing the raw images for the later calculations and discussions. It is only based on the intensity of single pixels.

Sometimes the dynamic range of a processed image far exceeds the capability of the display device or the human eye. In these cases, only the brightest parts of the image are visible on the display screen or by human eyes. The author often comes across such problems when displaying the Fourier spectra of images. An effective way to compress the dynamic range of pixel values is to perform the following intensity transformation:

$$s = c \cdot \log(1 + |r|) \tag{3-1}$$

where c is a scaling constant and the logarithm function performs the desired compression.

### 3.1.2 Histogram Processing

The histogram of a digital image with gray levels in the range [0, L-1] is a discrete function $p(r_k)=n_k/n$, where L is the number of gray levels in the digital image, $r_k$ is the kth gray level, $n_k$ is the number of pixels in the image with that gray level, n is the total number of pixels in the image, and k = 0, 1, 2, . . . , L-1.

Loosely speaking, $p(r_k)$ gives an estimate of the probability of occurrences of gray-level $r_k$. A plot of this function for all values of k gives us a global description of the appearance of the image. For example, a histogram with gray-levels concentrated toward the dark end of the gray-level range suggests a corresponding dark characterized image. The opposite is also true. A histogram with a very narrow shape indicates little dynamic range and thus corresponds to an

image having low contrast. A histogram with significant spread suggests an image with high contrast.

The shape of the histogram reveals the possibility for contrast enhancement. The following discussion develops methods for manipulating histograms in a consistent and meaningful manner.

### 3.1.3 Image Subtraction

The difference between two images can be expressed as

$$g(x,y) = f(x,y) - h(x,y) \tag{3-2}$$

where $f(x,y)$ and $h(x,y)$ are two images. The difference $g(x,y)$ is obtained by computing the difference between all pairs of corresponding pixels from *f* and *h*. Illustration is shown in Figure 3.1.

The difference $g(x,y)$ can be positive at one pair of coordinates and negative at another. The author has two approaches to deal with it. The first method (called Method One in this paper) is to take the absolute value. The resulting difference image has positive gray levels wherever there is difference. The gray levels are also proportional to how much the two images distinguish from each other at the specific coordinates (shown in Figures 3.2, 3.3, and 3.4). The second method (called Method Two in this paper) is more intended to deal with the images with actively moving ants that the author deals with. This approach requests a strong condition on the images. All the moving objects must be uniformly either darker or brighter than the backgrounds. When this condition is satisfied, we can keep only the positive differences and assign any coordinates with negative differences to zero (shown in Figure 3.5). The result shows only the moving objects from either Image 1 or Image 2, depending on the gray levels of the moving objects and the background. The latter approach turns out to be very useful in our research.

## 3.2 Spatial Filtering

In spite of the high quality equipment employed in this research, there is noise in the acquired final digital image. With a desire for less computing time and less coding, *spatial filtering* involving the use of spatial masks for image processing is performed to reduce the noise, instead of *frequency domain filtering* using the Fourier Transform. There are ready-to-use packaged functions from MATLAB®. These spatial masks are called *spatial filters*. Spatial filters operate on neighborhoods. There are linear and nonlinear spatial filters. Without explicit use of any coefficient, nonlinear spatial filters work directly on the values of the pixels in the considered neighborhood. In this section, we look at a nonlinear smoothing filter. Noise reduction can be achieved effectively with a nonlinear filter whose basic function is to compute the median gray-level value in the neighborhood in which the filter is located.

### 3.2.1 Median Filtering

Loss of sharp details is a principal difficulty of another smoothing method, lowpass filtering. With an objective to achieve noise reduction rather than blurring, *median filter* is a better alternative. In median filtering, the gray level of each pixel is replaced by the median of the gray levels in a neighborhood of that pixel, instead of by the average.

The median *m* of a set of values is such that half the values in the set are less than m and half are greater than m. The computing involves sorting the values of the pixel and its neighborhood, determining the median, and assigning this value to the pixel. The principal function of median filtering is to force points with distinct intensities to be more like their neighbors, actually eliminating intensity spikes that appear isolated in the area of the filter mask.

An example is shown to illustrate the effect of median filtering. Figure 3.6 is the result of median filtering with a 5-by-5 mask on Figure 3.5.

To show the effect in detail, small sections of both Figures 3.5 and 3.6 are cut at the exactly same location and put below (see Figures 3.7 and 3.8). They show one red imported fire ant. Readers can clearly see how a 5x5 mask has done its job.

The result of a huge mask (73x73) is shown in Figure 3.9. Although it does show better smoothing effect than the 5x5 mask, the computing time is significantly larger and unacceptable for the purpose we have currently. Therefore, the 5x5 mask has been the choice for our calculations in the RIFA image analysis.

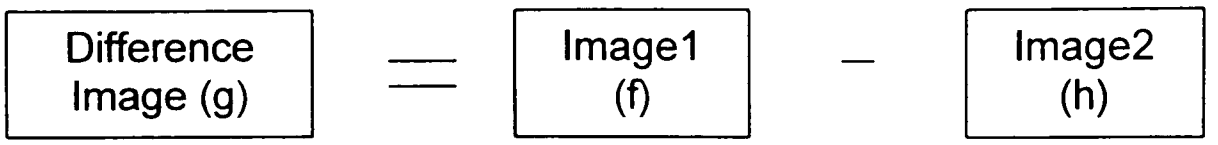



Figure 3.1: Illustration of image subtraction.

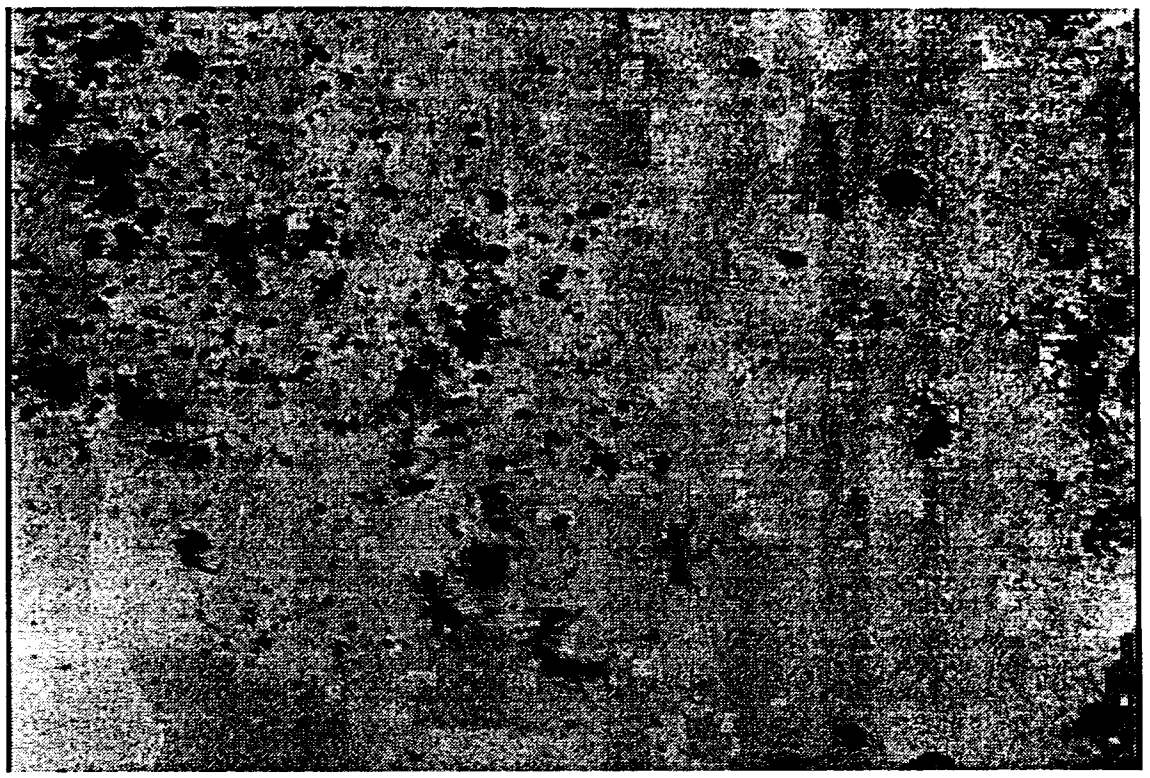


Figure 3.2: Image 1 taken at t = t1.

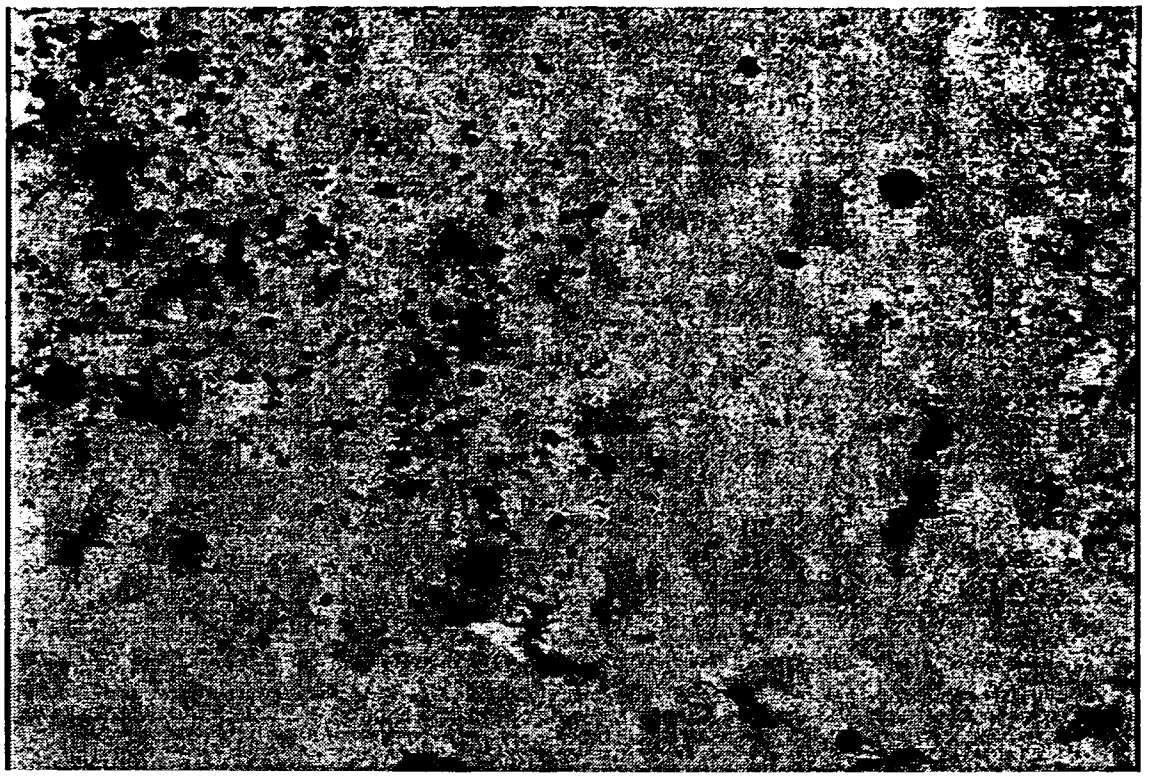


Figure 3.3: Image 2 taken at t = t2 = t1 + 0.5 sec.

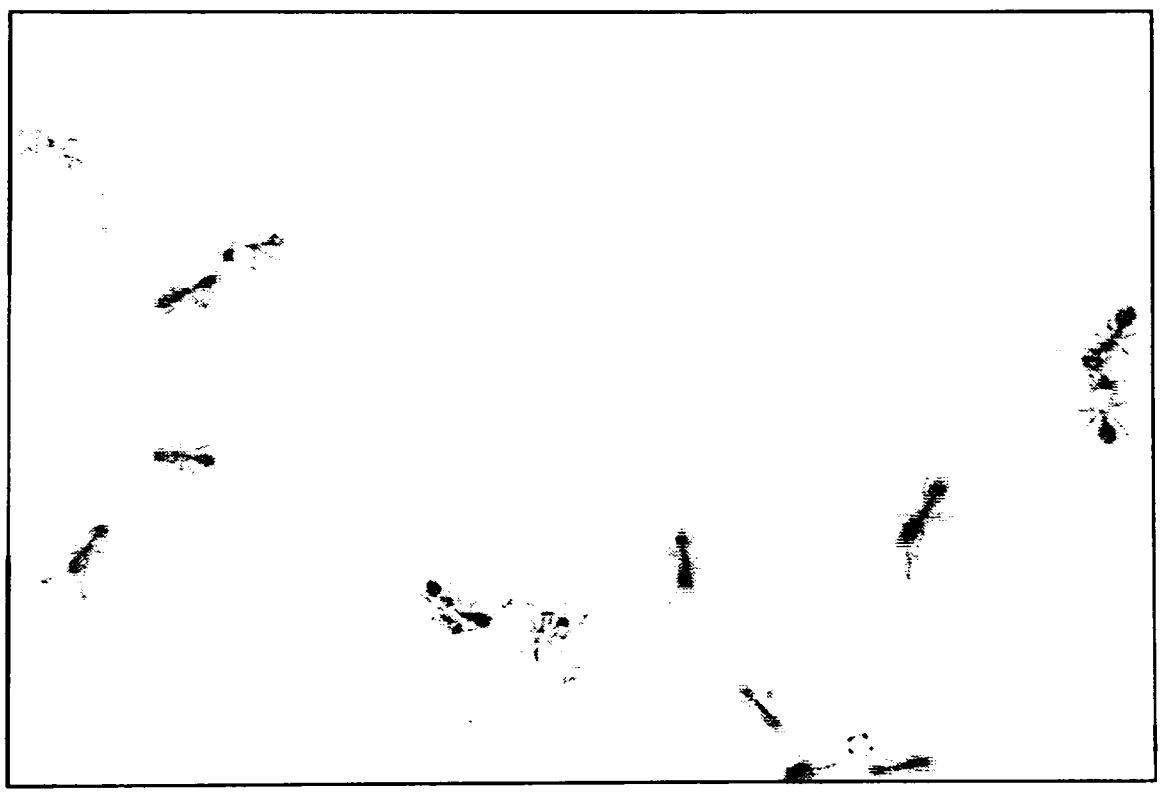


Figure 3.4: Difference between Figures
3.2 and 3.3 using method one.

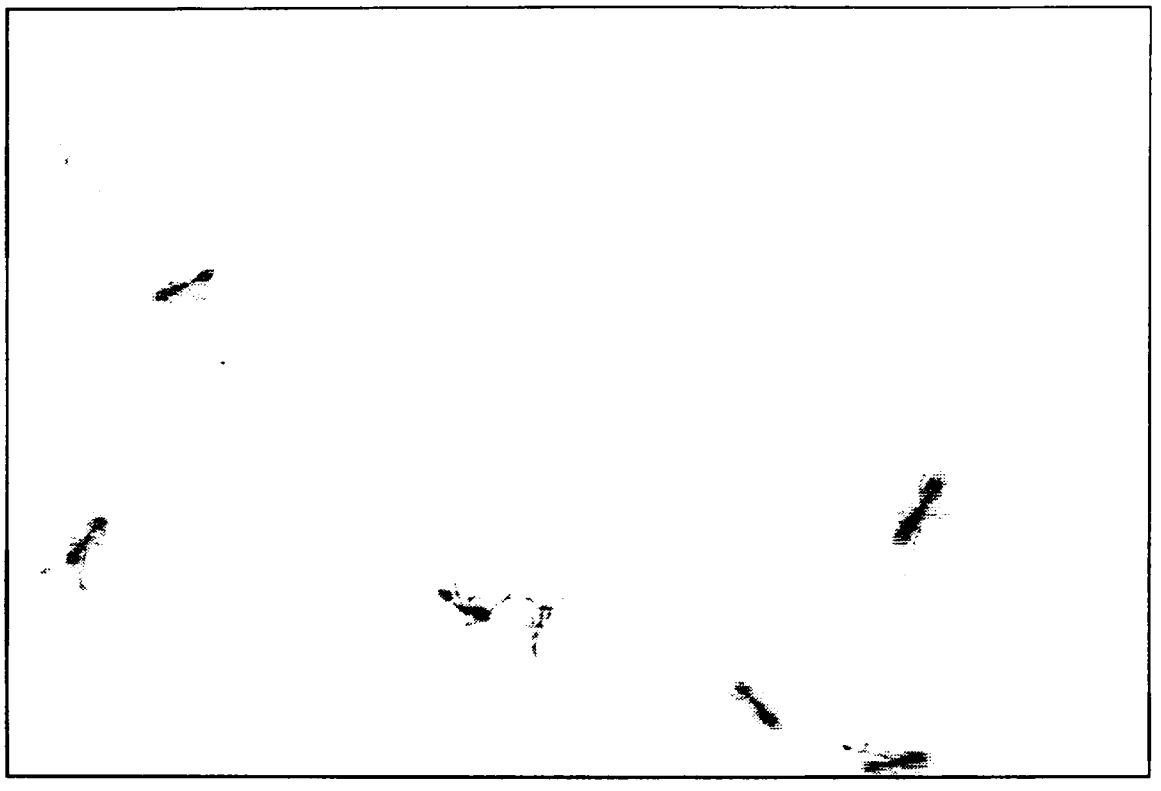


Figure 3.5: Difference between Figures 3.2 and 3.3 using method two.

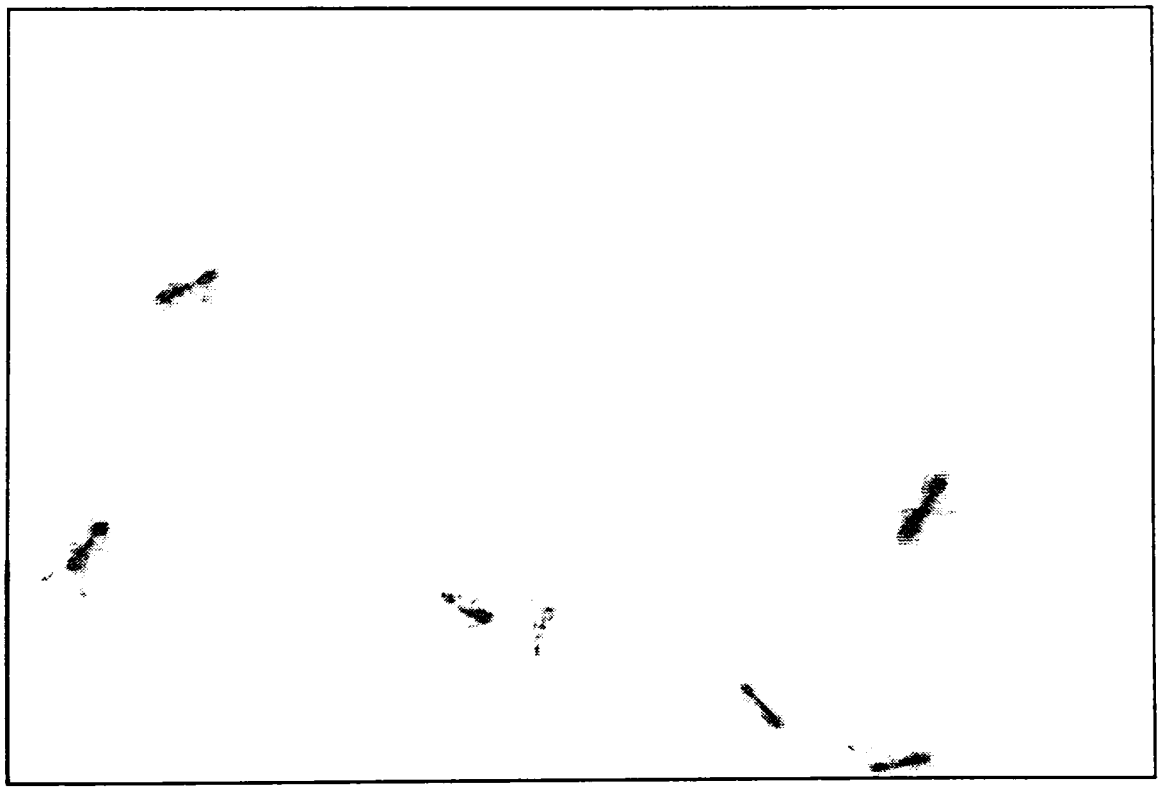


Figure 3.6: Figure 3.5 median-filtered using a 5x5 mask.

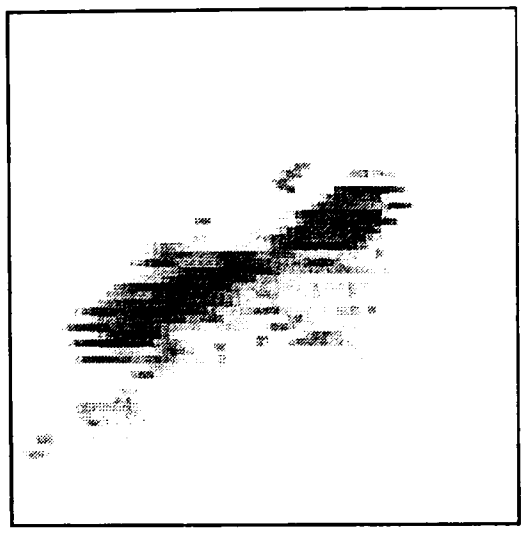


Figure 3.7: One ant from Figure 3.5.

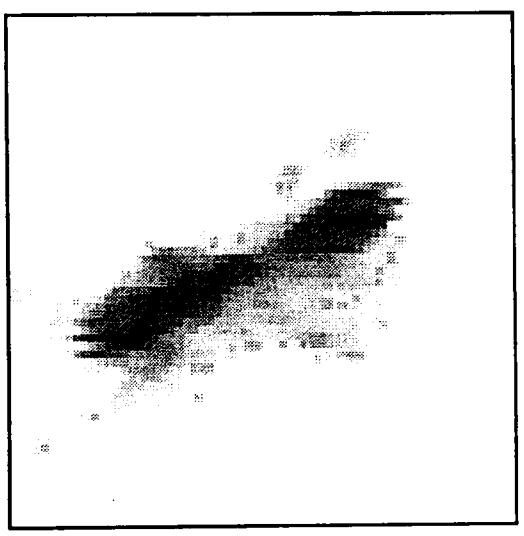


Figure 3.8: One ant from Figure 3.6

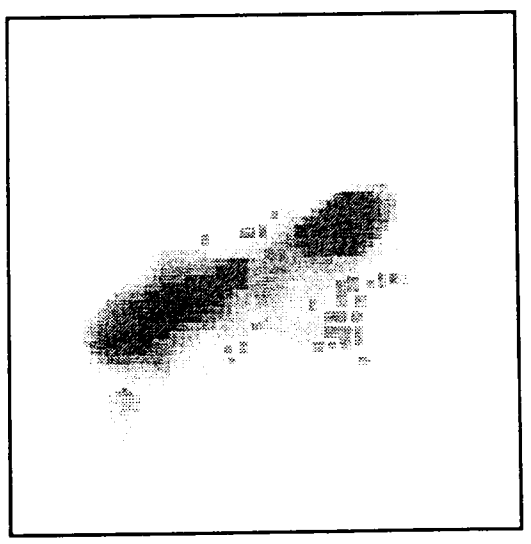


Figure 3.9: An example of 73x73 mask median-filtering result.

# CHAPTER IV
# FOURIER ANALYSIS

In the late 1960's, Blakemore and Campbell [8] suggested that the neurons in the visual cortex might process *spatial frequencies* instead of particular features of the visual world. In English, this means that instead of piecing the visual world together like a puzzle, the brain performs something akin to the mathematical technique of *Fourier Analysis* to detect the form of objects. While this analogy between the brain and the mathematical procedure is at best a loose one (since the brain does not really "do" a Fourier Analysis), whatever the brain actually does when we see an object is easier to understand within this context.

In this research, the equipment collects the light from a certain solid angle and records the light intensities from different directions onto different locations (pixels) on the image plane, with information of each direction into each *pixel* accordingly. It means, there is no *spatial frequency* involved until researchers look at these video clips. However, once people look at them, they might be doing *Fourier Analysis* with their eyes and brains as Blakemore and Campbell suggested. It is only natural to go further to directly carry out the *Fourier Analysis* and examine the result with the help from the mathematical software and computers we possess. Thus, a review of the basic concepts of Fourier Analysis will be very helpful starting point.

## 4.1 Fourier Theory Fundamental

### 4.1.1 Fourier Series

Fourier analysis was initially developed by the Physicist Joseph Fourier to study heat transfer problems. Fourier recognized that any function (see Figure 4.1), $f_p(x)$, whose graph displays a periodicity, T, can be considered to be an infinite sum of sinusoidal functions.

The Fourier series may be represented as the sum of a series of sine functions, cosine functions, complex exponential functions or any of several other sinusoidal representations [9, 10, 11].

4.1.1.1 Series Representation of a Limited Function

Consider a uniform and finite periodic function $f_p(x)$ with period p defined over a range

$$p = x_1 - x_0 \tag{4-1}$$

$$\omega = \frac{2\pi}{p} \tag{4-2}$$

of its variable x between $x_0$ and $x_1$. This function may be represented at any point over its range by a series of circular functions

$$f_p(x) = \frac{A_0}{2} + \sum_{n=1}^{\infty} (A_n \cos(n\omega x) + B_n \sin(n\omega x)) \tag{4-3}$$

where n is an integer.

$$A_0 = \frac{2}{p} \int_{x_0}^{x_1} f_p(x) dx \tag{4-4}$$

$$A_n = \frac{2}{p} \int_{x_0}^{x_1} f_p(x) \cos(n\omega x) dx \tag{4-5}$$

$$B_n = \frac{2}{p} \int_{x_0}^{x_1} f_p(x) \sin(n\omega x) dx \tag{4-6}$$

or

$$f_p(x) = \sum_{n=0}^{\infty} C_n \cos(n\omega x - \theta_n) \tag{4-7}$$

with $c_n$ and $\theta_n$ found from $a_n$ and $b_n$

or

$$f_p(x) = \sum_{n=-\infty}^{\infty} D_n e^{in\omega x} \tag{4-8}$$

$$D_n = \frac{1}{p}\int_{x_0}^{x_1} f_p(x) e^{-in\omega x} dx \tag{4-9}$$

where x is in an x-unit (in spatial case it should be a length unit), $\omega$ is an angular frequency in radiance per x-unit, $\theta_n$, is a phase shift angle, and $A_n$, $B_n$, $C_n$ & $D_n$ are amplitudes of the frequencies at $\omega_n = n\omega$.

When we choose $f_p(x) = \sin 2\pi x$, the waveform can be obtained in Matlab as shown in Figure 4.2 in addition, the corresponding spectrum is shown in Figure 4.3. Obviously, the period of $f_p(x)$ is p = 1 LU (Length Unit). Thus, frequency $f = \frac{1}{p} = 1$ $LU^{-1}$ or angular frequency $\omega = \frac{2\pi}{p} = 2\pi \frac{rad}{LU}$ and $b_1$ = 1 represent the only nonzero component in the spectrum.

The Fourier series of any periodic function may be represented in either the x-domain (spatial or time domain) as a function of f(x) or in the $\omega$-domain as a function of frequency $F(\omega)$.

$$f_p(x) = \sum_{n=0}^{\infty} F_n(\omega) \tag{4-10}$$

When the function is represented in x-space it is usually a reasonably continuous function in x. In $\omega$-space, the function is represented as an infinite series of amplitudes, $A_n$, $B_n$, $C_n$, or $D_n$, at discrete frequencies, $\omega_n$ , with discrete phase shifts $\theta_n$. The discrete frequencies are determined only by the period, T, of the periodic function. Each frequency contains a portion of the total energy or power in the function f(x). The total energy in the function f(x) is the sum of the amplitudes in each discrete frequency.

$$E_{f(x)} = \sum_{n=0}^{\infty} |C_n|^2 \tag{4-11}$$

or better

$$E_{f(x)} = \sum_{n=-\infty}^{\infty} \left| D_n e^{i(n\omega x + \theta_n)} \right|^2 = \sum_{n=-\infty}^{\infty} |D_n|^2 \,. \qquad (4\text{-}12)$$

The Fourier series moves a function back and forth between dual domains, frequency domain, $\omega$, and position domain, x, with different aspects of the function represented uniquely in each domain.

Shortly following the development of Fourier series of periodic signals the problem of non-periodic functions was addressed. It has been noted that Fourier series of periodic functions have discrete frequency content and an infinite sum of the discrete frequencies. The discrete frequencies are associated with the period as:

$\omega_n = n2\pi/T$; n= 0, 1, 2,3,...;

$\omega_0 = 0\times 2\pi/T$; Constant term, The non-oscillating term;

$\omega_1 = 1\times 2\pi/T$; the fundamental frequency, first oscillating term;

$\omega_2 = 2\times 2\pi/T$; the first harmonic, second oscillating term;

$\omega_3 = 3\times 2\pi/T$; the second harmonic, third oscillating term;

etc.

It can be seen, as the period, T, decreases, $\omega_1$ becomes larger and $(\omega_2 - \omega_1)$ grows larger, and the distance between $\omega_{n+1}$ and $\omega_n$ increases. As the period, T, increases, $\omega_1$ becomes smaller and $\omega_1$ and $\omega_2$ grow closer together. As, T, becomes very large, $\omega_1$ and $\omega_2$ move close together until, as, T, goes to infinity, the frequency becomes continuous. We have a non-periodic, f(x), in the space domain and a continuous function, $F(\omega)$, in the frequency domain. This leads us to a tool known as *Fourier analysis*.

### 4.1.2 Fourier Transform in One Dimension

A mathematical tool of great utility in the analysis of both linear and nonlinear phenomena is *Fourier Transform*. This tool is widely used in the study of electrical networks and communication systems.

Let us look at one-dimensional functions first. Any function, $f(x)$ (not limited to periodic function) may be considered composed of the superposition of a series of continuous periodic functions of suitable amplitudes and frequencies. Now a periodic function u can be represented by $F(u)e^{jxu}$, so that the original function, considered as a summation of periodic functions, becomes

$$f(x) = \int_{-\infty}^{\infty} F(u)e^{jxu}du\,. \tag{4-13}$$

The function

$$F(u) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-jxu}dx \tag{4-14}$$

which gives the amplitudes *F(u)* of the periodic terms of frequency u, is called the Fourier Transform of Equation (4-13). Whenever *f(x)* or *F(u)* can be determined, the other can be computed from the Fourier Transform relationship. This concept is of extraordinary power in many problems of analysis in optics as well as in various other fields.

The *Fourier transform*, in essence, decomposes or separates a waveform or function into sinusoids of different frequency which sum to the original waveform. It identifies or distinguishes the different frequency sinusoids and their respective amplitudes [6].

Since the Fourier transform *F(u)* is a frequency domain representation of a function *f(x)*, the *u* characterizes the frequency of the decomposed cosinusoids and sinusoids and is equal to the number of cycles per unit of *x* [7]. If a function or waveform is not periodic, then the Fourier transform of the function will be a continuous function of frequency [6].

### 4.1.3 Discrete Fourier Transform

Because a digital computer works only with discrete data, numerical computation of the Fourier transform of *f(x)* requires discrete sample values of *f(x)*. In addition, a computer can compute the transform *F(u)* only at discrete values of *u*, that is, it can only provide discrete samples of the transform.

Suppose that a continuous function $f(x)$ is discretized into a sequence

$$\{f(x_0), f(x_0 + \Delta x), f(x_0 + 2\Delta x), \ldots, f(x_0 + [N-1]\Delta x)\}$$

by taking N samples $\Delta x$ units apart. It will be convenient in subsequent developments to use *x* as either a discrete or continuous variable, depending on the context of the discussion. To do so requires defining

$$f(x) = f(x_0 + x\Delta x) \tag{4-15}$$

where *x* now assumes the discrete values $0, 1, 2, \ldots, N-1$. In other words, the sequence $\{f(0), f(1), f(2), \ldots, f(N-1)\}$ denotes any $N$ uniformly spaced samples from a corresponding continuous function. With this in mind, the discrete Fourier transform pair that applies to sampled function is given by

$$F(u) = \frac{1}{N}\sum_{x=0}^{N-1} f(x)\exp[-j2\pi ux / N] \tag{4-16}$$

for $u = 0, 1, 2, \ldots, N-1$, and

$$f(x) = \sum_{u=0}^{N-1} F(u)\exp[j2\pi ux / N] \tag{4-17}$$

for $x = 0, 1, 2, \ldots, N-1$.

In the two-variable case, the discrete Fourier transform pair is

$$F(u,v) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\exp[-j2\pi(ux / M + vy / N] \tag{4-18}$$

for $u = 0, 1, 2, \ldots, M-1$, $v = 0, 1, 2, \ldots, N-1$, and

$$f(x,y) = \sum_{u=0}^{M-1}\sum_{v=0}^{N-1} F(u,v)\exp[j2\pi(ux / M + vy / N] \tag{4-19}$$

for $x = 0, 1, 2, \ldots, M-1$, $y = 0, 1, 2, \ldots, N-1$.

These equations can be used to compute transforms and inverse transforms of appropriately-sampled data.

The discrete Fourier transform and its inverse are periodic with period N. Thus, only the N values of each variable in any one period are required to obtain f(x, y) from F(u, v). Or, only one period of the transform is necessary to specify F(u, v) completely in the frequency domain.

Let us consider the one-variable case. By translation property of discrete Fourier transform, the magnitudes of the transform values from (N/2)+1 to N-1 are reflections of the values in the half period to the left of the origin. Thus the actual highest frequency is N/2.

### 4.1.4 Fast Fourier Transform

The number of complex multiplications and additions required to implement Eq. (4-16) is proportional to $N^2$. Proper decomposition of this equation can make the number of multiplication and addition operations proportional to $N \log_2 N$ The decomposition procedure is called the *fast Fourier transform* (FFT) *algorithm*. The reduction in proportionality from $N^2$ to $N \log_2 N$ operations represents a significant saving in computational effort. The Fourier transform command in MATLAB is implemented this way as most other mathematical software is.

## 4.2 Frequency Analysis of Digital Images

After the initial development of Fourier series and Fourier transforms in one-dimensional space, it was clear that they could be extended to two-dimensional (2-D) space, three-dimensional (3-D) space, or to n-dimensions (n-d) space.

### 4.2.1 Optical Transfer Function

Optical engineers began to think about and develop a 2-D Fourier transform to calculate the Optical Transfer Function (OTF) of optical imaging systems. The

OTF is associated with the beginning of the field of study of diffraction optics. Williams and Becklund [12] define:

"The Optical Transfer Function **(OTF)**, is the frequency response, in terms of spatial frequency, of an optical system to sinusoidal distributions of light intensity in the object plane; the OTF is the amplitude and phase in the image relative to the amplitude and phase in the object as a function of frequency, when the system is assumed to respond linearly and to be space invariant."

The spatial frequency of the OTF is a measure of the spatial energy content of an optical image. As the spatial frequency increases the spatial energy of the image increases, as the spatial frequency decreases the spatial energy decreases. The 2-D Fourier transform, F(u, v), of an image calculates the spatial frequency content of the image. The 2-D transform is a complex amplitude and phase function in u, and v. The spatial energy in the image is

$$|F(u, v)|^2.$$

The complex optical transfer function (OTF) is made of two parts, the Modulation Transfer Function (MTF) and the Phase Transfer function (PTF):

$$F(u, v) = |F(u, v)| \angle F(u, v)$$

$$G(u, v) = |G(u, v)| \angle F(u, v)$$

$$G(u, v) = OTF(u, v) \bullet F(u, v)$$

$$OTF(u, v) = MTF(\omega)\, e^{\,iPTF(\omega)} \qquad (4\text{-}20)$$

$$OTF(u, v) = MTF(u, v) \angle PTF(u, v)$$

$$|G(u, v)| = MTF(u, v) \bullet |F(u, v)|. \qquad (4\text{-}21)$$

In order to change an image the MTF must change by at least 0.1 or 10% to get a recognizable change in the image [12]. We may think similarly that it will also require the object plane change by at least 10% to show a clear change in the image. More about this will be discussed in STV image analysis.

### 4.2.2 Spatial Frequency Filters

In communication engineering, the engineers are working with a broad spectrum of 1-D radio frequencies. To separate the frequencies into sets of high frequency signals and low frequency signals communication engineers have developed so called "filtering" techniques. The filtering techniques or filters are known as low pass filters, band pass filters, band reject filters, and high pass filters as illustrated in Figure 4.5.

$$F_O(\omega) = H(\omega)\, F_I(\omega) \tag{4-22}$$

$$|H_{\text{low pass}}(\omega)| = \begin{Bmatrix} 1, & \omega \le \omega_c \\ 0, & \omega > \omega_c \end{Bmatrix}, \quad |H_{\text{high pass}}(\omega)| = \begin{Bmatrix} 1, & \omega \ge \omega_c \\ 0, & \omega < \omega_c \end{Bmatrix}, \tag{4-23}$$

etc.

Low pass filters select for use, only frequencies below a specified cutoff frequency, $f_c$ in Hz or $\omega_c$ in r/s. High pass filters select for use, only frequencies above a specified cutoff frequency. A band pass filter selects frequencies between two specified frequencies, a low cutoff frequency and a high frequency cutoff. A notch filter discards a band of frequencies between a specified low frequency and a specified high frequency. All of the physically realizable filters for communications are low pass filters, with an absolute upper frequency cutoff. Given the absolute high frequency cutoff, all the remaining realizable communications filters are band pass filters with maybe a few notches in the pass band. The signals are all called band limited signals due the physical nature of the devices.

Spatial frequency filters for two-dimensional images can be extended from the above. For example, 2-D low pass filters can be defined as following:

$$\left|H_{lowpass}(u,v)\right| = \begin{Bmatrix} 1, D(u,v) \le D_0 \\ 0, D(u,v) > D_0 \end{Bmatrix} \tag{4-24}$$

where $D_0$ is a specified nonnegative quantity, and D(u, v) is the distance from point (u, v) to the origin of the frequency space; that is,

$$D(u,v) = \sqrt{u^2 + v^2}\ . \tag{4-25}$$

When we investigate an optical system, it is readily apparent the optical system is also band limited. There are two reasons for the optical systems to be band limited. The resolution of the optical system is limited due to imperfections of the imaging system, and the wavelength of the light illuminating the image has diffraction limits. The diffraction of the optical system is due to the optical imaging system having a finite surface area and a finite light gathering capability associated with the optical aperture. The diffraction limit of the illumination source is associated with the wavelength of the incident light, short wavelength high energy light, blue, violet, ultraviolet, x-rays and above are capable of imaging small object with dimensions approximately 5 or 6 times the wavelength of the incident light. Long wavelength light, orange, red, infrared, and longer will image gross features of an object. As a result red light cannot image sharp edges while blue light, UV-light, and x-rays can image sharp edges. So the imaging system is band limited and acts as an optical low pass filter and visible light is band limited and also acts as a low pass filter. Usually the optical system is the dominating effect and has a much lower cutoff than the diffraction limit of the light.

Digital video camera has very limited band due to the structure of the CCD chip it uses for imaging. However, we will show later that the cutoff of our digital video camera is much higher than that of average human eyes.

### 4.2.3 Frequency Domain Representation of Digital Images

Physical laws suggest that any conceivable object that can yield an image may always be represented by a series or by a simple or multiple Fourier integral. The amplitudes of the terms of the series or the integrand of the integral usually can be regarded as describing the spatial frequencies, which leads to a complete representation of the same object in a different domain rather than the spatial. There have been plenty of great examples in contemporary optics and mathematics.

It is often useful to think of functions and their transforms as occupying two domains. These domains are referred to as the upper and the lower domains in older texts, "as if functions circulated at ground level and their transforms in the underworld" [7]. They are also referred to as the function and transform domains, but in most physics applications, they are called the time and frequency domains respectively. Operations performed in one domain have corresponding operations in the other. For example, the convolution operation in the time domain becomes a multiplication operation in the frequency domain, The reverse is also true. Such theorems allow one to move between domains so that operations can be performed where they are easiest or most advantageous.

The earlier contents of this section prepare us for the most often used transform in our research, the two-dimensional fast Fourier transform on a discrete matrix (the digital image). The result of this transform is a discrete matrix whose elements represent the frequency domain amplitudes.

### 4.2.4 Using MATLAB

As a powerful mathematics software package, MATLAB provides very convenient function to perform the two-dimensional discrete Fourier transform and inverse, respectively FFT2 and IFFT2. The following example will show the transform of a 128-by-128-pixel picture at the center of which there is a white square hole in a black background (Fig. 4.6 shows its 3-D view) into its frequency space counterpart.

Using the command **FFT2** transforms the real space 2-D matrix into a complex 2-D matrix that is not as easily understood. Fig. 4.7 shows the real part of each element of the matrix, while Figure 4.8 shows the modulus.

Occasionally, people are also interested in the square of the modulus as power spectrum.

Let us look at Figure 1.4 again. The Fourier transform of this image is shown in the frequency domain (Figure 4.9).

It is common to illustrate Fourier spectra of images as intensity functions. The dynamic range of Fourier spectra usually is much higher than the typical display method is able to reproduce faithfully, in which case only the brightest (largest values) parts of the spectra are visible on the display screen. A useful technique that compensates for this difficulty consists of displaying the function

$$D(u,v) = c\log[1+|F(u,v)|] \tag{4-26}$$

instead of $|F(u,v)|$, where c is a scaling constant and the logarithm function performs the desired compression. Use of Eq. (4-26) greatly facilitates visual analysis of Fourier spectra, as shown in Figure 4.8 and 4.9.

### 4.2.5 Partial Volume Under 2D FFT Surface

Initially, we started doing FFT on an image, getting the absolute value for each element of the FFT matrix and drawing it in the 3-D view. The results look much alike. First, there is an extremely tall spike representing the DC component at the center of the frequency plane. Then, there is the mid-low-frequency range in which usually much more common height spikes can be found comparing to the high frequency range. Usually in the high-frequency range, there is nothing but relatively flat component representing noise. However, for some highly periodically featured pictures, there could be some unique features shown in this range. We try our best to investigate hoping to find something useful.

#### 4.2.5.1 The Idea

We assume that the volume under certain frequency range will tell us the amount of information in which we are interested. We followed the following steps to find out.

1. Prepare two images into black and white ones with 256 gray levels.
2. Subjectively judge which of the two has more amount of valuable information.
3. Find the size of the image.

4. Take the FFT of the image.
5. Divide the area of the 2-D FFT into equally spaced rectangular rings, shown in Figure 4.12, or equally spaced circular rings, shown in Figure 4.13.
6. Sum the magnitudes of the FFT pixels inside each ring to get the volume of that ring.
7. Import the sums to Microsoft® Excel 97.
8. Ignore volume of the center block.
9. Try combining rings in different ranges to analyze the relation between the partial volume under FFT surface and amount of valuable information according to the subjective result.

It should be noted that in step 5, we choose to use fixed $\Delta u$, $\Delta v$, or $\Delta D$. The direct result of this choice will be the linearly increase of ring areas as the ring number increases. We are comfortable with the possible effect of this behavior on the results

### 4.2.6 Circular Division or Rectangular Division?

Now let us discuss which method we should use to make the division in the discrete Fourier transform. There are two choices, circular division and rectangular division. The DFT given by the software is in the form of a matrix, usually a two-dimensional one for a two-dimensional image. From this point of view, the rectangular division is sort of 'intrinsic', and thus easy to compute. On the contrary, to make a circular division, we have to locate the center of the matrix, calculate the distance between this center and each element of the matrix. With the distance falling into our pre-determined grids, we then say this element belongs to group n instead of the next groups. Obviously, there will be much more computing involved. On the other hand, we are looking forward to find the distribution of the Fourier transform along the modulus frequency without regarding any direction. By this standard, circular would definitely be the choice. Thus, for a large size image

with more than 400 by 600 pixels, we will use the circular division as the better option. However, in situations where the images' sizes are less than 100 pixels, we choose to use rectangular division as the only choice because the difficulty in dividing and the deviation brought by the circular division in such small matrices.

### 4.2.7 Excluding the DC Component

Why in step 8 we should ignore the volume of the center block?

After examining the equations of 2-D discrete Fourier transform, it is obvious that the center of the frequency domain is the result of summing all the gray-level values of every pixel in the picture. It describes how bright the picture looks from a far distance rather than describing any feature of the picture that is detailed. Here we can borrow the electrical term, 'DC component', to refer to this value. Usually when a large size picture is involved, this 'DC component' tends to get extremely large comparing with other frequency components. What we want to make sure is that this value will not get in the way of our examining other potentially useful values. Our way of doing this is to simply ignore the center block of our divisions.

### 4.2.8 Defining the Frequency Range

Now that our available frequency domain rings start only at the second, there comes the question, 'Exactly what frequency range should we look at?'

Common sense tells us that in the high frequency range, noise usually takes up most of the activity instead of the real image. Nevertheless, at first we will not exclude any other frequency range. We will see some reasoning to take out some part of the frequency range in the chapter that deals with the STV. For right now, we will include every ring starting with the second ring and watch the result carefully.

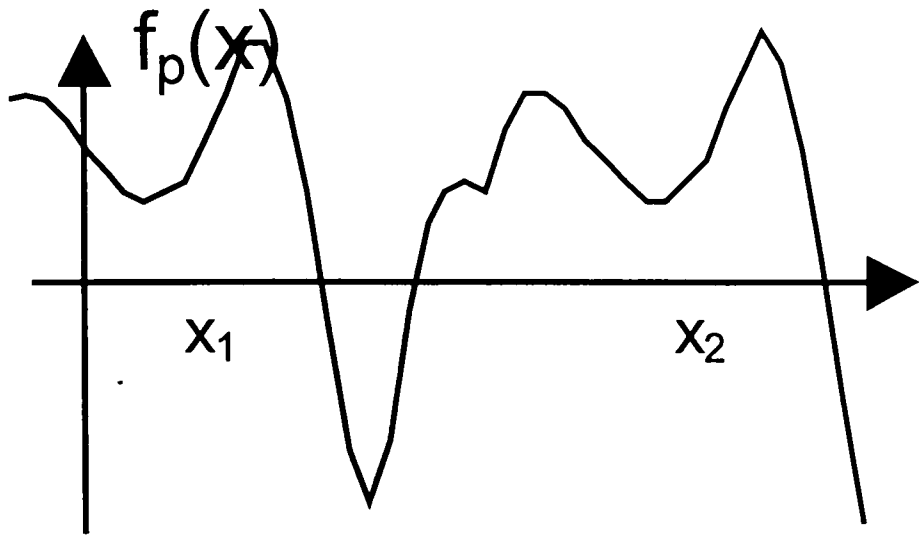



Figure 4.1: Figure for $f_p(x)$

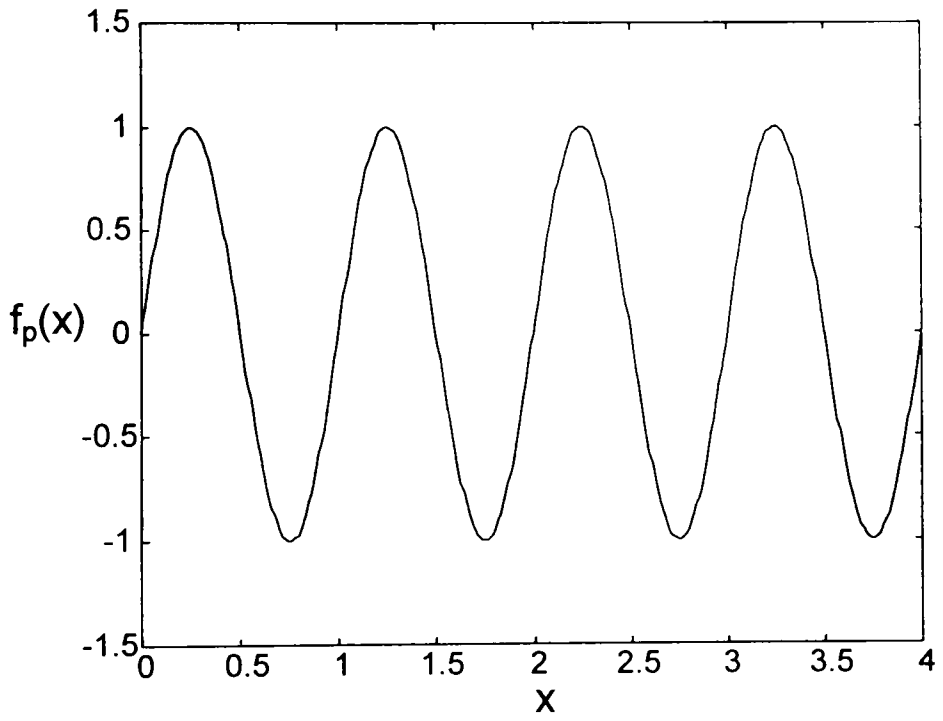



Figure 4.2: A single frequency wave. ( $f_p(x) = \sin 2\pi x$ )

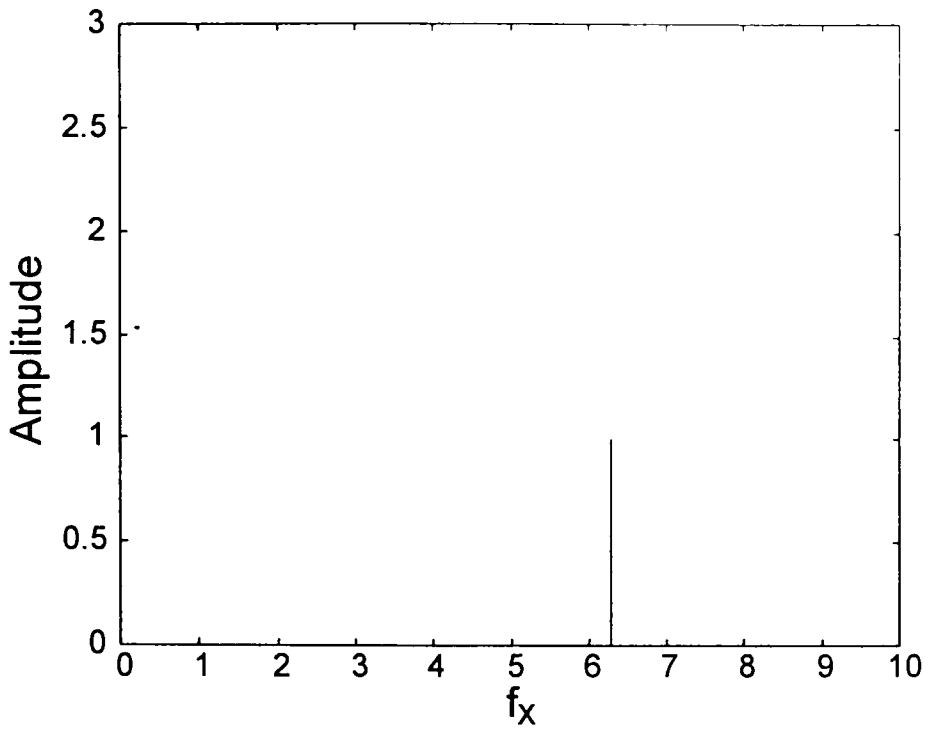



Figure 4.3: The spectrum of the single frequency wave shown in Figure 4.2.

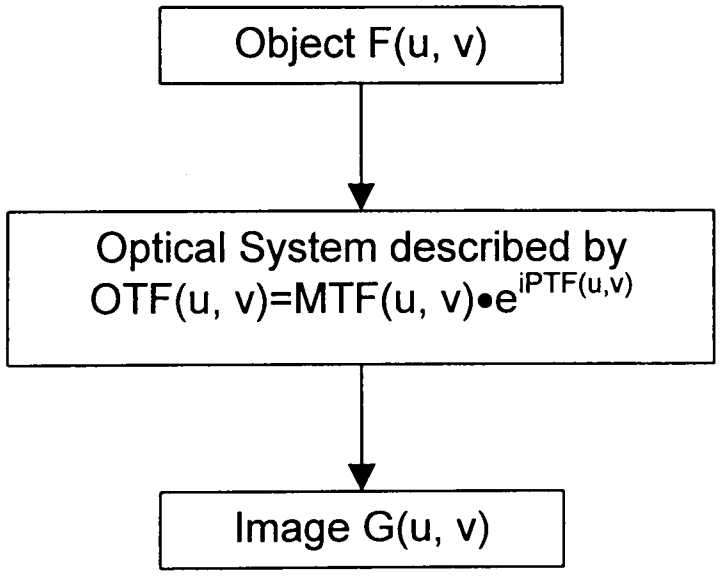



Figure 4.4: Illustration of OTF

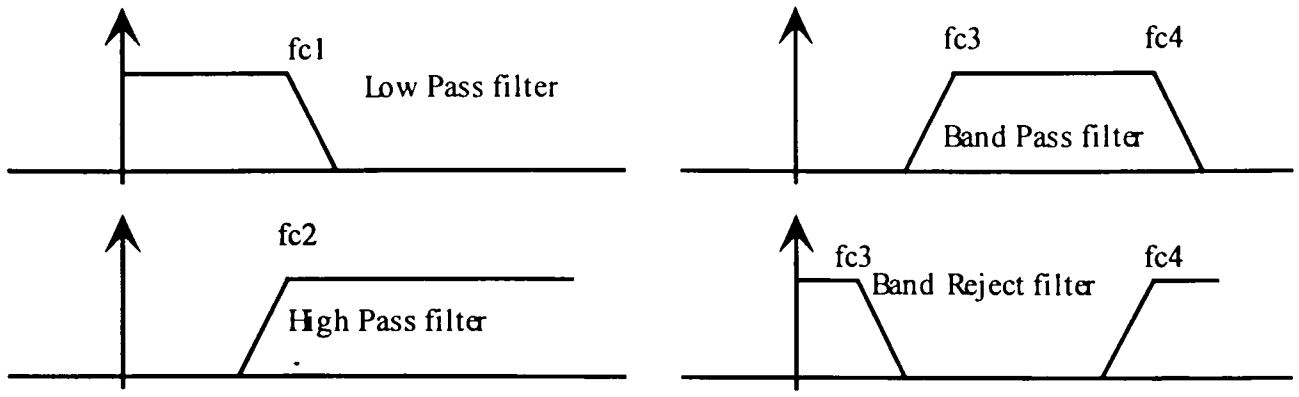



Figure 4.5: Sketches of four filters

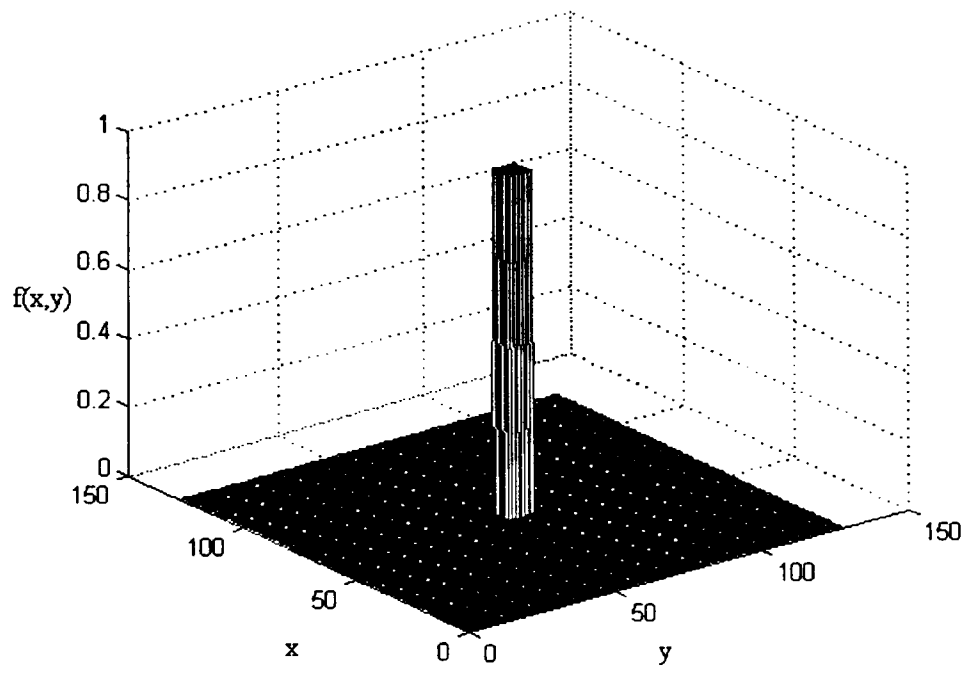



Figure 4.6: 3-D view of a square hole

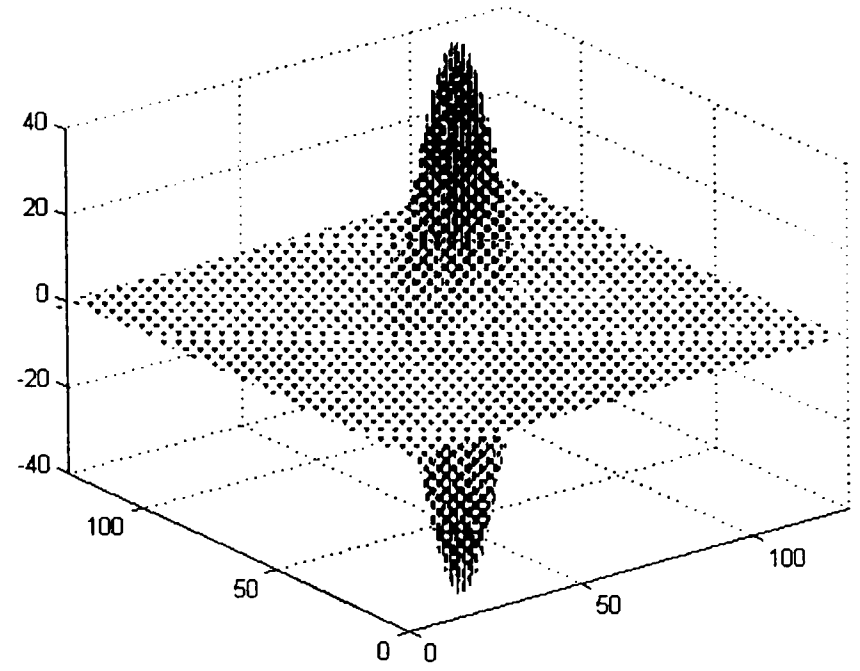



Figure 4.7: 3-D plot of the real parts of the square hole's FFT.

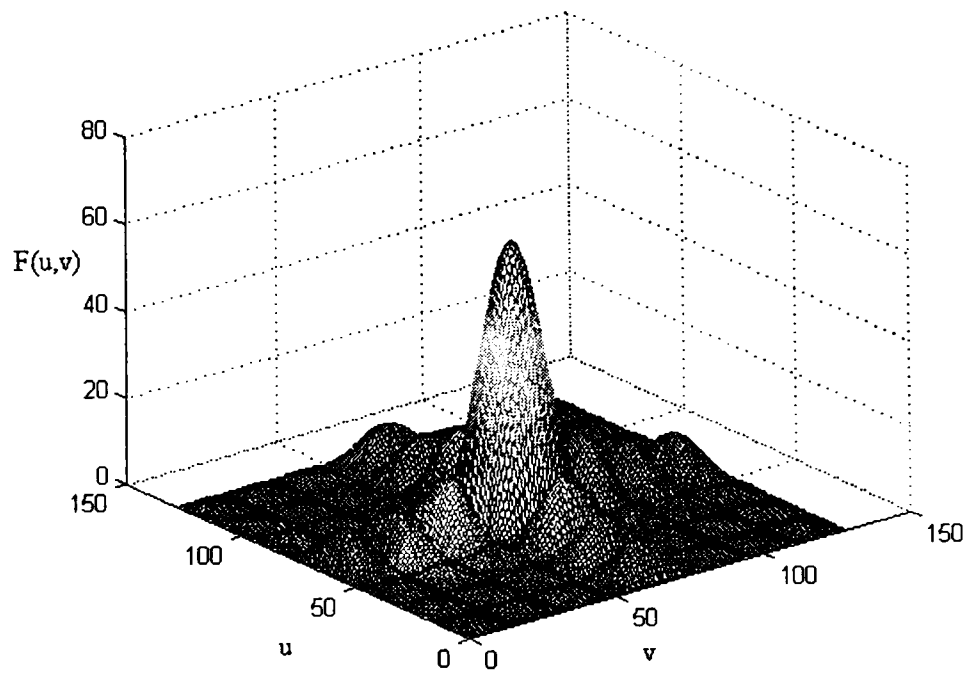



Figure 4.8: 3-D plot of the absolute values of the square hole's FFT.

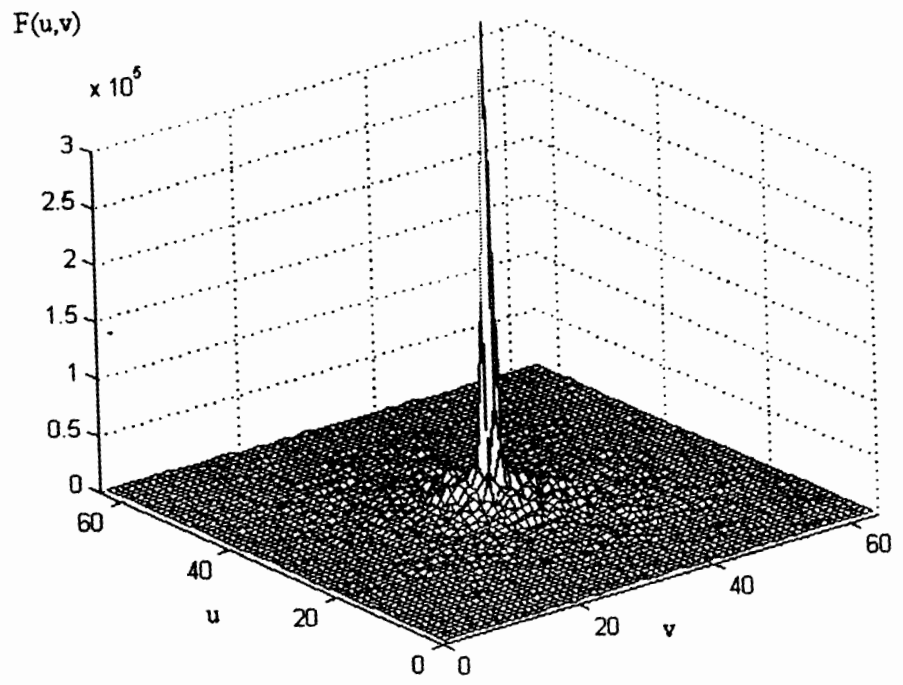



Figure 4.9: Absolute FFT result of Figure 1.4.

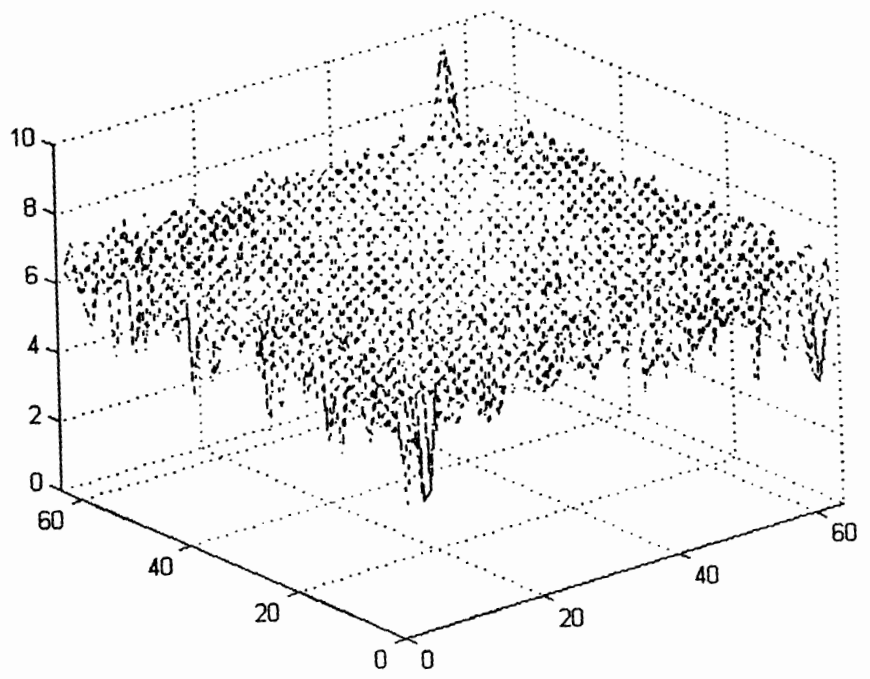



Figure 4.10: Log scaled 3-D view of Figure 4.9

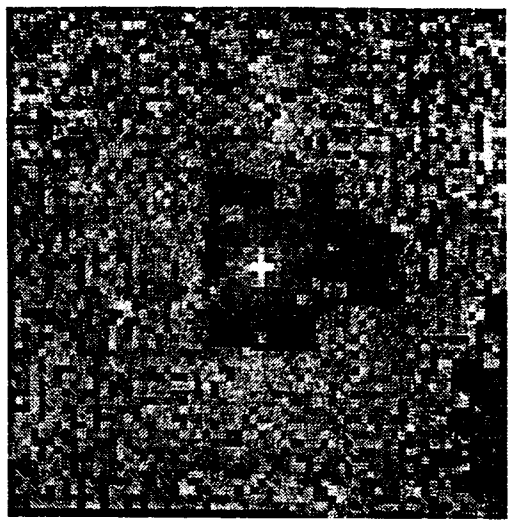


Figure 4.11: 2-D intensity illustration of the log scaled Fourier spectrum

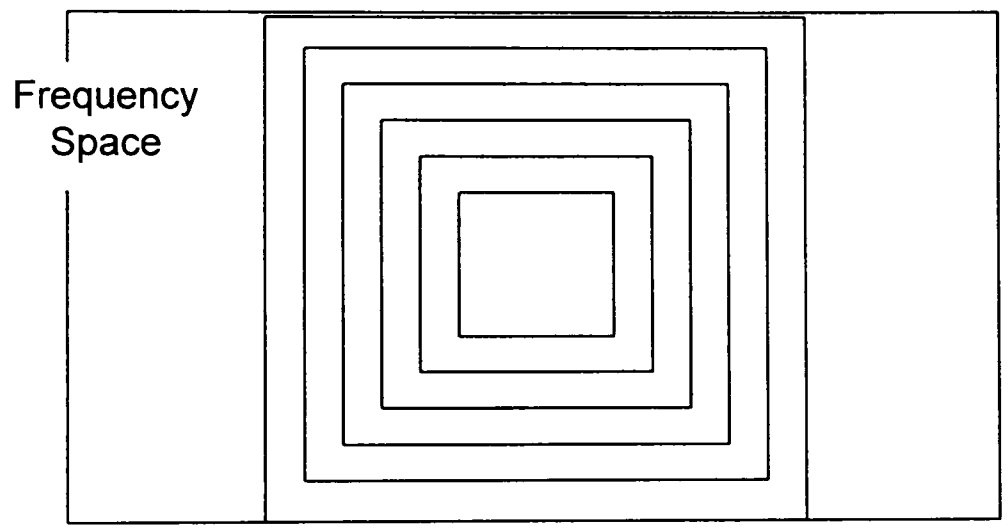



Figure 4.12: Square band division

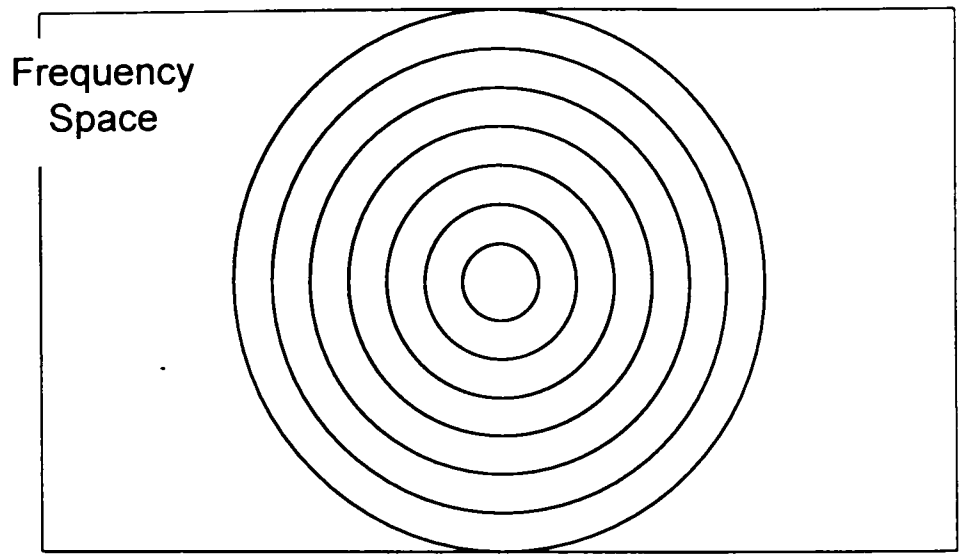



Figure 4.13: Circular band division

# CHAPTER V
# RIFA IMAGES ANALYSIS

## 5.1 Necessity of Fire Ant Imaging

### 5.1.1 Introduction

The red imported fire ant (*Solenopsis Invicta* Buren; Hymenoptera: Formicidae) or RIFA is native to South America. It was imported into the Mobile, Alabama, area between 1933 and 1941. Being extremely successful, RIFAs have spread throughout the southeastern part of the United States. The RIFA causes medical problems for humans, endangers domestic animals and wildlife species, is an agricultural pest, and invades both electrical and electronic equipment.

RIFA has several attributes that have enabled it to invade all of the southeastern and part of the southwestern United States among which are high fertility and fecundity rates, omnivorous diet and opportunistic behavior. Pathogens and parasites have failed to follow the RIFA from its native areas, and the RIFA's ability to sting allows it to compete and repel larger vertebrate competitors from resources.

One of the problems caused by RIFA is their invasion of electrical equipment, such as the transformer boxes of utility companies. It has been reported that researchers have found attraction of RIFAs to electrical fields that could be caused by magnetite in the RIFA. (Slowik, thesis, manuscript 1996)

### 5.1.2 Transition from Qualitative to Semi-quantitative Measurement

We want to help entomologists quantify their observations. First, we would like to provide a way for them to count the number of ants in one picture by clicking a computer mouse instead of killing the whole colony and count by hand. Then we are interested in finding a way to semi-quantitatively measure the trend of the ant activities with the combination of the video equipment in the field and the

computing power back in the laboratory. Traditionally, such jobs have been done qualitatively on a scale of 1 to 5 representing by expert entomologist going out to the field and using their experience and judgment. Of course, it will need great effort to standardize the video shooting of the ant activity footage in the field.

### 5.1.3 Develop a Non-invasive Method to Quantify Ant-transformer Activity.

Another benefit of the successful development of such a method will be that no physical change or harm will be done to the ants throughout the duration of measurement. We believe that entomologists will benefit from having the colonies consistently providing valuable data as time evolves.

After some observation and experiment, we learned that there is a rising activity period during which the ants become alert due to outside stimulation, a high activity period during which the ants actively search for enemies, and after a long term of no more stimulation, a slow return to the normal activity level.

An ideal solution is sketched based on our understanding of the problem. Instead of an expert, we will have someone inexperienced in evaluating the population of a RIFA colony to perform a hammer test on the transformer box and take video of the ants coming out of the gaps between the transformer box cover and the body. If we can determine the number of ants shown in each frame of the digital video, we may determine the rate of increasing and decreasing, which will be helpful to describe the size of the whole colony.

In this chapter, the author will introduce some results from an experiment performed in the lab of the Plant and Soil Science Department, Texas Tech University to check the feasibility of this idea.

## 5.2 Preparing the Fire Ant Images

After the digital video is taken in the field and brought back to the laboratory, it is necessary to make some pre-treatment to prepare the video clips into clear difference-images that only interesting information is kept. Our digital video clips taken in the lab will be treated as clips from field.

### 5.2.1 Timing the Digital Video Images

We download the digital video clips from the videocassette onto the hard disk of our computer as described in Chapter II. Opening the 'avi' files in Adobe Premier 4.2 allows us to 'export' any frame as a Bitmap picture. It also allows us to utilize the timing system that comes with the digital video taking. With that, we can extract frames at regular time intervals. In our first attempt, we took $15/30^{th}$ of a second, or 0.5 second as our time interval. The reason behind our picking such a length is that only after such a period, there are noticeable movement taking place between the two consecutive frames. The picture files are also indicated by numbers based on the sequence in which they are extracted from the clip.

### 5.2.2 Taking the Difference of Two Digital Images

We used several images from a RIFA video clip to test the idea introduced above. The video clip was taken in a lab of the Plant and Soil Science Department, Texas Tech University. The environment in which the RIFA lived was simple, a white tray with some soil and supplies of food and water. Later, we exported several images from the clip using the hardware and software introduced in Chapter II. Several typical images are shown below. Figure 5.1 and Figure 5.2 are 0.5 second apart, and so are Figure 5.3 and Figure 5.4. The time interval between Figure 5.1 and Figure 5.3 is two seconds. It is important to mention here that some images have the white tray as their background, whereas, in other images soil and occasionally even grass leaves are the backgrounds, especially those taken from field video clips.

We observed that in these images, most of the ants moved around rapidly while at the same time the background either did not change much or changes were relatively slow. Since we are only interested in finding the number of ants, we expected that differences between two consecutive images would simplify the problem. The resulting image would remove anything that shows up in the exact same place in both images, including most of the background and some motionless ants (most of the times only the dead ones stay still). Obviously, twice as many ant shapes usually appeared in the difference image as in either of the original images. Except some ants in the later image took the locations of some ants in the earlier one. These possibilities increased as the ant density in the image increased.

Fig. 5.5 shows the gray level result of the subtraction of Fig.5.4 from Fig.5.3. It is very exciting to see the clear shapes of the moving ants. It almost invites the observers to count by hand. Nevertheless, it is not so easy in a more complex environment with many ants moving around. Fig. 5.6 shows the same kind of subtraction of images 2 seconds later.

### 5.2.3 Simple Enhancement-Intensity Value Mapping

Series of Images acquired in this way were grouped for calculation using the method of frequency component analysis. By performing some simple enhancement techniques, such as intensity mapping, the vague images can be transformed into much clearer ones (see Figures 5.7 and 5.8).

## 5.3 On-Screen Counting

### 5.3.1 'Clicking' the Ants

Upon the request of an entomology graduate student, a GUI Matlab script was developed to achieve the on-screen hand-counting of RIFA numbers in still images. A user was able to point the mouse to and single-click at any place in the digital image on the computer screen (Figure 5.9). The number of clicks made was

output on a menu at the end of the program (Figure 5.10). It is still a tedious job to make all those clicks. Nevertheless, automatic counting makes the life of an entomologist a little easier.

### 5.3.2 Difference Images Make a Difference

An easy progress can be achieved by using the difference images instead of the original images. Without the background, the user can concentrate on just any white area without having to judge whether the thing he sees is an ant or not.

## 5.4 FFT Analysis of RIFA Images

The FFT calculations are carried out on two sets of RIFA images. The author also used the ant number count to divide all the images within one set into different groups. Each group included certain numbers of images that contained the same numbers of ants. Then the average was obtained among FFT calculation results of the images that belonged to the same ant number group.

### 5.4.1 Images of Less Than 10 ants and Picking the Frequency Range

First, let us look at a graph (Figure 5.11) showing the frequency component distribution of 51 images along the frequency magnitude. Each curve in the graph represents one of the images. Every dot of one curve represents the sum of the absolute frequency component in a certain frequency segment. In another way to state it, the sum of all the absolute FFT matrix elements between two adjacent rings.

With no assumption, the author first found the sum of the values represented by the dots of each curve. Next, the author plotted these summations against the corresponding ant number as shown in Figure 5.12. There is an up-going trend as the number of ants in the image increases. To see it more clearly, the author averaged the summations of the same population group and ploted the

averaged summation against the ant number (Figure 5.13). The result is not very good.

However, looking at Figure 5.11 suggests that it might give better result if the author calculates the summation of the FFT distribution in a narrower range. It turns out better (see Figures 5.14 and 5.15). Figures 5.14 and 5.15 were done by the circular method and the selected range was from ring 2 to ring 10. The need for a value that reveals the population in images can be better served by concentrating on the low frequency range rather than the whole spectrum.

### 5.4.2 Summary of Difficulties

The major problem the author encountered was that there was not sufficient number of pixels for each RIFA in several series of images. When the images were taken in the field, the camera was unable to get a close look at the scene that resulted in the insufficient numbers of pixels for RIFA. When this happened, the processed images only showed some non-uniform and irregular shapes that can not be well recognized.

The following are other problems encountered in the image acquisition and processing:

1. There was a color similarity between ants and soil when taking video in the field. Under normal lighting, they were hard to distinguish, especially in still images.
2. It was hard to stabilize the mounting of the video camera. The tilt and span can be easily affected by wind or ground shaking.
3. In natural lightingm, it was not easy to maintain constant lighting on different trials. Shade of the nearby trees or clouds was recorded in some images.

### 5.4.3 Conclusion

This chapter demonstrates the following:

Firstly, the approach is unique and sound mathematically.

Secondly, the result shown gives us confidence in furthering our digital image approach to discover a more reliable and universal way to reveal the number of the sampled ants in the camera screen.

Thirdly, this method puts strong requirements on the image acquisition, such as fixed setting of focal length of the camera, fixed pan and tilt of camera, constant lighting, stable camera mounting, same stimulation to the RIFAs, etc.

Figure 5.1: Image 1 taken at t = 0 (a11.bmp).

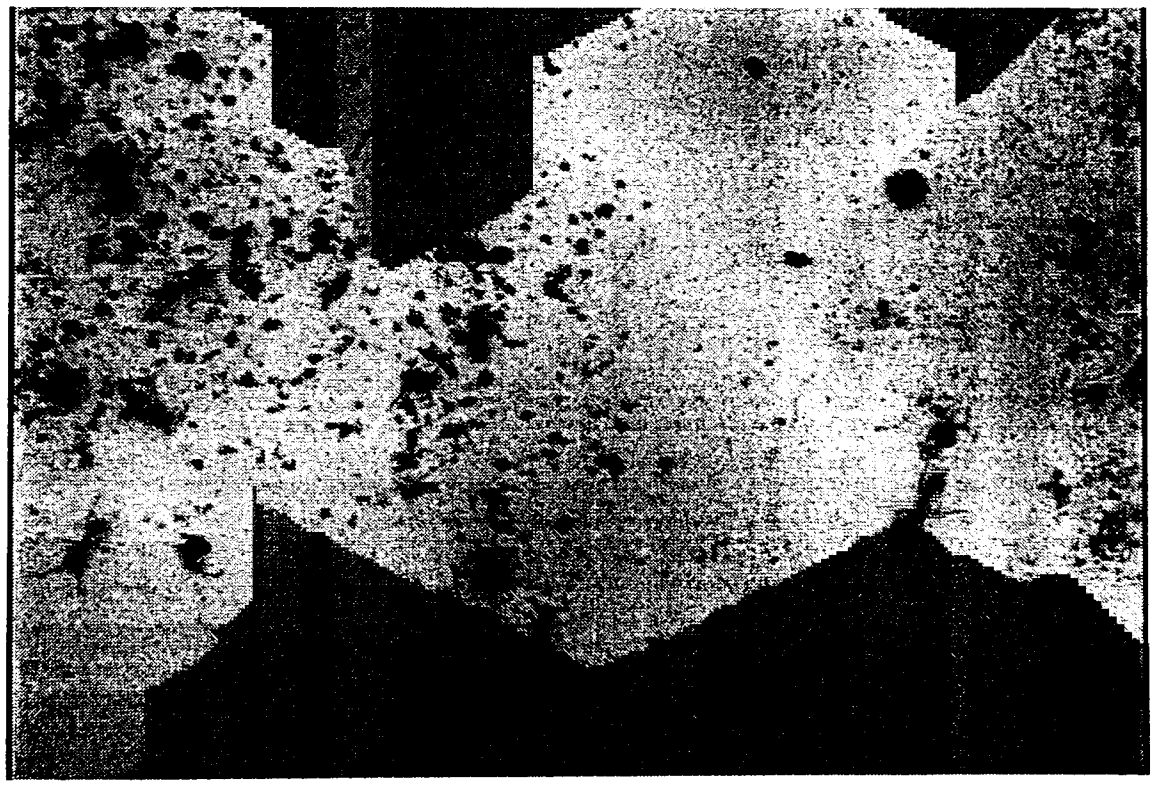


Figure 5.2: Image 2 taken at t = 0.5 sec (a12.bmp).

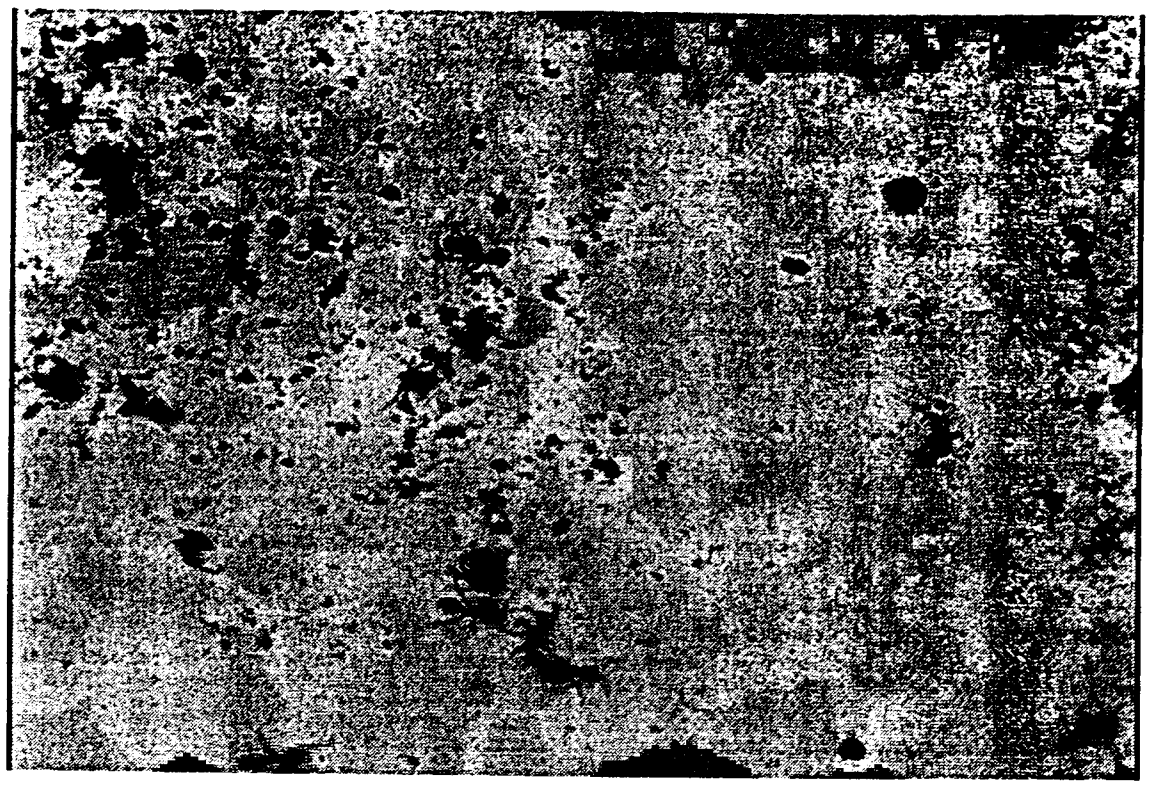


Figure 5.3: Image 3 taken at t =2.0 sec (a15.bmp).

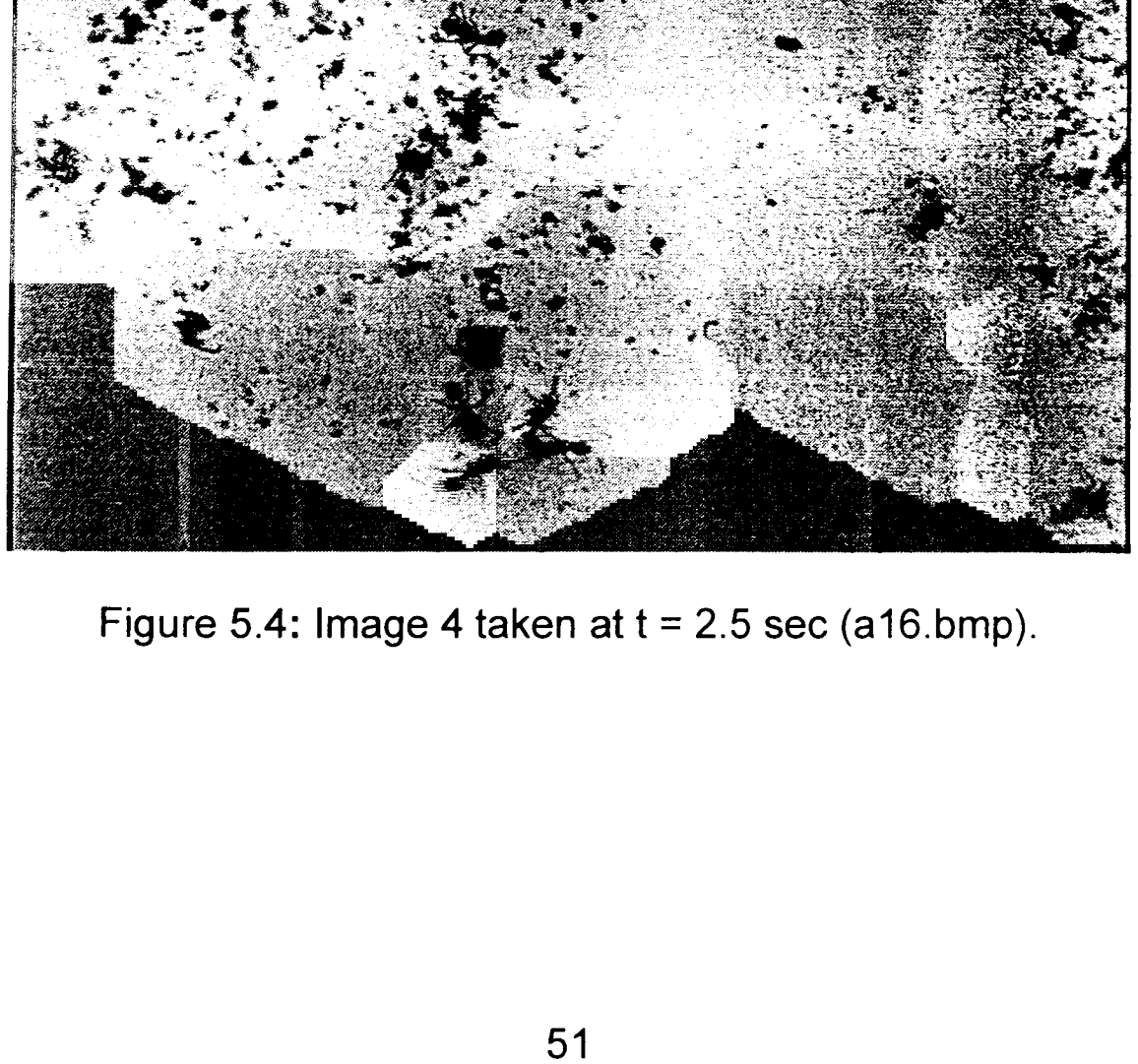


Figure 5.4: Image 4 taken at t = 2.5 sec (a16.bmp).

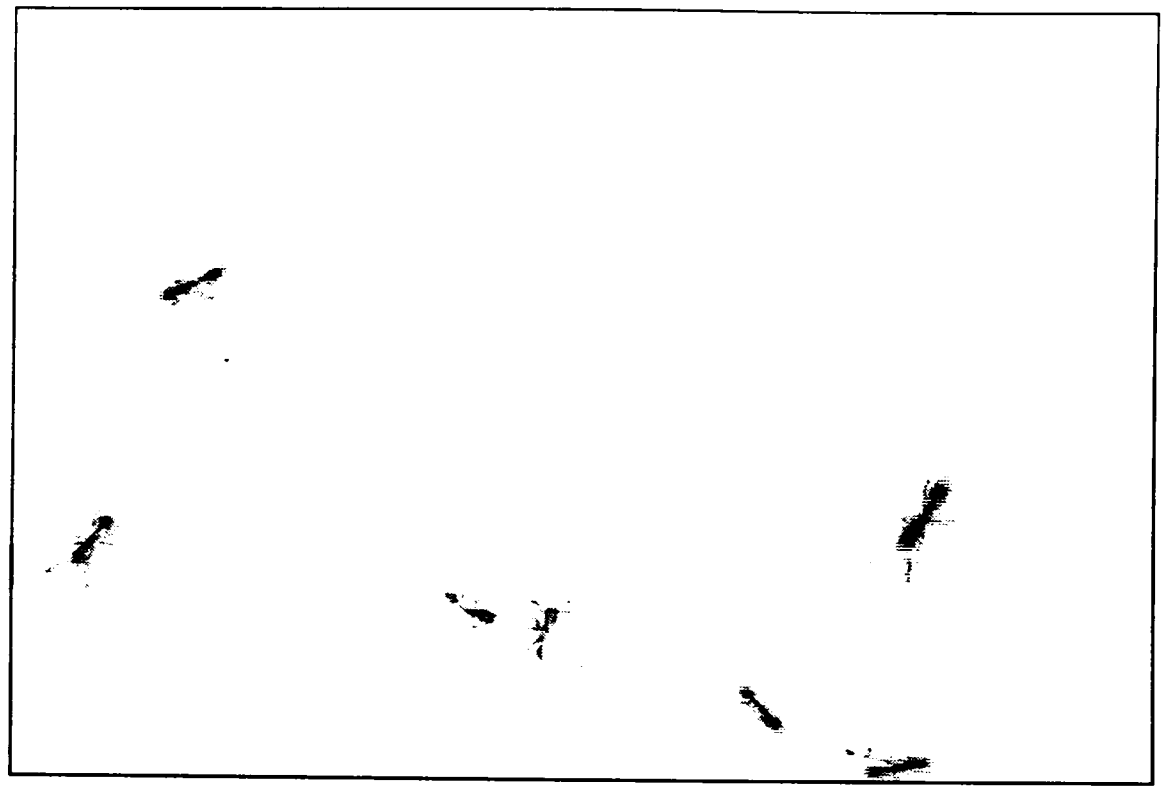


Figure 5.5: Image 1 with background removed
(black and white inverted).

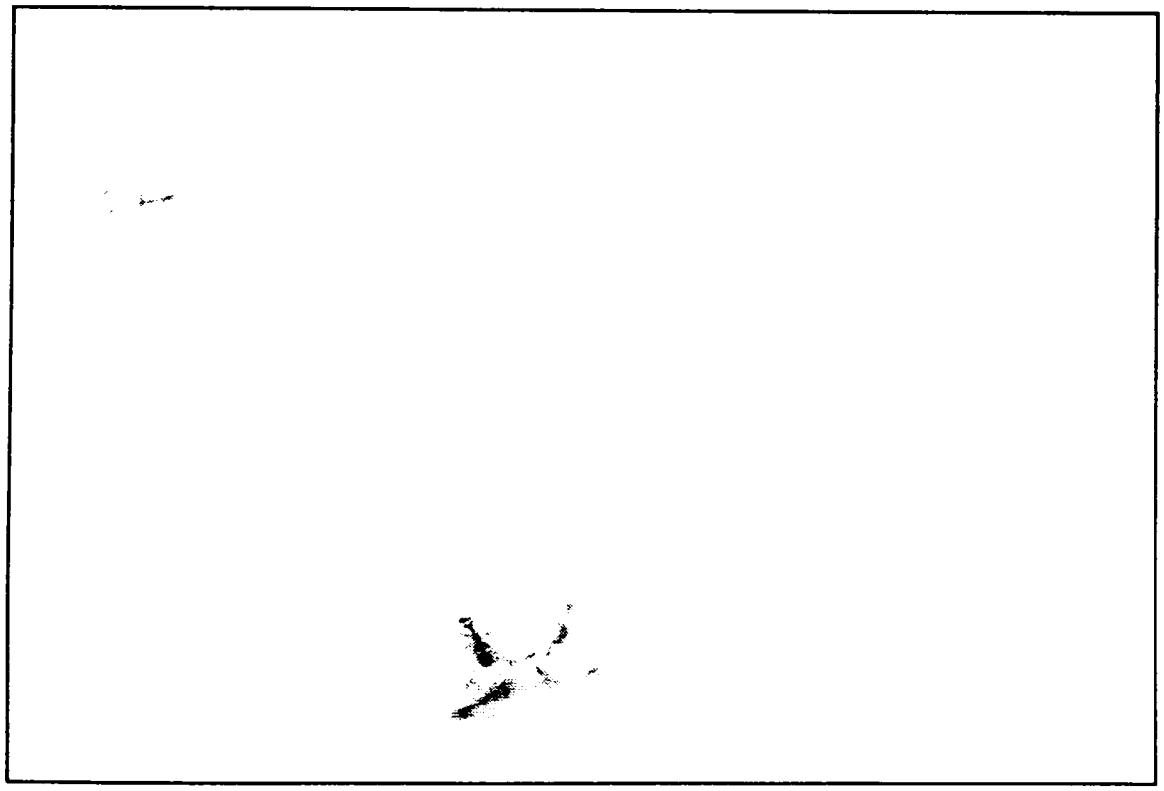


Figure 5.6: Image 3 with background removed
(black and white inverted).

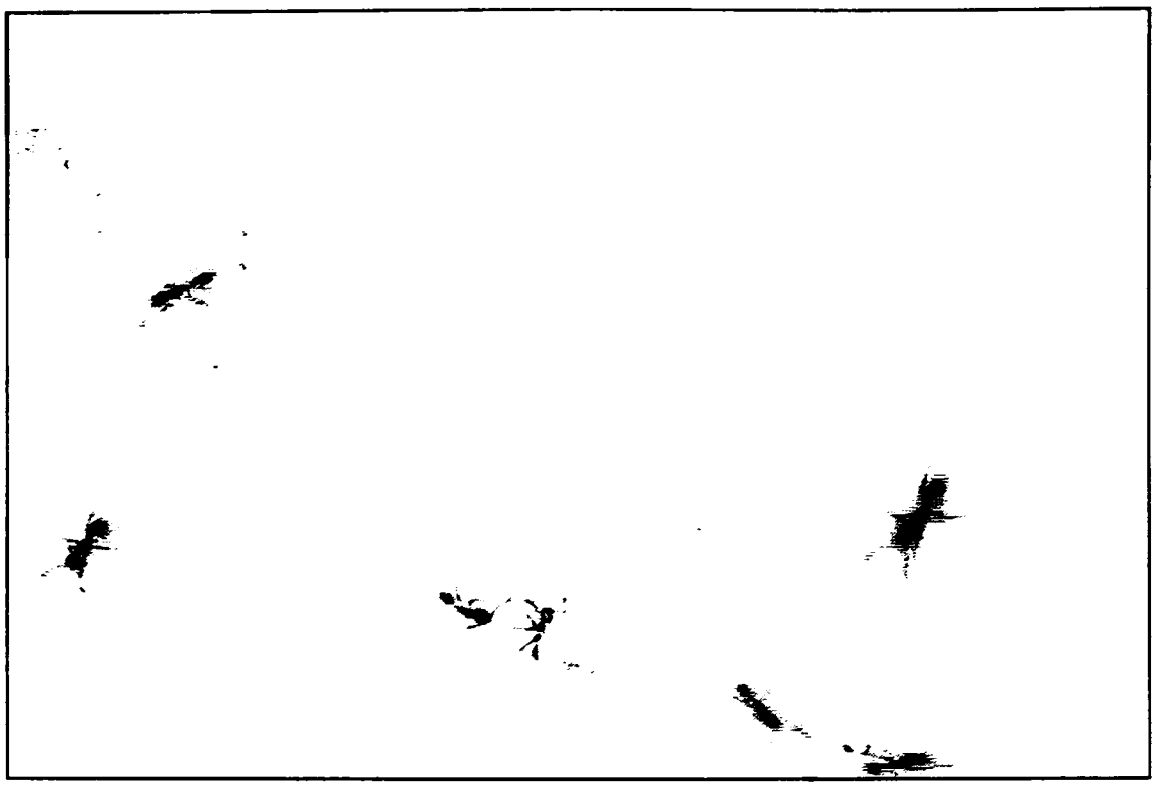


Figure 5.7: Enhanced Figure 5.5.

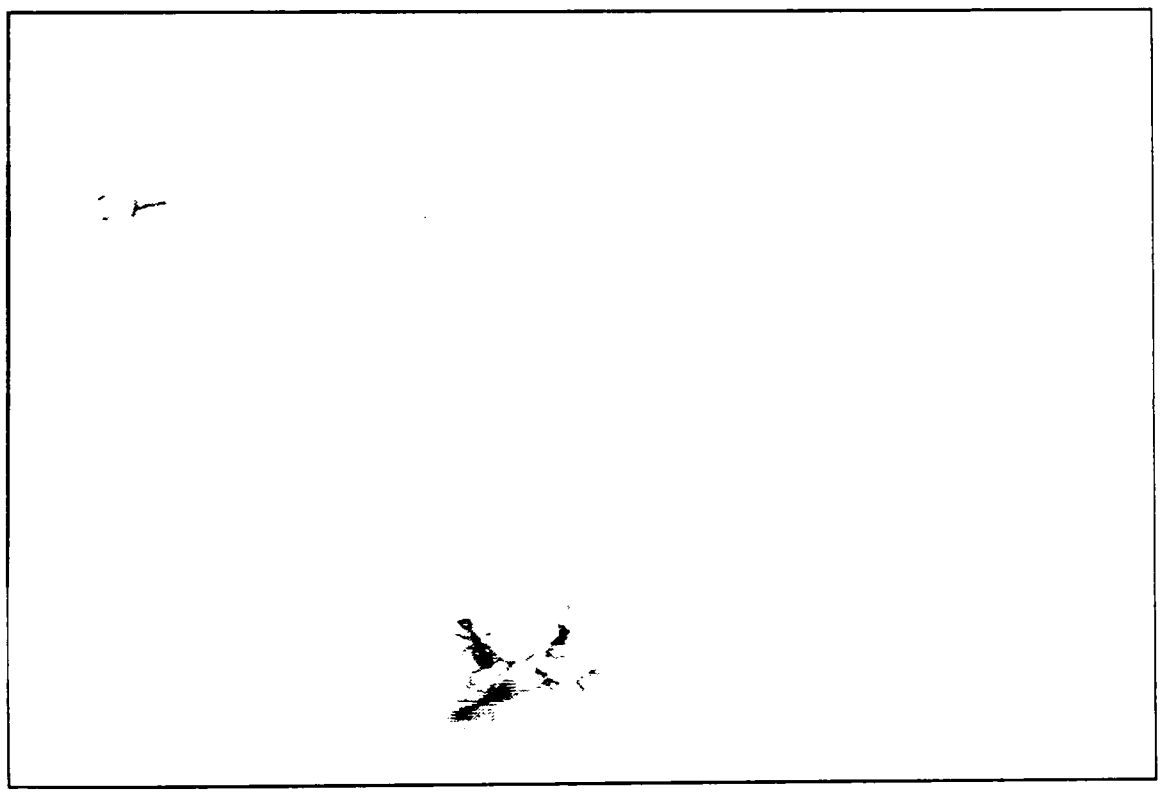


Figure 5.8: Enhanced Figure 5.6.

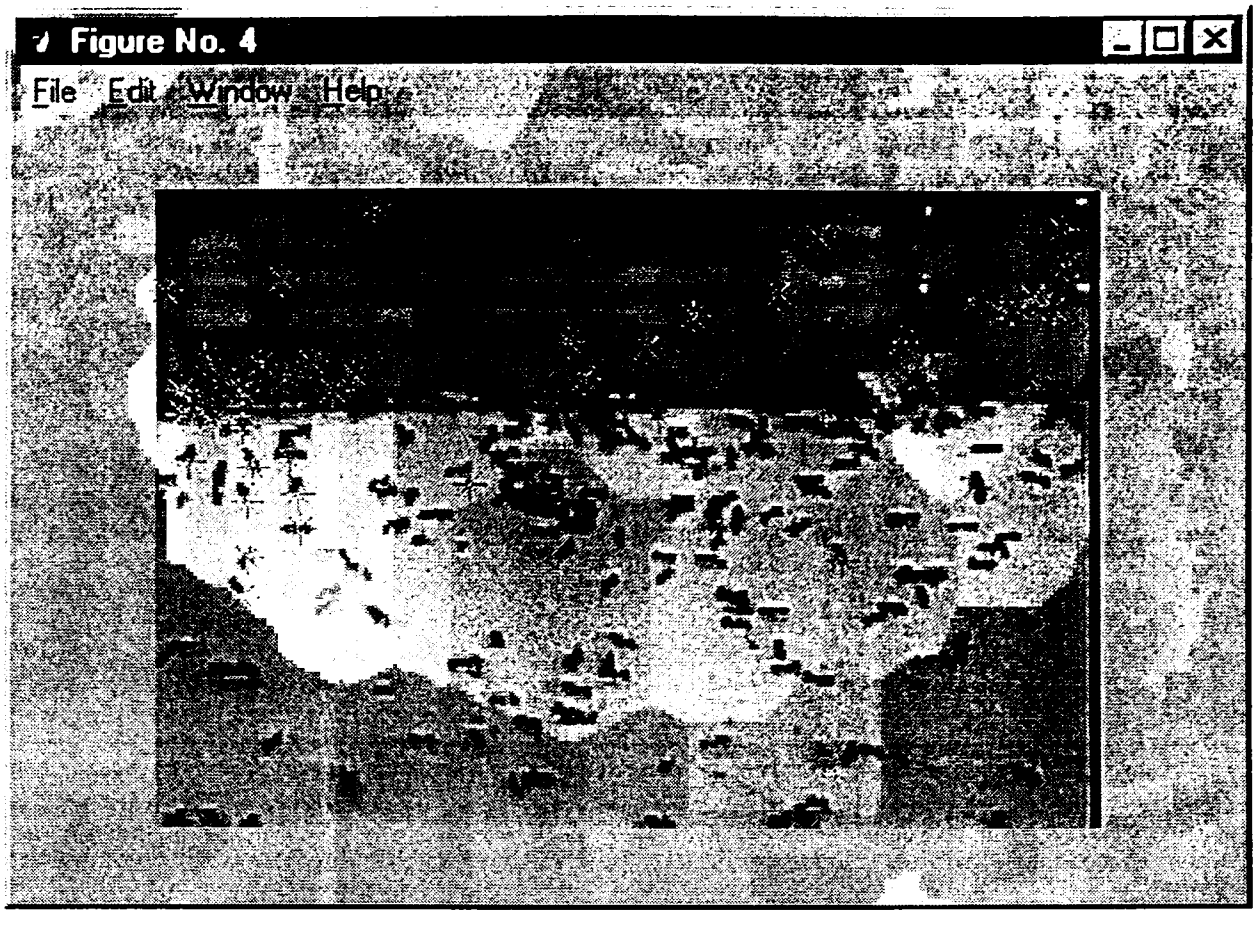



Figure 5.9: Click-counting the original.

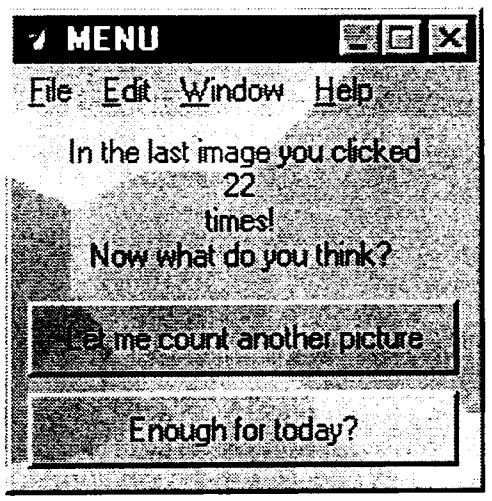



Figure 5.10: The counter tells the number of clicks.

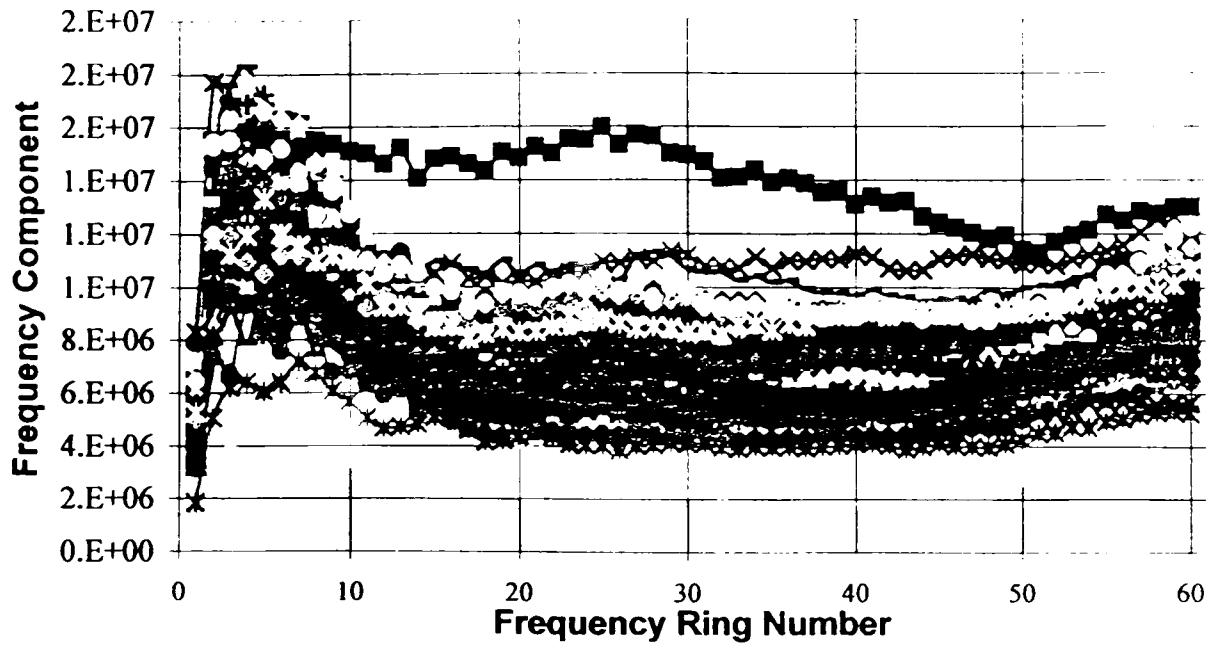



Figure 5.11: Frequency component distribution of 51 ant images.

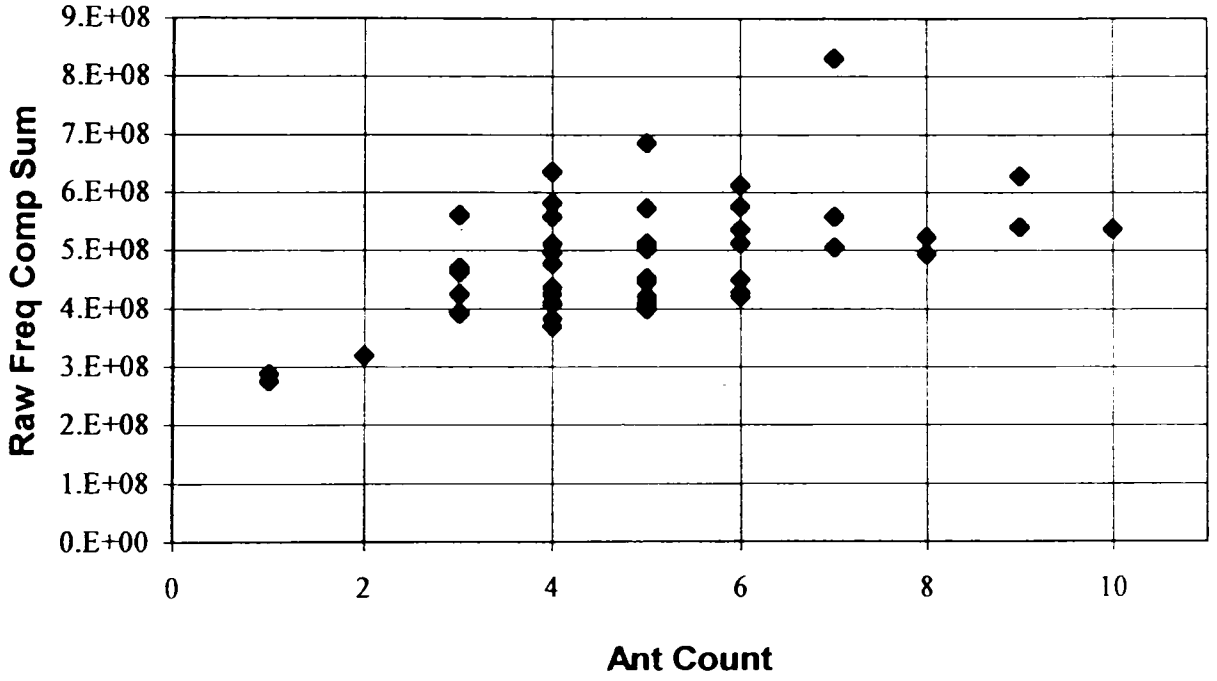



Figure 5.12: Total sum of amplitudes versus ant count.

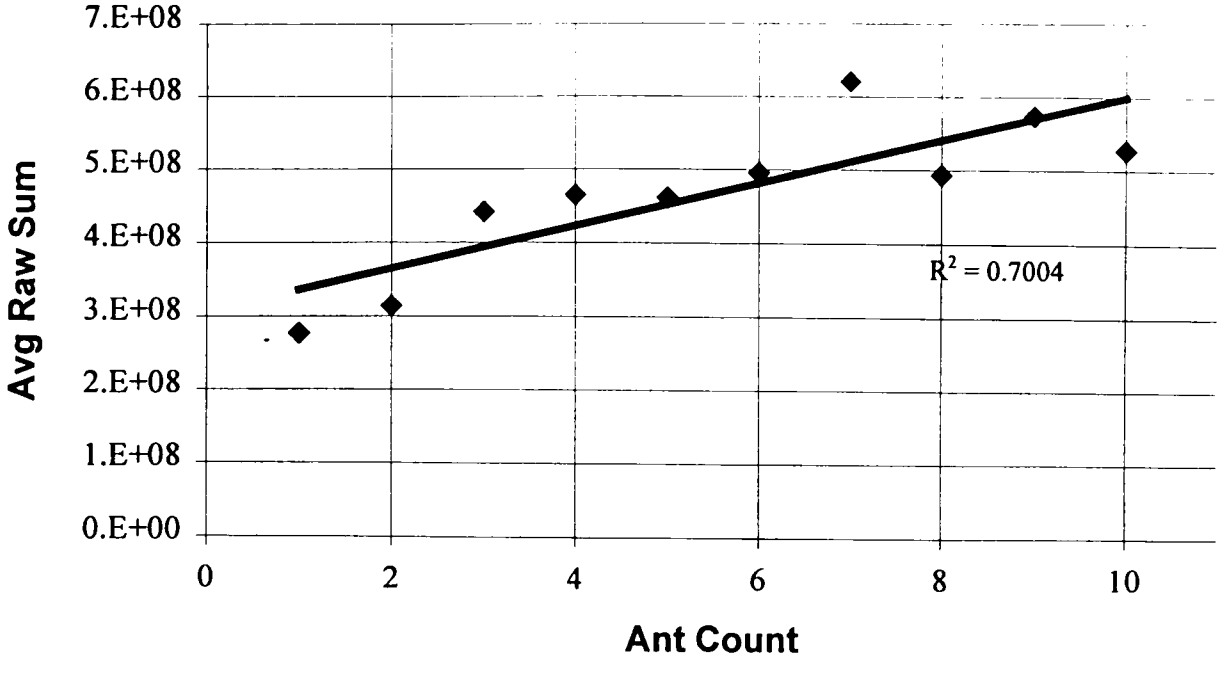



Figure 5.13: Averaged total sum of amplitudes versus ant count.

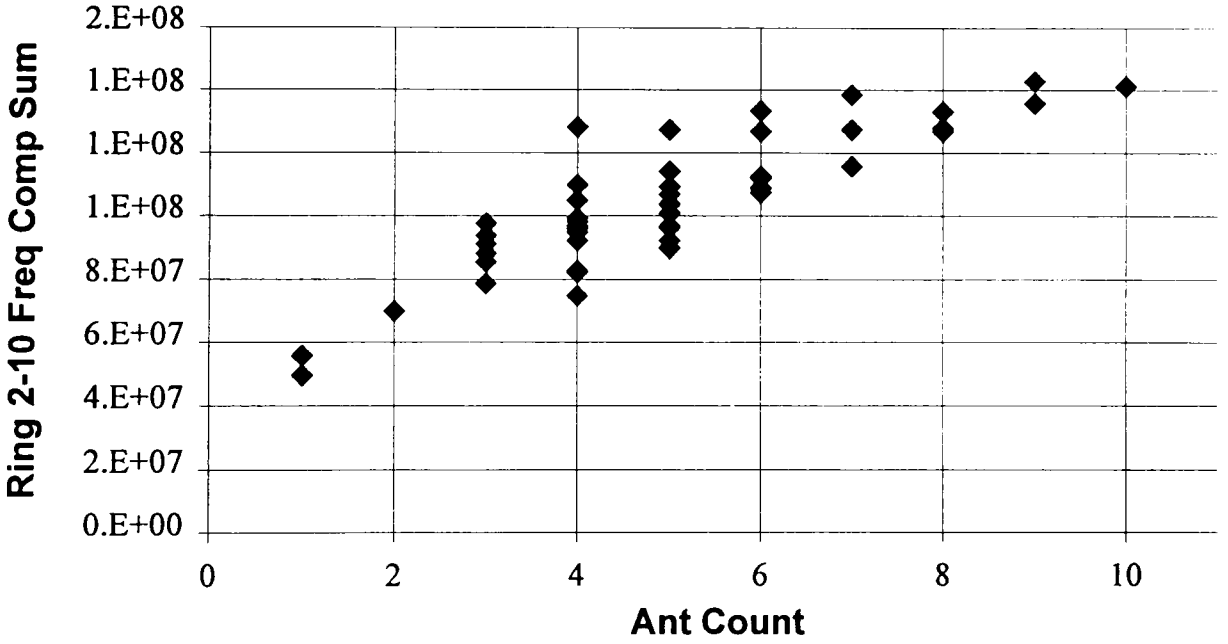



Figure 5.14: Selected (band 2-10) sum of amplitudes versus. ant count (Bandwidth = 4).

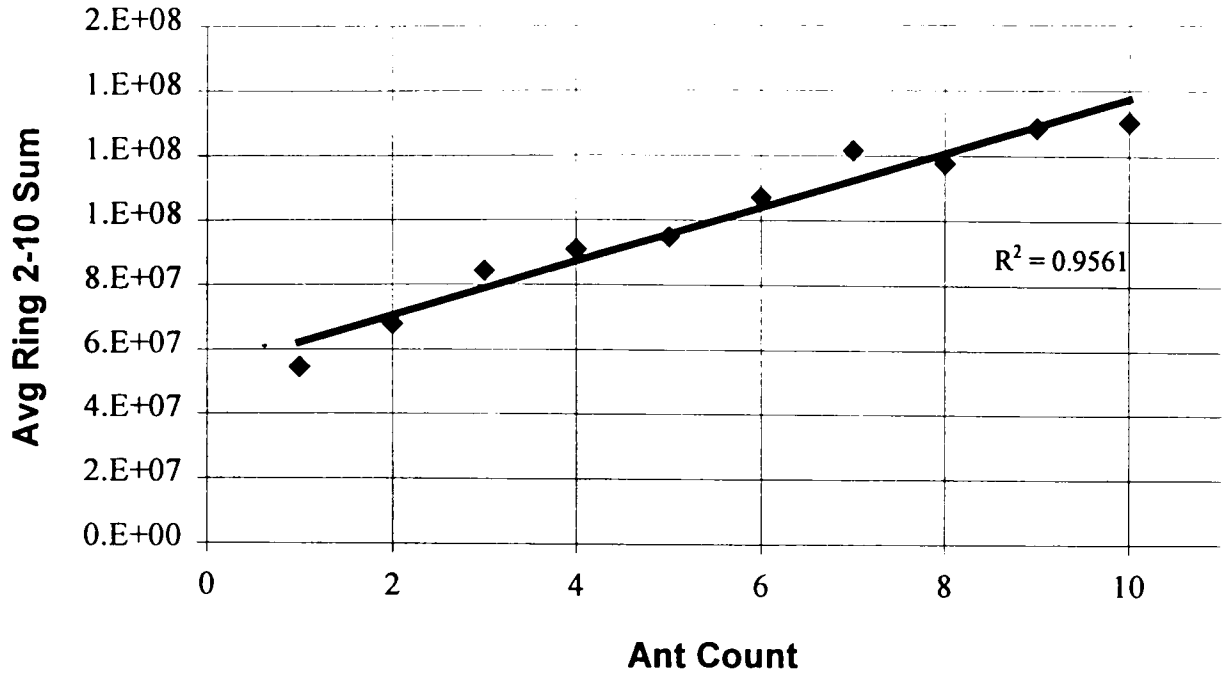



Figure 5.15: Averaged selected (band 2-10) sum of amplitudes versus ant count.

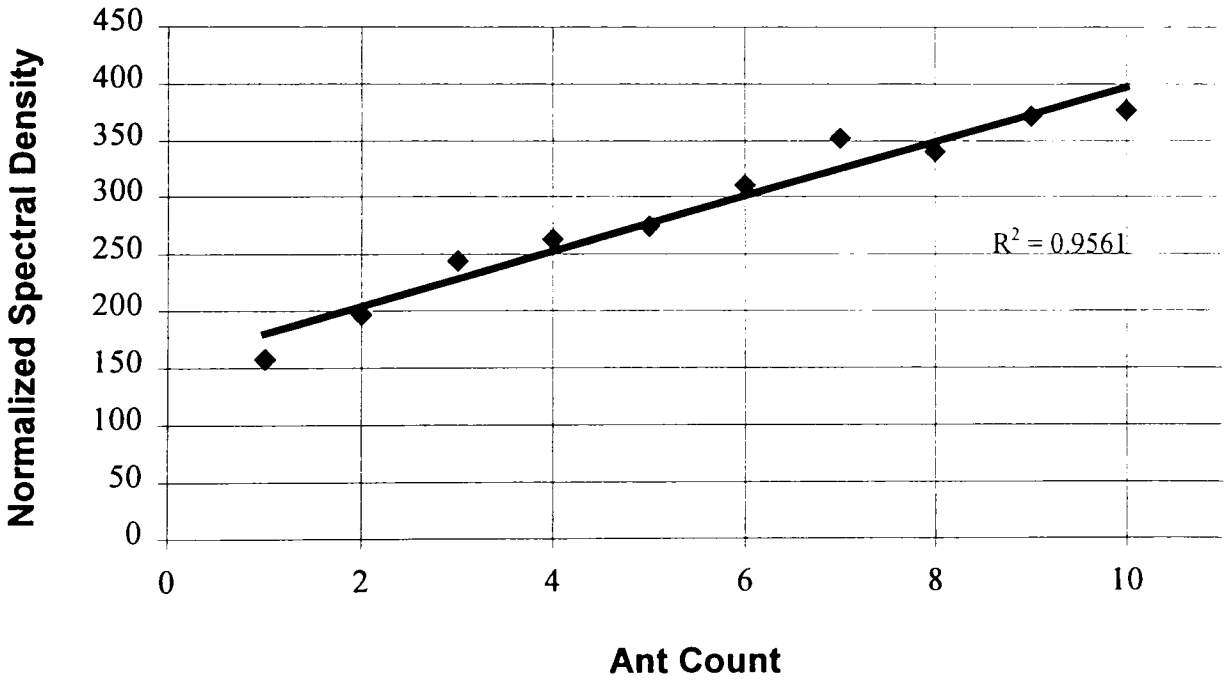



Figure 5.16: Normalized plot of Figure 5.15.

# CHAPTER VI
# STV IMAGE ANALYSIS

Researchers have been studying the visibility of small targets on highways and have developed special hardware to perform experiments in the real environment. They have also developed equations to calculate visibility. What the author presents here hopefully will be another perspective of the problem.

## 6.1 Roadway Lighting Background

The ultimate safety goal of a roadway lighting system is to reduce the number of visibility-related nighttime accidents to daytime levels. The main idea is to make the nighttime roadway environment more closely resemble its daytime counterpart that normally has best visibility.

Engineers have partially simulated daylight lighting by designing lighting systems that produce constant horizontal illuminance at the road surface. When a single light source, such as the sun, is used to light a roadway, not only the horizontal illuminance remains stable but also the vertical illuminance and the pavement luminance. However, in roadway lighting, where multiple light sources are arrayed along the road, the best that can be hoped for is merely a near-constant horizontal illumination; both pavement luminance and vertical illuminance will vary along the lighted section of roadway.

It was once decided that most objects on the road were very dark (low reflectance). Engineers knew that for these objects which are best seen in negative contrast against the pavement, high-level and uniform pavement luminance was the best solution.

For objects that are not dark, visibility relies strongly on contrast and that depends on the pavement luminance. Objects that are darker than the pavement will still be seen in negative contrast, objects lighter than the pavement will be seen

in positive contrast. Nevertheless, objects of luminance similar to the pavement will be invisible.

## 6.2 Small Target Visibility Research Background

Researchers in STV research have developed a special gadget (Figure 6.1) to be placed in the middle of a road segment illuminated by certain lighting systems. It has a square, flat surface perpendicular to both the road surface and the observer's line of sight. The square comes in different sizes. The target reflects light in a Lambertian (ideal diffuse) manner with a reflectance of 20%. Two poles that come out from the top corners hold two LED's with an on/off switch. In some extreme cases when the contrast is quite low, researchers need these tiny red lights to correctly aim their camera or other light measuring devices.

Figure 6.2 is a plot of the locations of the small targets. Figure 6.3 is an image of the same target (shown in Figure 6.1) at one of the locations under normal illumination.

Following is a method STV researchers have been using.

$$VL = \frac{L_T - L_B}{DL_4} \tag{6-1}$$

Visibility Level (VL) is a ratio and unitless. Target luminance ($L_t$) is calculated for one point at the center of the target. Background luminance ($L_b$) is the average of the pavement luminance as viewed by the observer at a point adjacent to the center of the top of the target and at a point adjacent to the center of the bottom of the target. $DL_4$ is a factor that is related to a number of factors including observer's age, etc.

As an alternative to measuring the luminance in the field, researchers have been using video cameras to take digital pictures of the target at different situation and measuring the pixel values around targets. The result has been known as Video Target Contrast or VTC.

$$VTC = \frac{L_T - L_B}{L_B} \tag{6-2}$$

$$nVTC = \frac{\Sigma L_T - \Sigma L_B}{\Sigma L_B} \tag{6-3}$$

## 6.3 Lights Off and Lights On Images

Digital images acquired under dramatically different lighting conditions prove to be great tools to learn the relation between image content and the corresponding frequency component.

### 6.3.1 Image Acquisition

In a typical night time roadway situation we would like to compare the information available to an observer under no light conditions, off road lighting conditions, partial roadway lighting conditions, and under full roadway lighting conditions.

All the images were made using an analog CCD video camera and a digital CCD video camera. The images were transferred to computer disk and Fourier transforms of the images were calculated. The analog and digital video images were compared to determine the similarities and differences in the digital and analog images then Fourier transforms of both types of images were calculated and compared. Either image represents the scene very well so we used the most convenient image capture and transfer technology available to us. The image capture and transfer technology was in the mid price range state-of-the-art electronics. More expensive technology would not improve the analysis significantly.

### 6.3.2 Understanding Frequency Domain

#### 6.3.2.1 No Light Condition

The series starts with the image under no light conditions. As shown in Figure 6.37, there is no light for the CCD camera to record, the scene is black. However, the Fourier transform of the CCD image shows some high frequency components and the amplitude of the components increase with increasing frequency. That shows that the elements of the digital image matrix are not simply all zeros. Otherwise, we would have nothing in the spatial frequency domain. Upon a very close inspection of the no light CCD image matrix, it was discovered that the gray scale pixels were indeed different from "dead black" or 0.

We know that the CCD camera has 256 levels of gray available and all the cells in the CCD array are not registering 0 for no light. The fact that the CCD has a dark current always present in each cell of the CCD array determines that some of the cells will be read at a level above 0. The values above 0 are distributed randomly throughout the CCD array.

There are two choices to fix the dark current problem. We can cool the CCD array to -40°F, which is very expensive. Or, we can determine what the Fourier transform of the dark current looks like and filter the dark current out of the Fourier transforms. We characterized the dark current frequency response, as shown above, to determine the dark current's contribution to the Fourier transform. The dark current's contribution was very small compared to the Fourier transform of images under other conditions. Therefore, we choose not to filter it out, but filtering it out is an easy task if it becomes necessary.

#### 6.3.2.2 Off road lighting conditions

After the zero light condition had been established the off road lighting conditions were imaged. The site we used for our roadway lighting conditions was adjacent to a lighted rest stop on U. S. Interstate 27, north of Abernathy, Texas. The off road lighting was too low for the CCD video camera to clearly image the

site, however the oncoming automobile head lights, and other light sources in the field of view of the camera were ready for imaging.

In Figure 6.38, the spatial frequency content in the high spatial frequency region has increased slightly but the low frequencies have remained small. There is an increase in the amplitudes of the frequencies from range 4 to 14 due to the auto and other point source lights. The point sources produce enough light for the camera to image themselves but not enough energy to illuminate any portion of the site. The contribution of the off road lighting from the rest stop lighting is visible in the original image and calculated in the Fourier transform, but the details are not visible in Figure 6.38 due to the resolution of the reproduction in the paper. The rest stop off road lighting was just at the threshold of the CCD's low light imaging limit. Under such circumstance, the human observers were in their lower mesoptic vision region. The average background luminance and the low frequencies from range 1 to 3 are very small because there is almost no contrast in the image.

We also imaged the off road lighting conditions with an STV target in the field of view. It was a standard 20% reflective, 18-cm x 18-cm target. Two fixed LED diodes were mounted on the target to make it locatable in the CCD's field of view and for future luminance calibration techniques. In Figure 6.39, the LED's are clearly visible in the low middle center of the image and the auto headlights have moved. The single unmoving source in all the frames is an outdoor light at a farm house The addition of the LED's and the changes in both locations and intensities of the auto head lights have changed the Fourier transform slightly. The more intense auto headlights in the middle right increase the amplitude of the midrange frequencies, but the higher frequency components still remain very low.

#### 6.3.2.3 Lighting Conditions

Under partial roadway lighting conditions, the roadway lights behind the STV target were energized to determine the frequency response to this contribution of partial illumination of the site. As can be seen in Figure 6.40, there

is an image of the top portion of the STV target and a portion of the roadway surface and its surrounding area are illuminated.

From the Fourier analysis of the scene the lower range frequencies, from range 1 to 3, have increased because the average background luminance has increased, but it has not increased a great deal. The mid frequencies, from range 4 to 14, have also changed slightly reflecting the more detail available in the image. The high frequency content has increased slightly.

Under full roadway lighting conditions shown in Figure 6.41, all the details of the roadway are clearly visible, the STV target is clearly visible along with the auto headlights. From the Fourier analysis, it can be seen that the lower range frequencies, from range 1 to 3, have increased significantly reflecting the large increase in the background luminance of the image. The midrange frequencies, from range 4 to 14, have increase significantly reflecting the availability of much more detail of the scene in the video image. The high frequencies, from range 15 to about 35 have also increased significantly showing the increase in available detail in the higher frequency ranges.

### 6.3.3 Conclusion

The graphs of in the previous figures were all scaled but they are still hard to compare. When all the graphs are placed in the same scale as shown in Figure 6.42, we can see the changes in spatial frequency content from a lights off to a lights on situation. The changes represent the differences in the information available to an observer as more light is added to the scene. As more light is added to the scene the average background luminance of the scene increases as shown in the region, from range 1 to 4. From range 4 to 38, as more light is added to the scene, the amplitudes of the higher frequencies increase, reflecting the added detail available to an observer.

The curve associated with Figure 6.41, has the largest partial volumes above each partial area. The total volume, the sum of all the partial volumes, is a

measure of the total information available to the observer. An observer has a visual transfer function with a cutoff frequency of about one minute-of-arc (1') or less, so an observer cannot resolve the higher frequencies past about, range 12 to 13 (according to the calculation in earlier section). As the background average luminance increases from a few cd/m$^2$ to thousands of cd/m$^2$ , the range 1 to 4, an observers visual response changes from sotopic to mesoptic to photopic. So the absolute value of the background luminance from the Fourier transform determines the visual lighting level response and the higher frequencies determine the details available to the observer.

Now it should be noted that in order to change the scene noticeably to an observer the modulation transfer function (MTF) must change by at least 10% or the changes in the scene are not noticeable to an observer. We also believe if MTF is kept unchanged, then the information provided by the scene must change by at least 10% for a clear recognition. From Figure 6.42, it is clear that there is a greater than 10% change from no light to some lights and there is a greater than 10% change from some light to full roadway lighting. The fully lighted roadway has almost a maximum background luminance. In order to increase the MTF by 10% the background luminance and all the higher frequency components must be increased or decreased by 10%. Since an observer's eyes operate over eight orders of magnitude in background luminance differences, from night vision to day time vision, 10% may not be discernible in mesoptic vision. However, we are satisfied when the changes in the spatial frequency are discernable to the observer.

## 6.4 Ten Image Sections with Targets

### 6.4.1 The Images

Ten images are carefully cut down from images taken of targets at different locations on the near side of the road lamps. Each of them is of size 64 by 64

pixels, with the target located at the center and some pavement as the background.

First, we judge the visibility of these targets by looking at a pair of these cut images and subjectively judge in which the target can be better seen. This method is known as 'subjective bubble sorting'. The result are shown in Table 6.1

Steps of 'subjective bubble sorting': (Distance between eyeballs and computer screen is about 2 feet.) We should be aware that human judgement error exists in such methods.

### 6.4.2 Sampling Theorem

A *bandlimited signal* is a signal, *f(t)*, which has no spectral components beyond a frequency *B* Hz; that is, *F(s) = 0* for $|s| > B$. The *sampling theorem* states that a real signal, *f(t)*, which is bandlimited to *B* Hz can be reconstructed without error from samples taken uniformly at a rate $R > 2B$ samples per second. This minimum sampling frequency, $R_N = 2B$ Hz, is called the *Nyquist rate* or the *Nyquist frequency*. The corresponding sampling interval, $T = 1/2B$ (where $t = nT$), is called the *Nyquist interval*. A signal bandlimited to *B* Hz which is sampled at less than the Nyquist frequency of *2B*, *i.e.,* which was sampled at an interval $T > 1/2B$, is said to be *undersampled*.

### 6.4.3 Human Eye Resolution and Cut-off Frequency

An interesting question emerged from examining the above images. What does the driver see instead of the video camera? To answer this question, first we need to find out what resolution a human eye has. It is related to the cutoff frequency. In other words, we will be able to decide what frequency range is effective in the frequency domain representation of a digital image. If that frequency is lower than the maximum frequency, we can remove all of the higher frequency components. Furthermore, we can use inverse FFT to reconstruct a digital image that shows what a human eye should see.

The resolutions of human eyes vary. However, it is said that roughly most people have a bar-gap resolution of a minute of arc. We will try to determine the upper cutoff frequency of human eye based on this resolution value and the measurement among our images.

Consider a line segment of 18 centimeters (because this is the length of a side of the standard square target) located at 83 meters away from the observer. We can estimate the span of the line segment from the observer's (eyeball or camera) point of view as:

$$\frac{18cm}{8300cm} \bullet \frac{180^\circ}{\pi} \bullet \frac{60'}{1^\circ} = 7.46 \text{ minute of arc.} \qquad (6\text{-}4)$$

Then we measure the target in the image to be 22 pixels. This gives us:

$$\frac{22pel}{7.46'} = 2.9 \text{ pixel per minute of arc.} \qquad (6\text{-}5)$$

Obviously, we now are able to find in our particular example the one minute of arc human bar-gap resolution is equivalent to about 3 pixels. Meanwhile, the resolution of the digital video camera is 1 pixel by definition. The number of pixels per min of arc is related to digital video camera's magnification or zooming. In a video cam that has a "zooming" capability, the images should be taken at the same zooming settings.

We can claim that under the settings we picked for these images, an average human eye has a sample rate as low as approximately one-third of the camera.

$$R_{human} = \frac{1}{3} R_{camera} \qquad (6\text{-}6)$$

This shows that the cutoff frequency of eye will be one third of that of the camera according to the Nyquist frequency.

$$B_{human} = \frac{1}{3} B_{camera} \qquad (6\text{-}7)$$

With this knowledge in hand, we can examine the Fourier transform equations and estimate the corresponding frequency limit. We know that the frequency domain representation of a 30 by 30 pixel image is a 30 by 30 matrix

with constant term located at center. That gives $B_{camera}$=[-15, 15]. From Equation (6-7), we have $B_{human}$ =[-5, 5].

In an FFT of a digital image with 30 by 30 pixels, the effective frequency component representing a typical human eye gain under the same situation will be from the center to the 5$^{th}$ ring introduced in section 4.2.3 and illustrated in Figure 4.12. If the image size is 480 by 480, the frequency component within 80 rings (ring introduced in section 4.2.3 and illustrated in Figure 4.13) from the center should be considered what human eyes get.

Following are the result of the reconstructed images with higher frequency component outside of the 5$^{th}$ ring removed.

The following points can be seen when Figures 6.19 through 6.28 are compared with Figures 6.5 through 6.14 :

- The resolution of the video cam is greater than that of the human eye.
- The images may be filtered to determine the cutoff spatial frequency intrinsic of individual human eye. Suppose the images' frequency components are filtered one ring at a time from the highest ring. When the observer first detects a change of content judging only by his eyes, we say he has a cutoff frequency related to that ring number.
- For most people, filtering the component above the 5$^{th}$ ring does not significantly reduce the resolution for their eyes.

### 6.4.4 Frequency Component Analysis

For 20 of the cut target images, we perform the calculations introduced in Chapter 5. We pick square division due to the limited pixels within the digital image. A plot of the FFT calculation result is shown in Figure 6.29. It also has a size of 30 by 30 pixels. The square at location (16, 16) represents the DC value. The bright squares at adjacent locations represent the frequencies that have relatively larger component. We also find that these high value point are mainly located around the center or on the axes of the $f_x$=0 and $f_y$=0. Because of these

missing frequency component at off-axis frequency space, our choice of the square division method introduced in Section 5.2.3 becomes a very good approximation of circular division.

First, let us look at a 3-D plot of the frequency component distribution of the series of target images (30 by 30 pixel) of the ten left locations (Figure 6.30). There are two general features. First, the later locations have higher components than the early ones. Second, the frequency component is mainly distributed within the first 5 ring. For higher frequencies, it is hard to say whether the value is more related to signal or more to noise. Figure 6.31 is the same data presented in a two-dimensional graph.

The second observation provides us another justification for us to neglect the higher frequency components.

The results of the frequency component analysis of both left side and right side are shown in Figure 7.13. What we can see from this plot are:

- The frequency component analysis values at further locations are usually higher than those at nearer locations.
- The frequency component analysis values of the left side of the road are usually higher than those of the right side.
- The highest frequency component analysis value of either side does not occur at the furthest location.
- The lowest frequency component analysis value of either side occurs at the nearest location.

Now we can compare with VTC(Video Target Contrast) and n-point VTC results shown in the following two plots.

To compare these data better, we normalize all series of results to make them fit into interval [0, 1] so that they can be shown in a single graph. From both the left side comparison and the right side comparison, we can see the dynamic relations of the results given by different methods agree generally. They almost have the same slope. The differences occur at the beginning and end. The shape

of the curve of our frequency component analysis is closer to the n-point STV curve.

### 6.4.5 Conclusion

This section demonstrates the following feature of the Frequency Component Analysis:

Firstly, theoretically the frequency component analysis does provide another perspective to look at the STV problem.

Secondly, the example images we studied show that the relative visibility levels of the 20 target images revealed by the frequency component analysis match the STV results. In other word, we can roughly reproduce the same behavior of the different target locations.

Thirdly, we believe that dynamically FCA and VTC are equivalent in revealing the visibility in the specified circumstances.

Finally, in order to be comparable, standardized image acquisition method should be carefully designed and carried out. We are interested in maintaining fixed focal length, fixed tilt of camera, constant lighting, etc.

Table 6.1: Subjective Bubble Sorting

| Step number | Finding | Ranking Adjustment Starting with low visibility |
|---|---|---|
| 0 | None | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 |
| 1 | 2<1 | ***2, 1,*** 3, 4, 5, 6, 7, 8, 9, 10 |
| 2-7 | 3>1, 4>3, 5>4, 6>5, 7>6, 8>7 | ***2, 1, 3, 4, 5, 6, 7, 8,*** 9, 10 |
| 8-11 | 9<8, 9<7, 9<6, 9>5 | ***2, 1, 3, 4, 5, 9, 6, 7, 8,*** 10 |
| 12-14 | 10<8, 10<7, 10>6 | ***2, 1, 3, 4, 5, 9, 6, 10, 7, 8*** |

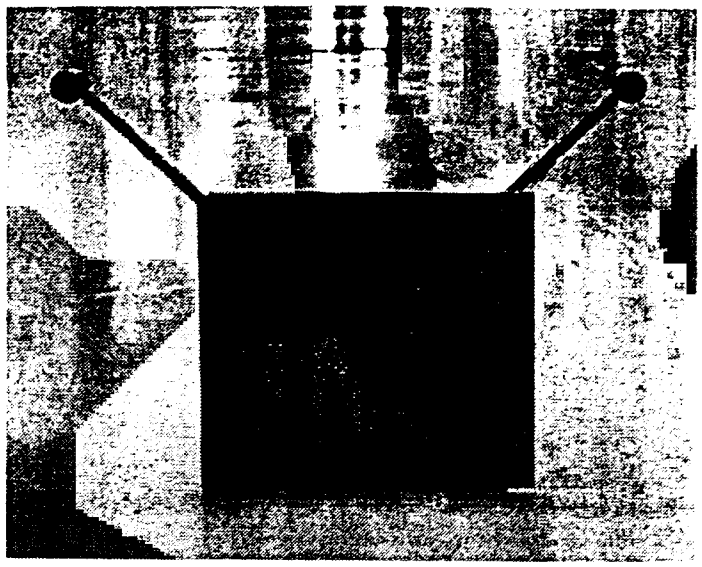


Figure 6.1: The 18cm by 18cm target.

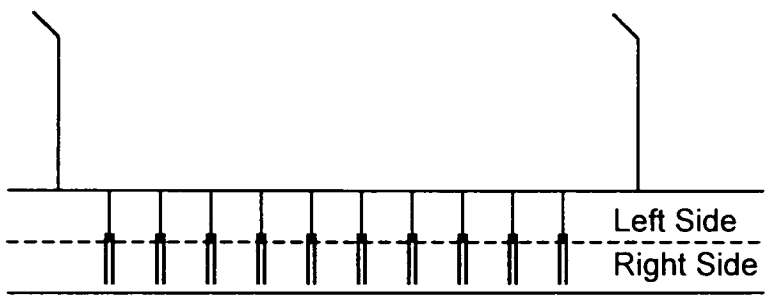



Figure 6.2: Illustration of the lighting system and the target setup.

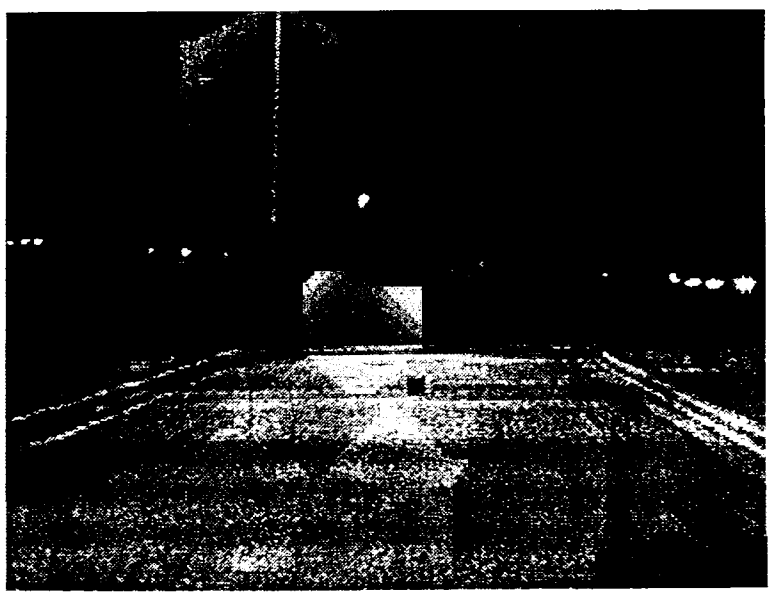


Figure 6.3: An example of VTC setup.

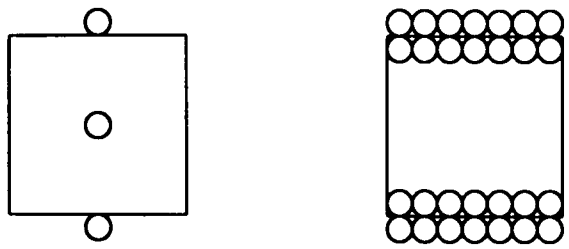


Figure 6.4: VTC and n-point VTC.

Figure 6.5: Original left side location # 1.

Figure 6.6: Original left side location # 2.

Figure 6.7: Original left side location # 3.

Figure 6.8: Original left side location # 4.

Figure 6.9: Original left side location # 5.

Figure 6.10: Original left side location # 6.

Figure 6.11: Original left side location # 7.

Figure 6.12: Original left side location # 8.

Figure 6.13: Original left side location # 9.

Figure 6.14: Original left side location # 10.

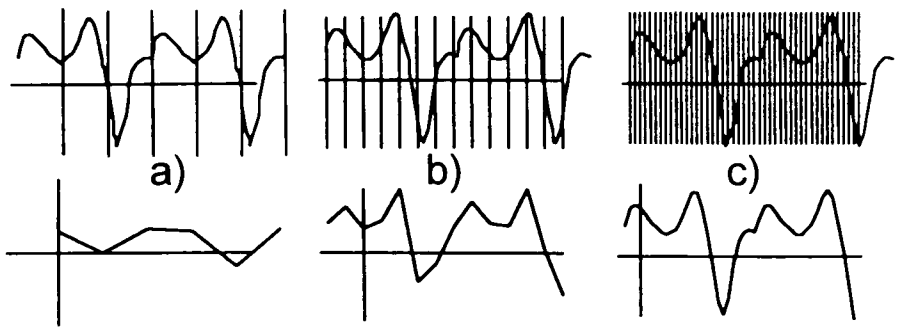



Figure 6.15: f(x) sampled.
(a) below the Nyquist, (b) below the Nyquist,
(c) at or above the Nyquist rate and the
reconstructed $f_R(x)$'s from the Nyquist samples.

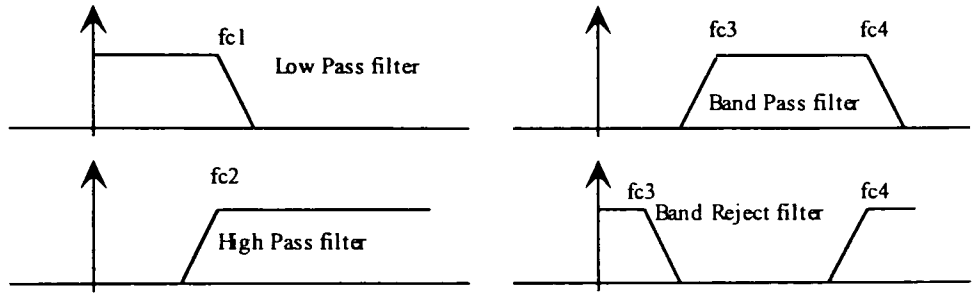



Figure 6.16: Different kind of filters.

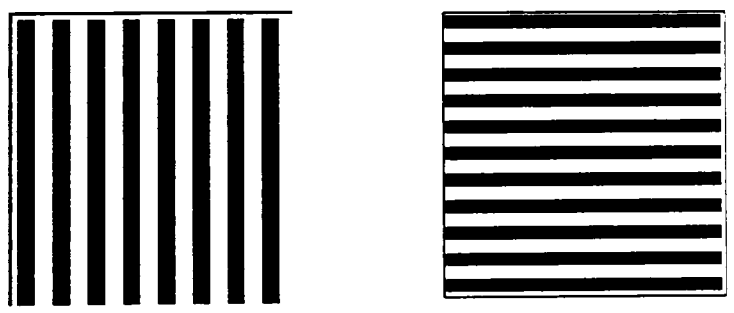


Figure 6.17: Vertical and Horizontal bars.

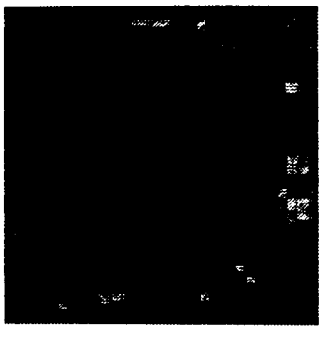


Figure 6.18: Magnified original left side location # 10.

Figure 6.19: Filtered left side location # 1.

Figure 6.20: Filtered left side location # 2.

Figure 6.21: Filtered left side location # 3.

Figure 6.22: Filtered left side location # 4.

Figure 6.23: Filtered left side location # 5.

Figure 6.24: Filtered left side location # 6.

Figure 6.25: Filtered left side location # 7.

Figure 6.26: Filtered left side location # 8.

Figure 6.27: Filtered left side location # 9.

Figure 6.28: Filtered left side location # 10.

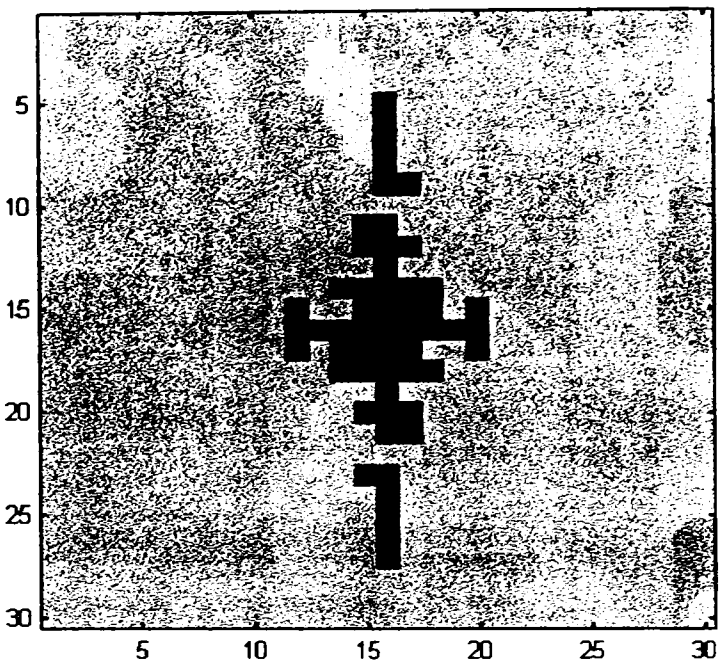



Figure 6.29: Frequency domain left side location #7.

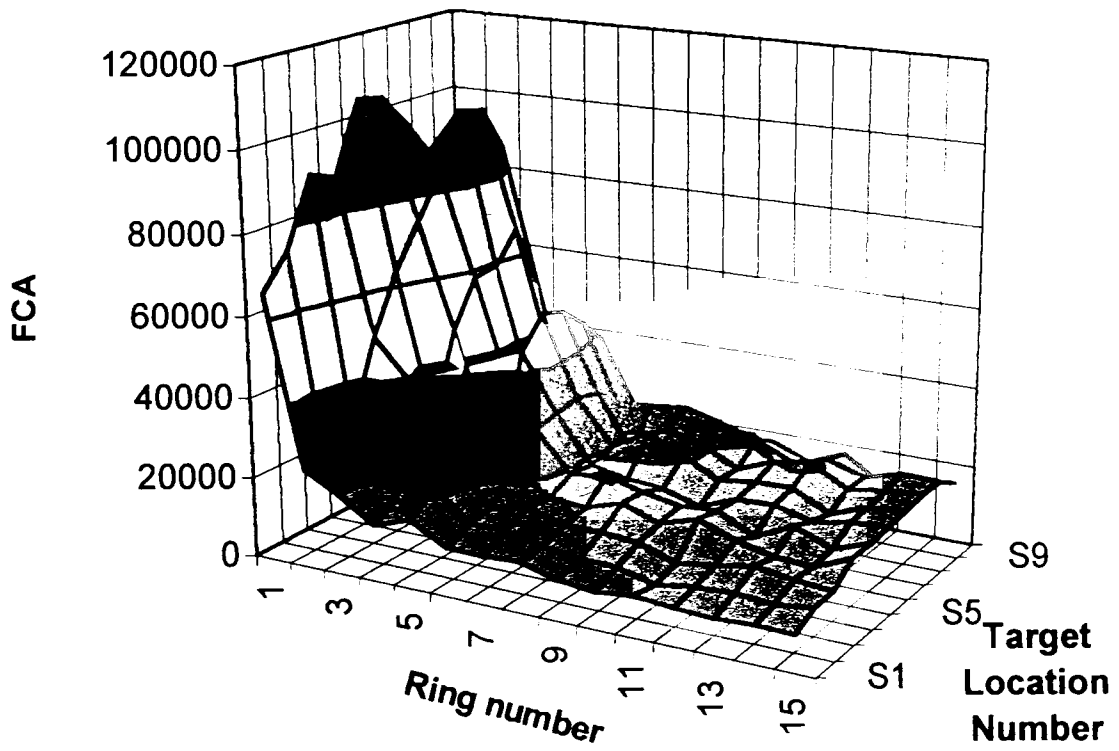



Figure 6.30: 3-D frequency component distribution
( of the 10 targets on the left side ).

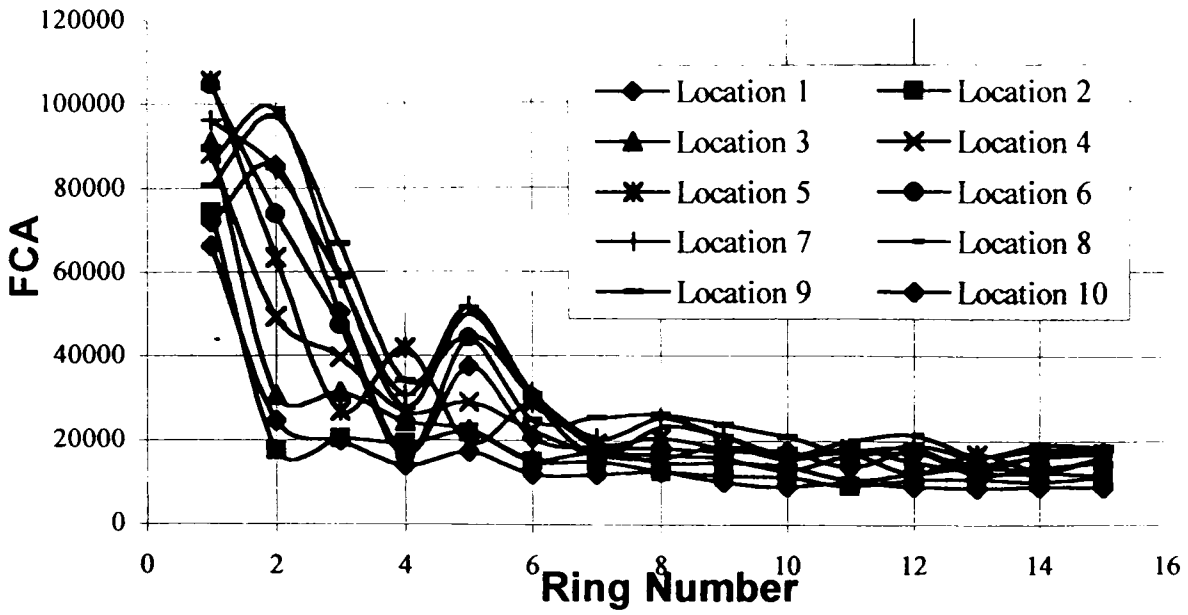



Figure 6.31: 2-D frequency component distribution
(of the 10 targets on the left side ).

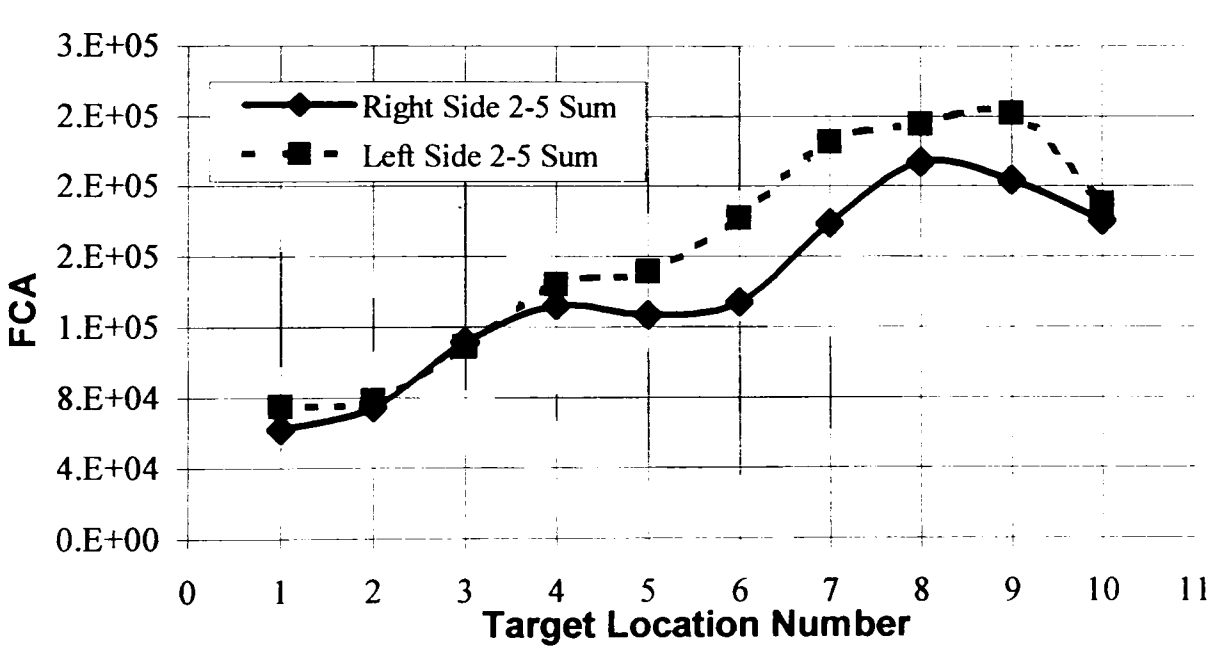



Figure 6.32: Frequency Component Analysis –
Comparison between left and right targets.

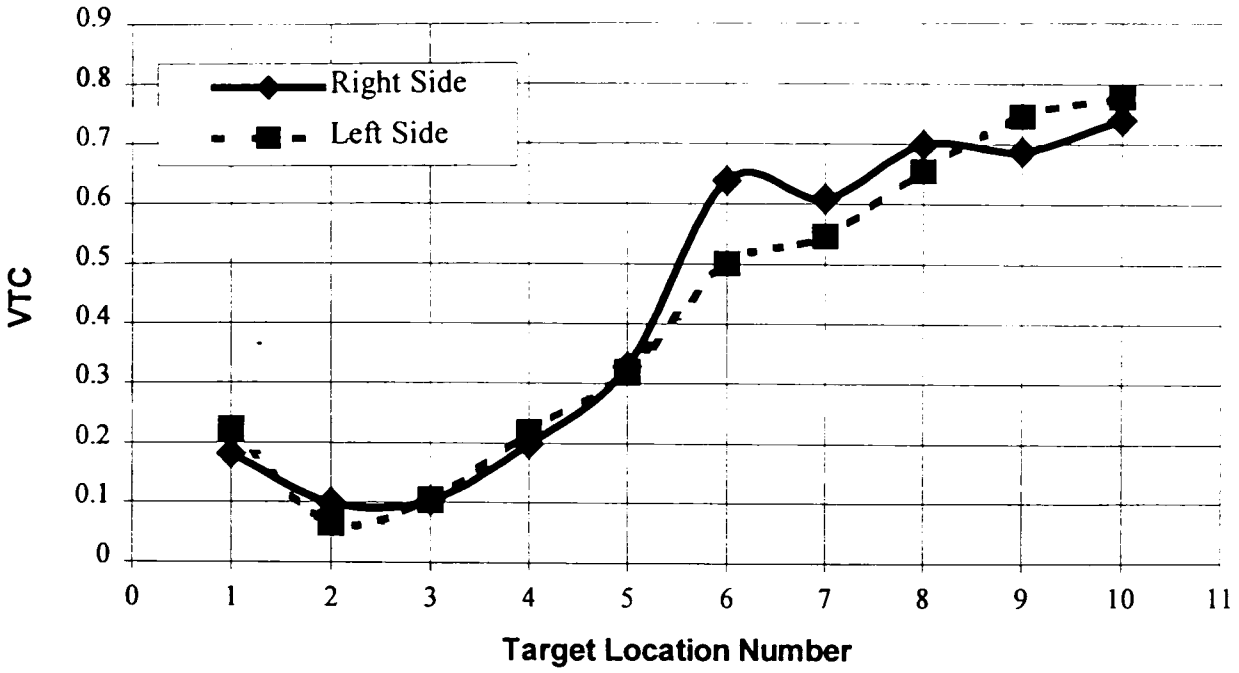



Figure 6.33: Video Target Contrast –
Comparison between left and right target.

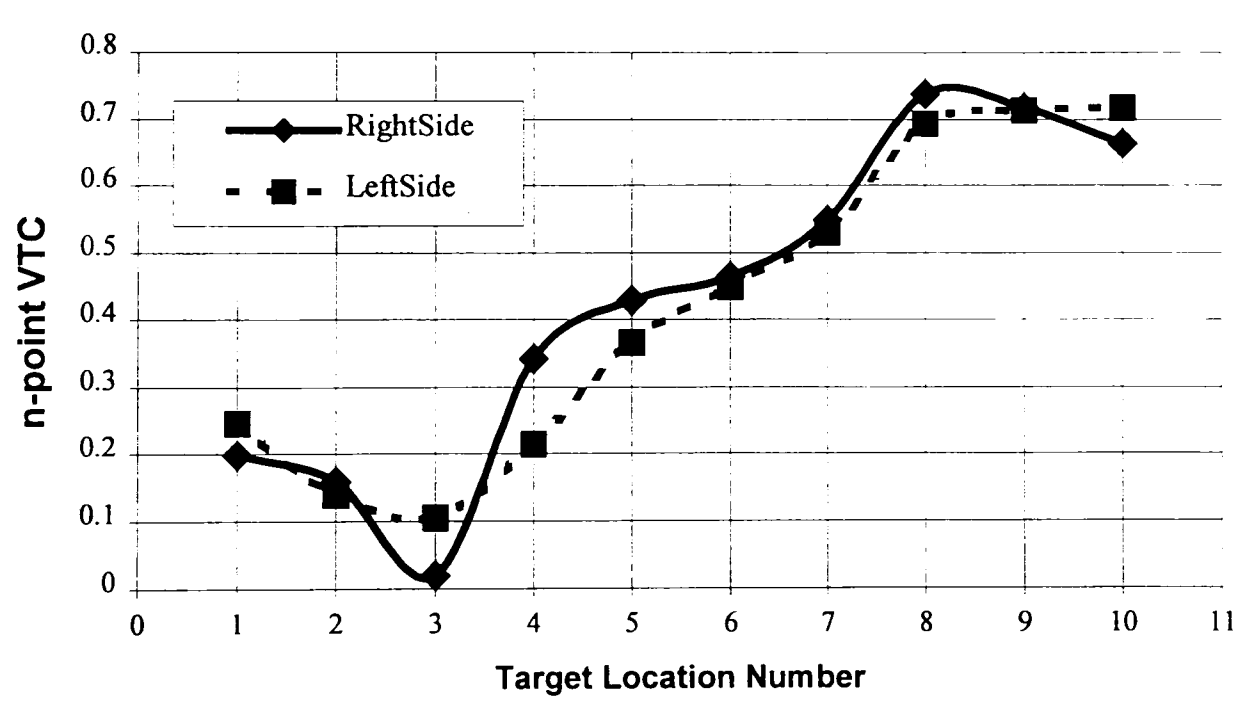



Figure 6.34: n-point VTC –
Comparison between left and right targets.

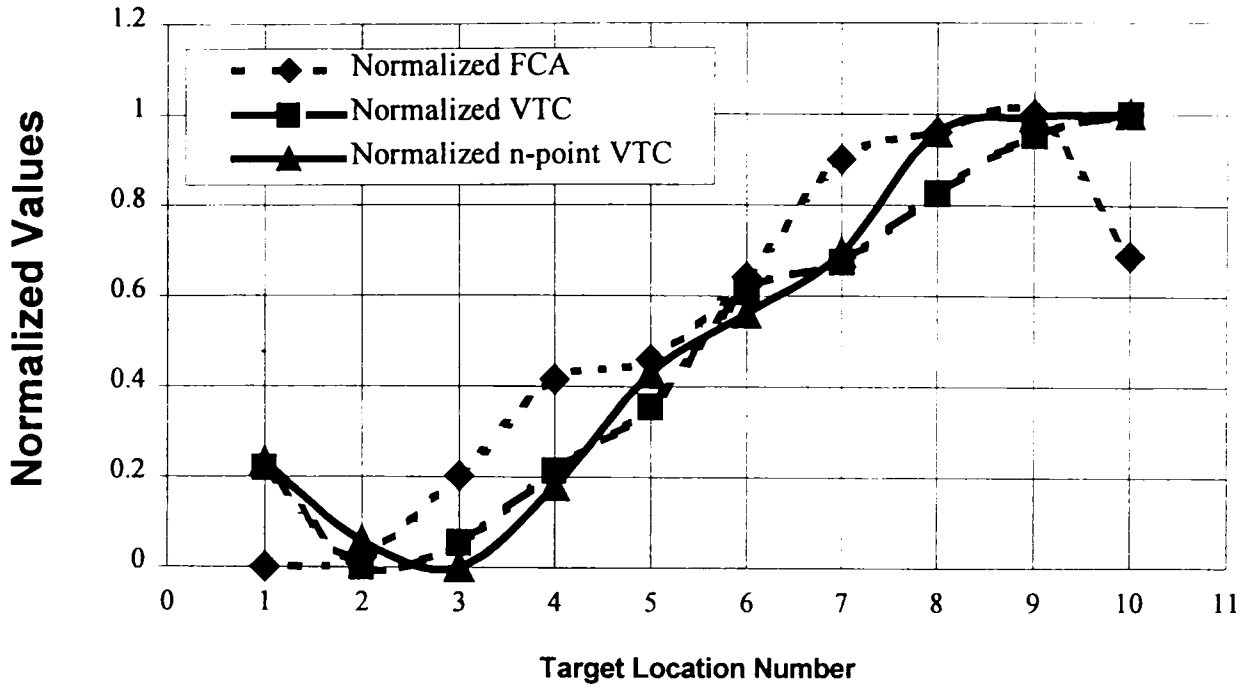



Figure 6.35: Comparison between methods (left side).

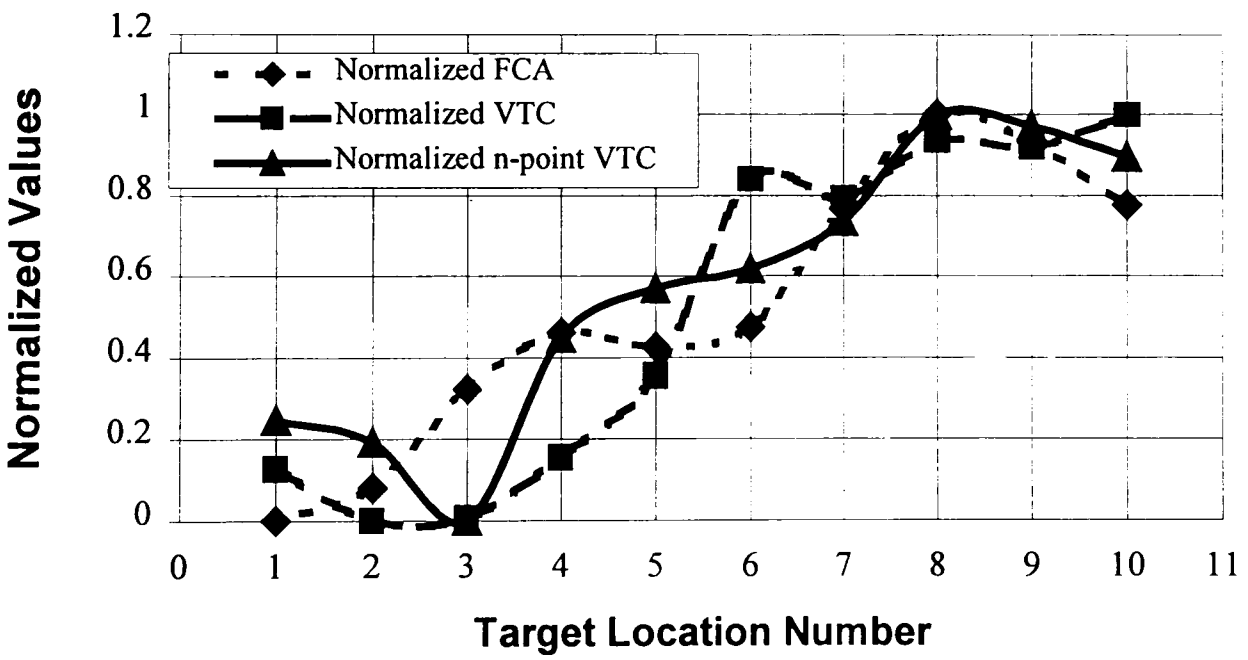



Figure 6.36: Comparison between methods (right side).

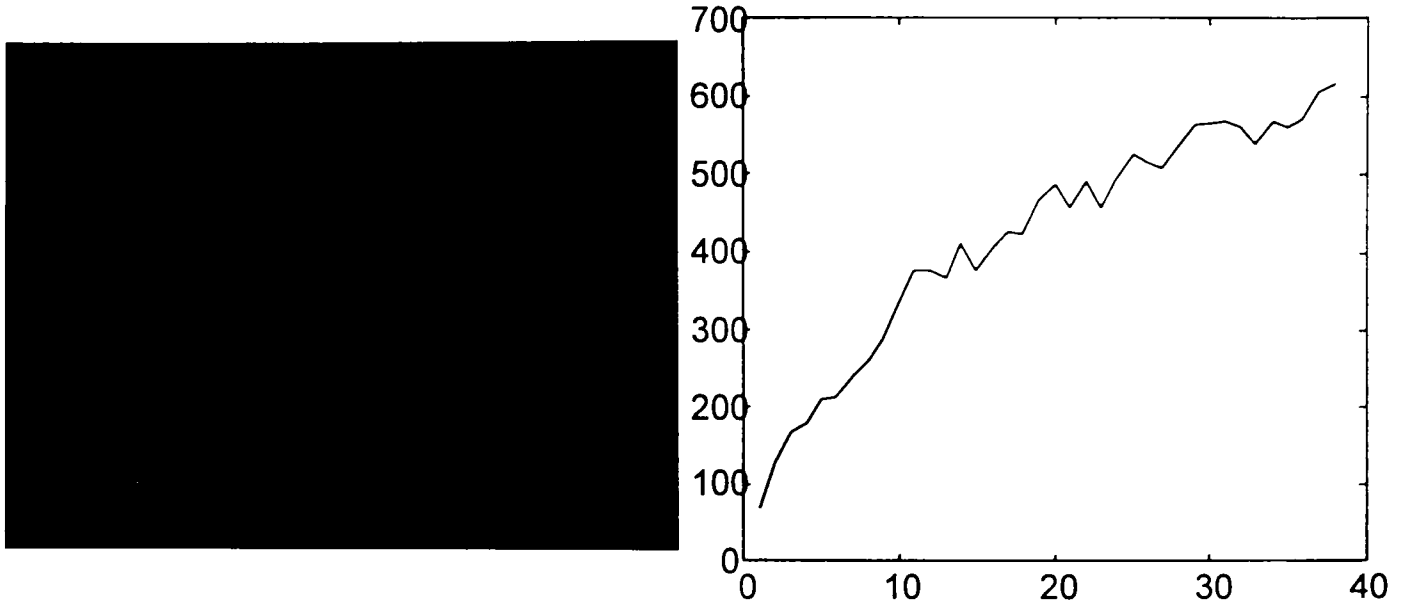



Figure 6.37: No light image
(and frequency component distribution).

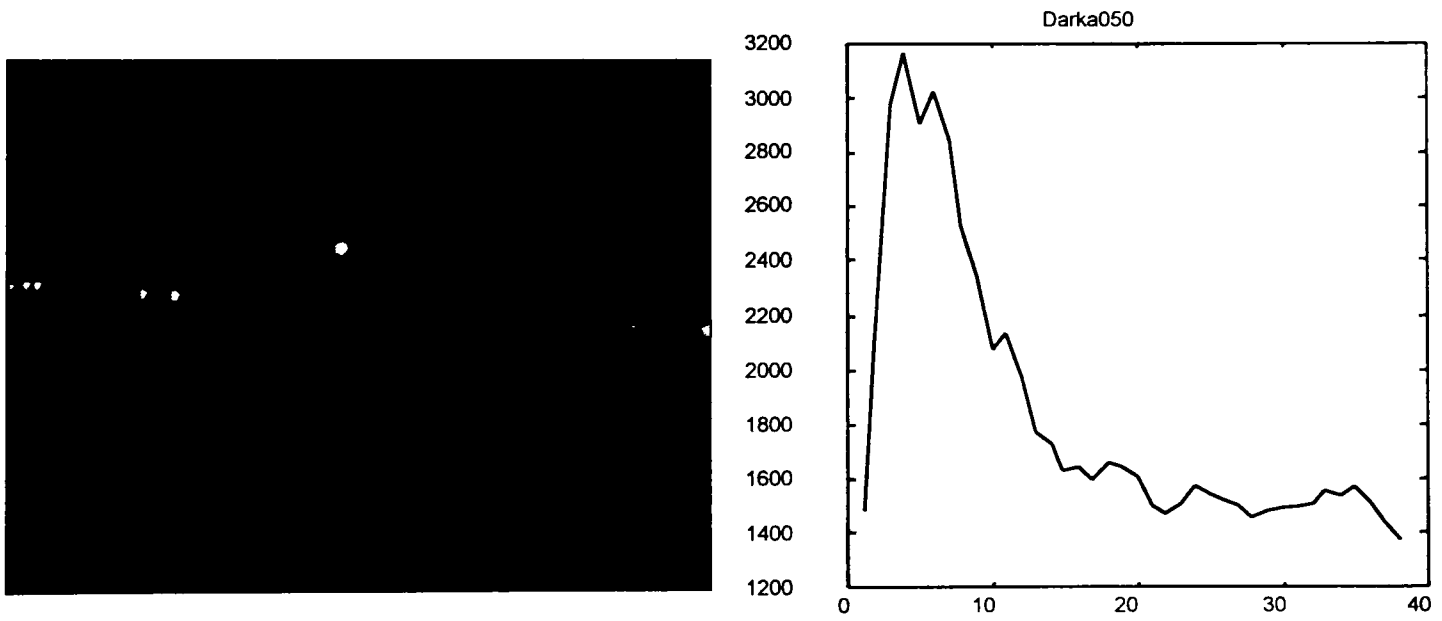



Figure 6.38: Off road lights image
(and frequency component distribution).

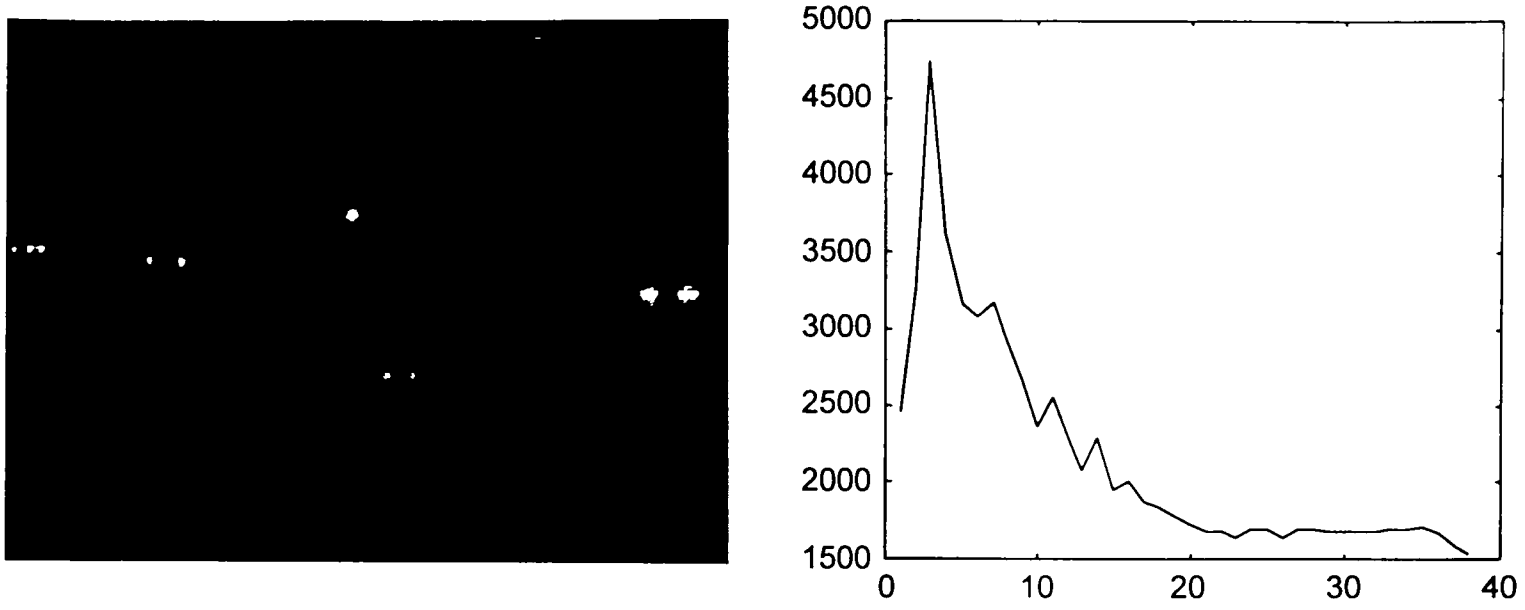



Figure 6.39: Off road lights and LEDs image
(and frequency component distribution).

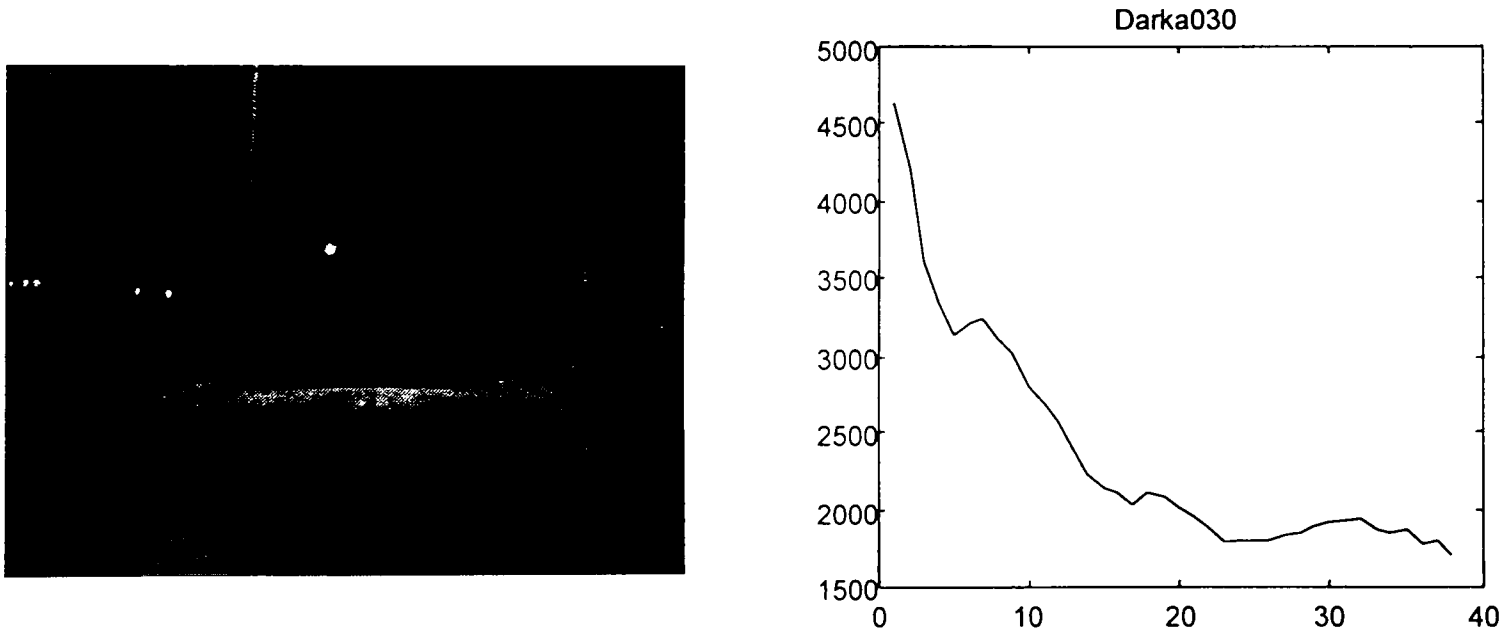



Figure 6.40: Partial roadway lighting image
(and frequency component distribution).

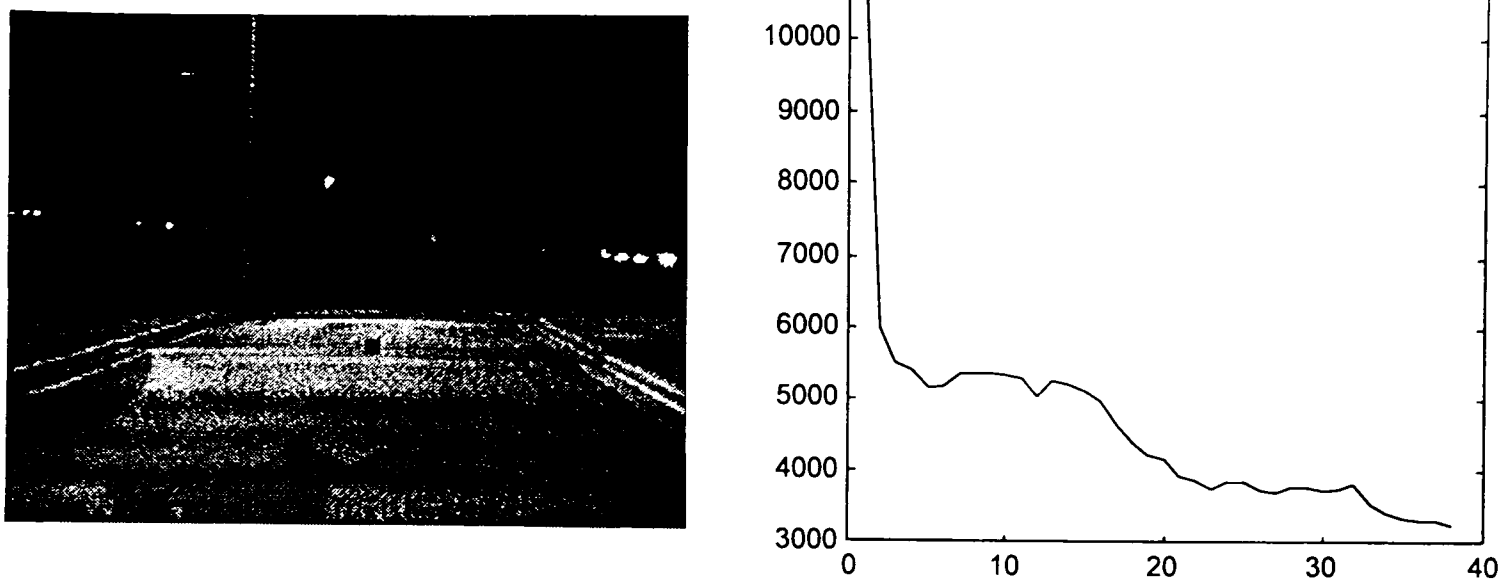



Figure 6.41: Full roadway lighting image
(and frequency component distribution).

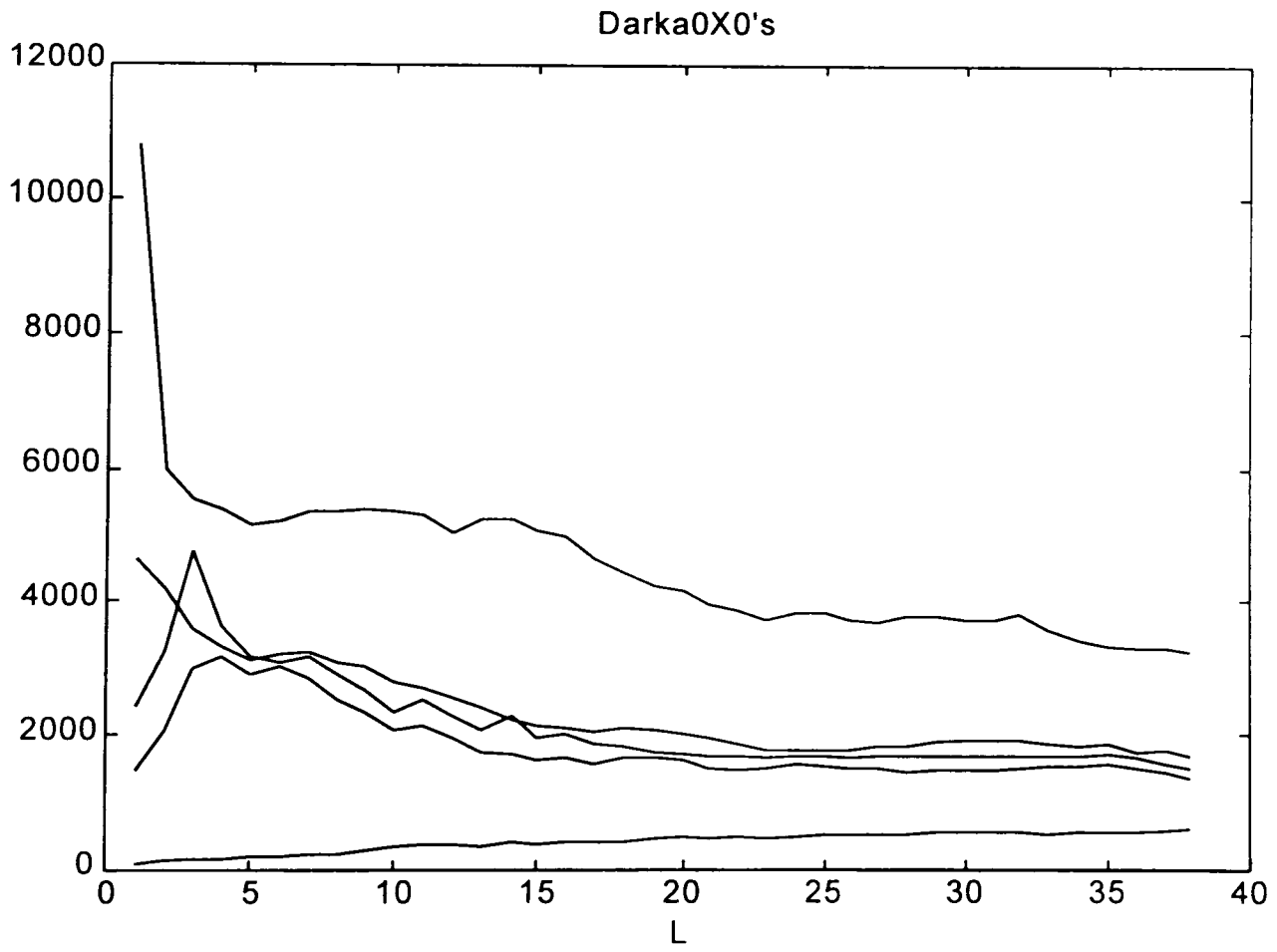



Figure 6.42: Frequency component distribution
(comparison of variable lighting situations).

# CHAPTER VII
# CONCLUSION

This thesis presents an overlook at the integrated use of digital video camera, capture card, computer and various software to acquire, store and analyze digital images. The procedures have been made very simple due to the off-the-shelf hardware and software products.

This thesis also demonstrates the feasibility of frequency component analysis on some specified digital images of different topics. It is based on the powerful mathematical tool of Fourier transform and the rapidly growing computing power of modern computers, which make its future result even promising. It has proven its universal ability of analyzing and getting useful information from the digital images of the subjects studied both in engineering and science research projects. It has also been successfully tried in other civil engineering research on roadway cracks and beam cracks.

To effectively use this approach, one needs to set up the experiment under strict standards, to make sure the digital images filmed contain the right information instead of interference from the environment, and to make sure they are comparable under different situations.

Currently, difficulties are encountered in comparing batches of images taken at different times of the day, different lighting situations, different background environments. Better formulas still need to be developed to compare series of images taken on different situations on the same subject.

# BIBLIOGRAPHY


1. Solari, Stephen J., *Digital Video and Audio Compression.* McGraw-Hill, New York, 1997.
2. Duffieux, P.M., *The Fourier Transform and its Applications to Optics*, Wiley-Interscience, New York, 1983.
3. Mansuripur, Masud, *Introduction to Information Theory.* Prentice-Hall, Englewood Cliffs, NJ, 1987.
4. Abramson, Norman, *Information Theory and Coding.* McGraw-Hill, New York, 1963.
5. Gonzalez, Rafael C, Woods, Richard E. *Digital Image Processing.* Addison-Wesley Publishing Co., Reading, MA, 1987.
6. Brigham, E. Oren, *The Fast Fourier Transform and Its Applications*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1988, 448 pp.
7. Bracewell, Ron N., *The Fourier Transform and Its Applications*, McGraw-Hill Book Company, New York, 1965, 381 pp.
8. Blakemore, C. and Campbell, F. W., On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images, *Journal of Physiology,* 203, 237-260. 1969.
9. Wilson, Raymond G.; *Fourier series and Optical Transform Techniques in Contemporary Optics, an Introduction,* John Wiley & Sons, Inc., New York,1995.
10. Baher, H., *Analog & Digital Signal Processing,* John Wiley & Sons, New York, 1990.
11. Lathi, B. P., *Modern Digital and Analog Communication Systems, second edition,* Holt, Rinehart, and Winston, Inc., Chicago, 1989.
12. Williams, Charles S. and Becklund Orville A., *Introduction to the Optical Transfer Function,* John Wiley & Sons, New York, 1989.
13. Buren, W. F., G.E. Allen, W. H. Whitcomb, F. E. Lennartz & R. N. Williams. Zoogeography of the imported fire ant. *J. New York Ent. Soc.* 82: 113-124. 1974.
14. Slowik, T.J. *Response of the red imported fire ant (Solenopsis invicta Buren) to electricity and magnetism. Thesis, Texas Tech University, Lubbock, Texas, USA.* 1996.

# APPENDIX

# GLOSSARY

**AVI**.

An acronym for Audio/Video Interleaved. It refers to the file format used by Video for Windows™ in the recording and playback of interleaved digital video and audio data.

**Bucket Brigade**.

A line of people formed to fight a fire by passing buckets of water from a source to the fire.

**CCD**.

Stands for charge-coupled device. It is the imaging device made of light sensing pixels, which are used in the camera to capture images for video. Higher pixel counts usually mean better image and detail quality.

**CODEC**.

An abbreviation for COmpressor/DECompressor. Hardware and/or software which takes digitized video and produces a compressed file. It also decompresses the file back into digitized video. The same method must be used to decompress the data as was used to compress it.

**Compression**.

Compression is a process where a mathematical algorithm is applied to the pixels within an image (intraframe) or across multiple images (interframe) so that the amount of data needed to display the image(s) is reduced. Common compression formats are JPEG, MPEG and Wavelet.

**DCT**.

Discrete Cosine Transform.

**DV Video**.

DV is the digital video recording system used as a basis for the Mini DV, DVCAM and DVCPRO videotape formats. It uses 4:1:1 sampling and 5:1 DCT (Discrete Cosine Transform) interfield motion compression.

The term DV device is used generically when referring to either a DV camcorder or videocassette recorder (standalone deck) equipped with a FireWire interface.

**FFT**.

Fast Fourier Transform.

**JPEG**.

A standardized image compression mechanism. JPEG stands for Joint Photographic Experts Group, the original name of the committee that wrote the standard.

**Mesoptic**.

Mesoptic vision uses a combination of rods and cones. Used during dawn or twilight.

**MPEG**.

MPEG (the Moving Pictures Experts Group) is a group of people that meet under ISO (the International Standards Organization) to generate standards for digital video (sequences of images in time) and audio compression. In particular, they define a compressed bit stream, which implicitly defines a decompressor.

**Pan**.

The angle of camera rotating within the horizontal plane.

**Photopic**

Photopic vision is the use of the cones. Instinctively people look at objects in the day in a manner which focuses that object on the fovea.

**Primary Colors**.

Commission Internationale de l'Eclairage (the International Commission on Illumination) designated the following specific wavelength values to the three primary colors: blue = 435.8 nm, green = 546.1 nm, and red = 700 nm.

**PCM**.

An acronym for Pulse Code Modulation. This is a widely used system for encoding analog waveforms into digital data.

**RIFA**.

Red Imported Fire Ant.

**Scotopic**

Scotopic vision is the use of rods. Night vision. Must keep object off fovea and scan to keep image visible. If you stare you will lose the image due to a photochemical equilibrium of the rod which will stop sending nerve impulses to the brain.

**SCSI**.

An acronym for Small Computer System Interface. The is a interface standard which is used to control hard drives, CD-ROM's and tape drives. It is usually faster than the IDE and Enhanced IDE disk controllers. Original SCSI can access up to seven peripherals. SCSI-2 can access up to 15 peripherals with an increased bandwidth speed.

**SMPTE Time Code**.

SMPTE is an acronym for the Society of Motion Picture and Television Engineers. This is a time code standard used in professional video production. It is a digital code recorded on a videotape which identifies each frame as a unique and unchanging time value in HH:MM:SS:FF format, where HH is hours, MM is minutes, SS is seconds and FF is frames.

**Sony DV**.

One format of Sony Digital Video Camcorders.

**Spatial Domain**.

The aggregate of pixels compromising an image, and spatial domain methods are procedures that operate directly on these pixels.

**Tilt**.

The angle of camera rotated in a vertical plane.

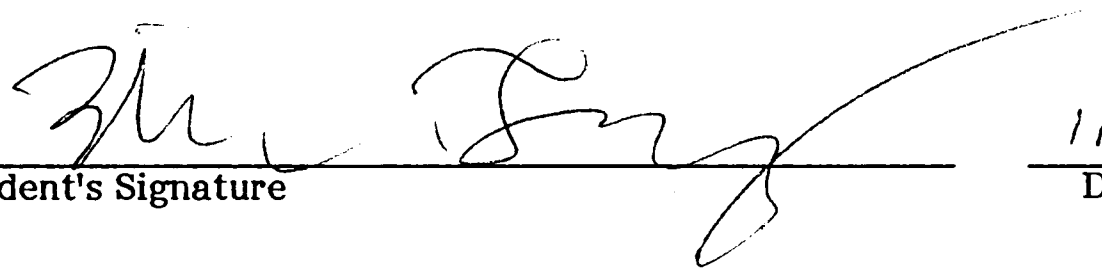
**Video**.

Sequences of images in time.

PERMISSION TO COPY

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Texas Tech University or Texas Tech University Health Sciences Center, I agree that the Library and my major department shall make it freely available for research purposes. Permission to copy this thesis for scholarly purposes may be granted by the Director of the Library or my major professor. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my further written permission and that any user may be liable for copyright infringement.

Agree (Permission is granted.)



Student's Signature

11/15/1999

Date

Disagree (Permission is not granted.)

Student's Signature

Date