

**DESIGN AND DEVELOPMENT OF VIDEO ACQUISITION SYSTEM FOR AERIAL
SURVEYS OF MARINE ANIMALS**

by

Sagar Agghera

A Thesis Submitted to the Faculty of
The College of Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degree of
Master of Science

Florida Atlantic University

Boca Raton, Florida

December 2011

**DESIGN AND DEVELOPMENT OF VIDEO ACQUISITION SYSTEM FOR AERIAL
SURVEYS OF MARINE ANIMALS**

by

Sagar Aghera

This thesis was prepared under the direction of the candidate's thesis advisor, Dr. Hari Kalva, Department of Computer and Electrical Engineering and Computer Science, and has been approved by the members of his supervisory committee. It was submitted to the faculty of the College of Engineering and Computer Science and was accepted in partial fulfillment of the requirements for the degree of Master of Science.

SUPERVISORY COMMITTEE:

Hari Kalva, Ph.D.
Thesis Advisor

Ankur Agarwal, Ph.D.

Bassem Alhalabi, Ph.D.

Borko Furht, Ph.D.
Chair, Department of Computer and Electrical Engineering and Computer Science

Mohammad Ilyas, Ph.D.,
Interim Dean, College of Engineering and Computer Science

Barry T. Rosson, Ph.D.
Dean, Graduate College

Date

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor, Dr. Hari Kalva, for his support and direction. He has been patient, encouraging and a great source of motivation throughout my tenure here at FAU. He has inspired me from his calm and serene demeanor even in toughest of situations. His technical and editorial advice was essential to the completion of this master's thesis and has taught me innumerable lessons and insights.

I would like to sincerely thank my committee members Dr. Ankur Agarwal and Dr. Bassem Alhalabi for their valuable comments, suggestions and input to the thesis. Thanks a lot for your patience and time.

I express my gratitude and love to my dearest parents – Mohanbhai and Bhanuben for their immense caring and support throughout my life. Thanks to valuable support and encouragement by my brothers – Akash and Parth.

I would like to thank Rafael, Waazim, Reena and Asif for their valuable support for design of system. I would also like to thank all my colleagues from Multimedia lab – Velibor, Alexander, Rashid and others. Thanks to my friends - Saravana, Kandarp, Vaibhav, Ranjith, Venkat and Payoj for supporting me.

ABSTRACT

Author: Sagar Aghera

Title: Design and Development of Video Acquisition System for Aerial Surveys

Institution: Florida Atlantic University

Thesis Advisor: Dr. Hari Kalva

Degree: Master of Science

Year: 2011

The planned deployment of turbines in the Florida Gulf stream could disturb the pattern of sea-turtle migration and reproduction. Therefore, a comprehensive study of the distribution of sea turtles is essential. Aerial counts and associated tasks are currently performed using human observations and distance sampling theory. The results of the observations made by human observers during flights are error-prone due to visibility biases. Video recordings have been introduced to circumvent some of the problems mentioned above and ensuring a systematic approach, frame-by frame if needed, to minimizing count errors. The process is still time-consuming, tedious, and error prone, due to observer's fatigue. The goal of the proposed work is to improve such methodology by approaching the problem from an automated video analysis perspective. This work focuses on the first step in automating aerial surveys: design and development of video acquisition system for aerial surveillance. The proposed solution consists of camera mounted on airplane which captures high resolution images at up to 15 frames per second

and the captured images are to be compressed and saved to data recorder in real time.

GPS is also integrated with the system to tag the images with GPS locations for biologists to crosscheck turtle sightings. This work also evaluates image compression schemes and the tradeoffs in compression performances based on computing required and storage capacity that will help in compressing the recorded data in real time.

DEDICATION

This thesis is dedicated to my wonderful family.

“You have been with me every step of my way, through good and bad times. Thank you for all the unconditional love, guidance, and support that you have always given me, helping me to succeed and instilling in me the confidence that I am capable of doing anything I put my mind to. Thank you for everything. I love you all!”

DESIGN AND DEVELOPMENT OF VIDEO ACQUISITION SYSTEM FOR AERIAL SURVEYS OF MARINE ANIMALS

FIGURES	viii
TABLES	ix
I.INTRODUCTION	1
1.1 Motivation	1
1.2 Background	1
1.3 Scope of the thesis.....	3
1.4 Glossary of terms.....	4
II.PROBLEM DESCRIPTION	7
2.1 Biologist requirements	7
2.2 Environmental challenges	9
2.3 Geometry layout.....	10
2.4 System requirements	10
III.SYSTEM DESIGN AND ARCHITECTURE	12
3.1 Selection of camera	13
3.1.1 SLR cameras vs. industrial cameras vs. camcorders	13
3.1.2 Global vs. rolling shutter	14
3.2 Selection of lens	15
3.2.1 Manual vs. auto iris.	16
3.2.2 Filters usage.....	17
3.3 Recording system	18
3.4 Software design	19
3.4.1 Multithreaded application.....	21
3.5 GPS.....	22
3.6 Camera mounting	23
3.7 Power supply	25
3.8 Camera calibration experiments.....	27
3.8.1 Geometry verification.....	27

3.8.2 Camera parameters verification.....	29
3.8.3 Hardware verification.....	31
IV.SYSTEM MODELING	33
V.DATA STORAGE.....	39
5.1 Saving raw images	39
5.2 Jpeg implementation	40
5.2.1 JPEG implementation using OPENCV	41
5.2.2 JPEG implementation using INTEL IPP	43
5.3 Lossless compression of bayer image	50
5.3.1 Bayer image compression using INTEL IPP.....	50
5.3.2 Preprocessing of bayer data.....	52
5.3.2.1 Frame difference method.....	53
5.3.2.2 Mean difference method.....	54
5.3.2.3 Reduced scale method	56
VI.CONCLUSION AND FUTURE WORK	59
6.1 Conclusion.....	59
6.2 Future work	61
BIBLIOGRAPHY	62

FIGURES

Figure 2.1 Survey area and transect line path	8
Figure 2.2 Geometry layout	10
Figure 3.1 System design and architecture	12
Figure 3.2 Rolling shutter vs. global shutter.....	15
Figure 3.3 Software sequence diagram.....	22
Figure 3.3 Camera mounting on wing of aircraft	25
Figure 3.4 Schematic of geometry experiment.....	28
Figure 3.5 Snapshot of geometry experiment	28
Figure 4.1 Sample geometry layout	33
Figure 4.2 Field of view for two cameras and spatial overlapping.....	35
Figure 4.3 Angle of view and geometry layout	36
Figure 4.4 Focal length and angle of view.....	37
Figure 5.1 Compression scheme	40
Figure 5.2 JPEG compression scheme.....	46
Figure 5.3 Performance analysis for JPEG compression.....	49
Figure 5.5 Histograms of R, G and B channels	55

TABLES

Table 1 Motion blur vs. Exposure time	30
Table 2 Calibration results	31
Table 3 Pixel distribution over 90 meter strip for different camera resolutions	34
Table 4 Focal length and angle of view calculations	37
Table 5 Storage requirements for saving tiff format.....	39
Table 6 Performance analysis – OPENCV implementation	41
Table 7 Storage requirements (JPEG format).....	42
Table 8 Performance analysis of YUV JPEG compression.....	44
Table 9 Performance analysis of RGB JPEG compression	44
Table 10 Performance analysis for half resolution RGB JPEG.....	47
Table 11 Performance analysis for half resolution YUV JPEG	47
Table 12 Performance analysis for quarter resolution RGB JPEG.....	47
Table 13 Performance analysis for quarter resolution YUV JPEG	48
Table 14 Comparison between different JPEG compression methods	49
Table 15 Performance analysis – execution time and compression ratio	51
Table 16 Huffman vs. Huffman +RLE	52
Table 17 Performance analysis – frame difference method.....	54
Table 18 Performance analysis – mean difference method	56
Table 19 Performance analysis – reduced scale method	57
Table 20 Storage requirement for different compression methods.....	58
Table 21 System requirements addressed	60

I.INTRODUCTION

Sea-turtles are currently facing major threats such as the destruction and alteration of their nesting and foraging habitats. The proposed deployment of underwater turbines in the Florida Gulf Stream could affect the pattern of sea-turtle migration and reproduction. Therefore, a comprehensive study of the distribution of sea turtles is planned.

1.1 MOTIVATION

The current state of the art is to have biologists fly on an aircraft and record the sightings of sea turtles. The population counts are extrapolated using the distance sampling theory. The results based on the observations made by human observers during flights are error-prone and the counts can be inaccurate, due to visibility bias. The other approach currently used involves taking video footage from airplane and then the biologists observe these tapes offline and record the sightings. These tapes are usually five to six hours long making it overall fatigue experience for the biologists. Automated and reliable video acquisition system is required here to assist biologists doing aerial surveys.

1.2 BACKGROUND

Aerial survey is a widely used technique based on distance sampling theory, which allows estimating species population and distribution (Buckland et al., 1993). It is very

useful in marine environments where displacement of observers is compromised at least when comparing against ground environments. Marine animal surveys have been done by biologists before, and some of the recent work include surveying of sea turtles in North Carolina (Chester et al., 1995) and Western Mediterranean (Cardona et al., 2005), as well as whales in Alaska (Hobbs et al., 2000). All of these studies implement a similar survey design with some difference depending on the target and environment conditions. The process is entirely manual where biologists make observation from aircraft, looking down through the aircraft window. These surveys are often repeated regularly to keep track of changing populations.

These aerial surveys have similar characteristics such as airplane speed, altitude, number of observers, distance of observations, and length of surveys. For instance, turtles observation in North Carolina (Chester et al., 1995) and West Mediterranean (Cardona et al., 2005) have similar airplane speed and altitude around 75 knots and 500 ft respectively, while seals and whales observation are performed faster and higher, around 100 knots speed and 600 to 800 ft altitude. Regarding minimum weather conditions, all surveys use the Beaufort scale to monitor sea conditions and are carried out if the estimated value is less than 3 (gentle breeze). Other common characteristics found were the use of GPS coordinates to register sightings locations and the use of inclinometers to determine the location of sightings with maximum precision.

Some attempts to automate the process have been tried by recording surveys in videotape for manual post analysis as is in Belugas counting in Alaska (Hobbs et al., 2000). In this case, two cameras with different zoom area were mounted in the airplane.

Initially, the videotape camera models used were Cannon Hi-8 814XL-S and Ricoh Hi-8 R800-H with 8x magnification for the zoomed camera. Later, video camera models were substituted with Sony Digital 8 DCR-TRV103 and Sony Digital Camcorder DSR-PD100A with 5x magnification. Improvements in the process were achieved by manually analyzing the whole recorded video and comparing against live counting during the trip. To the best of my knowledge, there is no other related work with the same broad scope as the current project.

1.3 SCOPE OF THE THESIS

The goals and scope of the thesis are as follows:

The goal of this project is to design and develop a system to record aerial video, analyze the video, and establish an accurate population count of the sea turtles. The video analysis can be used to determine an automatic count or to aid biologists to verify and correct their visual observations.

The scope of this thesis is limited to the design and development of an aerial video surveillance system. The system presented will be mounted under the wings of a Cessna 337 Skymaster (a four seat aircraft) flying 500-550 feet above the Atlantic Ocean to record the footage of the sea turtles in ocean. The operating environment, the requirements placed by biologists, and the need for robustness and reliability make this a very challenging task.

As the survey duration is 6 hours, a lot of data is expected and hence management of the data in real time is very crucial. The scope of thesis also includes evaluation of different compression methods implemented to help in managing the data in a cost effective way.

The proposed thesis serves as a guide for building aerial surveillance system irrespective of domain of use.

1.4 GLOSSARY OF TERMS

Some key terms which will be useful in understanding of the thesis are given below:

- **Angle of view/field of view:** It is the extent of the view taken in by a lens. The focal length of a lens, in conjunction with film size or image sensor size, determines the angle of view.
- **Aperture size:** A circle-shaped opening in a lens through which light passes to strike the image sensor or the film. Enabling the aperture to be made uniformly wider or narrower will let in more or less light.
- **Bayer Image:** It is raw output of the Bayer filter cameras. Since each pixel is filtered to record only one of three colors, the data from each pixel cannot fully determine color on its own.
- **Beaufort scale:** is an empirical measure for describing wind speed based mainly on observed sea conditions. Lower the number then better is environmental condition over sea coast making it helpful for carrying out aerial surveys.
- **Burrows Wheeler transform (BWT):** It is also known as block sorting algorithm. The Burrows-Wheeler transform is applied on blocks of input data (symbols). It is usually the case that larger blocks result in greater compressibility of the transformed data at the expense of time and system resources.
- **DISK IO speed:** It is a measure of Read/Write time onto the disk drive.

- **Exposure time:** The effective length of time a camera's shutter is open to expose the sensor to light for capturing image.
- **F-Number:** F-numbers indicate the size of the aperture in relation to the focal length of the lens. A smaller number indicates a larger aperture diameter.
- **Gain:** It is factor to compensate the low exposure level done post exposure to sensor by camera software. Excessive amount of gain introduces noise to an image as it is software generated pixel regions which were missed by sensor.
- **General Interval Transform (GIT):** This lossless data compression algorithm compresses blocks of data at ratios that are close to Huffman encoding compression ratios. However, with low entropy source data, the GIT algorithm typically provides coding at higher compression ratios as compared with the Huffman algorithm.
- **Huffman Encoding:** Huffman coding refers to the use of a variable-length code table for encoding a source symbol (such as a character in a file) where the variable-length code table has been derived in a particular way based on the estimated probability of occurrence for each possible value of the source symbol.
- **Iris:** It is a device inside a lens of thin overlapping metal leaves that move inwards or outwards, creating an aperture of variable size. The aperture size controls the amount of light passing through the lens to the sensor.
- **Move To Front transform (MTF):** It is a technique of encoding data which is used along with other data compression algorithms to improve compression performance. The main idea is that each symbol in the data is replaced by its index in the stack of “recently used symbols”. For example, long sequences of identical symbols are

replaced by as many zeroes, whereas when a symbol that has not been used in a long time appears, it is replaced with a large number.

- **PSNR:** Peak Signal-to-Noise Ratio is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. It is most commonly used as a measure of quality of reconstruction of lossy compression codecs.
- **RGB JPEG image:** It is referred to the resultant JPEG image out of the compression of R, G and B color channels.
- **Transect line:** A transect is a path along which one records and counts occurrences of the phenomena of study. It requires an observer to move along a fixed path and to count occurrences along the path and, at the same time, obtain the distance of the object from the path.
- **White Balance:** The camera adjusts the overall scene's color balance so that the areas meant to be reproduced as white in the picture will be white, thereby also adjusting all the other colors in the scene using the same color shift values, so that all colors are accurately represented.
- **YUV JPEG image:** It is resultant JPEG image out of the compressing of YUV buffer with Luminance (Y) and Chrominance components (U & V).
- **Ziv-Lempel-Storer-Szymanski (LZSS):** It is a lossless data compression algorithm which uses dictionary encoding technique. It attempts to replace a string of symbols with a reference to a dictionary location of the same string.

II.PROBLEM DESCRIPTION

2.1 BIOLOGIST REQUIREMENTS

A primary requirement is to emulate human observation so that established distance sampling techniques can be used in estimating population distribution. The aerial survey technique basically consists of flying an aircraft along a well-defined path known as a transect line to record sea turtle sightings found in an observation area known as the Strip at a given distance from the transect line. The current application includes 250 meters wide Strip of ocean to be observed on each side of transect line. Biologists tend to observe turtles away from line of transect due to the fact that the turtle don't come up for breathing just below flight because of noise of airplane. It is also difficult to observe directly below the flight as observation is carried out through side windows. Cameras mounted can help the biologists to capture data just below the flight where human observers are not able to collect data.

.

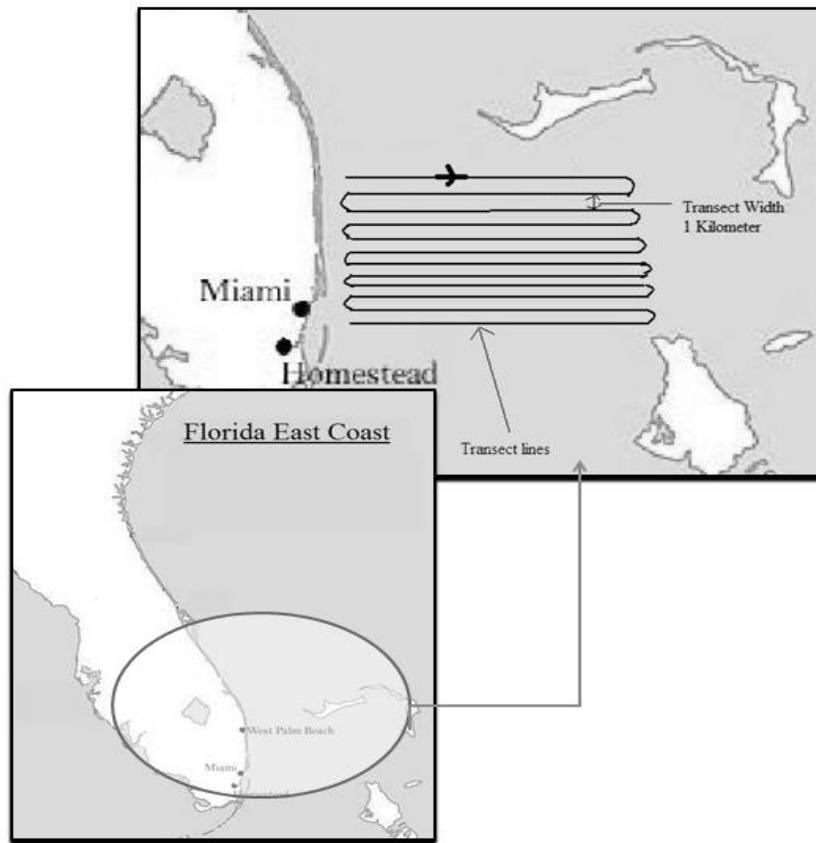


Figure 2.1 Survey area and transect line path

In the proposed system, cameras in the airplane must follow the transect lines and must record the target objects found on both sides of the transect line inside the strip area. To obtain accurate population estimate using the distance sampling theory, it is also very important to record as accurate as possible the location at which the object is found and the distance of target from transect line. Since the biologists also record the species of the sea turtles and their observed location, the proposed solution should have very high resolution color cameras to distinguish among different species and a GPS to record the position. Human observers on board do not have the time or resources to configure or monitor the system. The result is the system should record upon power-up and adjust automatically to changes in environmental conditions.

2.2 ENVIRONMENTAL CHALLENGES

The quality of the images collected will depend on many factors that can be classified as physical and optical. The physical factors such as airplane vibration, wind, waves, altitude and speed pose as environmental challenges. The optical factors are represented as sunlight, reflection from ocean surface, and brightness variations. All these factors need to be addressed with the correct selection of camera suitable to handle each of these problems. Environmental conditions affect directly both the physical and optical factors and subsequently the quality of the images recorded. Designing to address the environmental condition become the most important challenge since selecting an appropriate type of camera and other settings in the system, the system should be able to adapt to any expected variation in environmental conditions. Ocean conditions can change during the trip and the system has to automatically adjust to these variations without human intervention.

2.3 GEOMETRY LAYOUT

As shown in Figure 2.2, the strip is 250 meters wide on each side of transect line. The altitude of airplane is 500-550 feet with a speed of 115-125 mph.

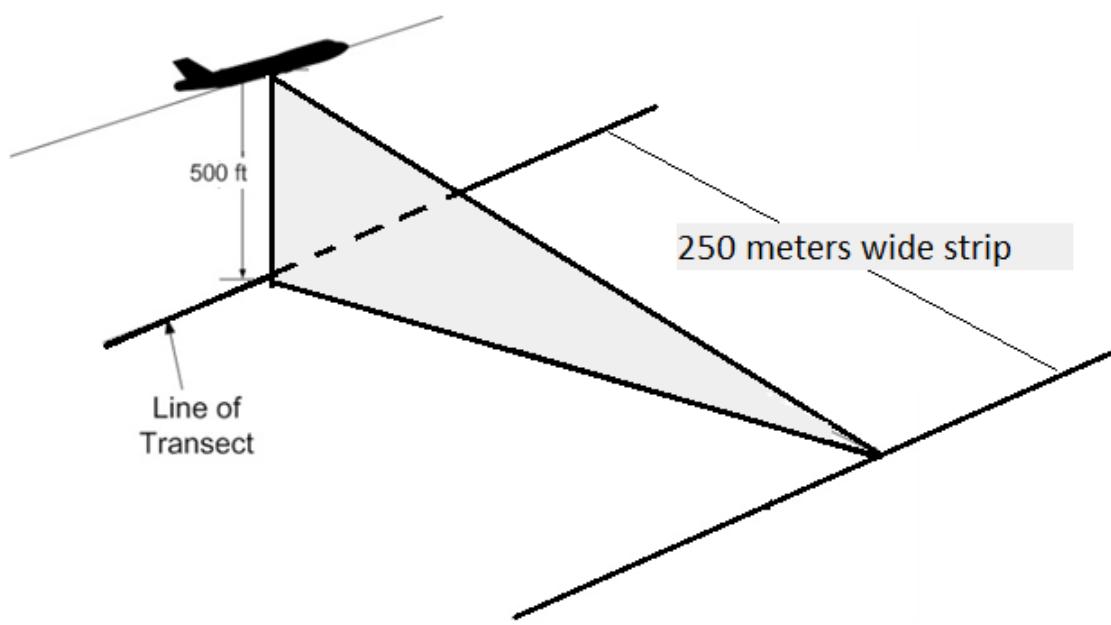


Figure 2.2 Geometry layout

2.4 SYSTEM REQUIREMENTS

Based on the biologists' requirements and taking into consideration the environmental conditions, the system requirements are

- **Sea turtle identification and classification:** Camera should be able to identify and possibly classify the turtle species found in the captured images with a given altitude of flight and size of turtle.

- **Robust system:** The system should be robust enough to operate for six continuous hours without showing any signs of overheating of sensor. Programmatically controlling camera parameters is a challenge; six hours of survey will produce a wide variety of intensity of input light. System should also withstand interruptions created due to lose electrical connections or power failures.
- **Autonomous operation:** An automatic system with minimal interference from a human observer is desired.
- **Offline video reviews:** The collected data from survey should be helpful to biologists to crosscheck their findings with the recorded data which includes the location of sightings and possibly verify the species of the turtle. Eventually, the system should be able to automate the process of aerial surveys to replace human observers and hence enhance the aerial surveillance procedures.
- **Camera protection:** The camera system should not get damaged or the loss of data should be avoided due to presence of debris and moisture or water particles present in air.
- **Minimize system size and weight:** As huge amount of data is expected, efficient compression algorithms are necessary for compressing data in real time to manage with disk memory available.

III.SYSTEM DESIGN AND ARCHITECTURE

In this chapter, we present system design and architecture of the proposed aerial surveillance system. As shown in figure 3.1, the system comprises of a console that is interfaced with multiple cameras, GPS and WIMAX ADAPTER. WIMAX capability is not included in current design and is planned to be implemented in future. The components of the system are described in the following sub-sections.

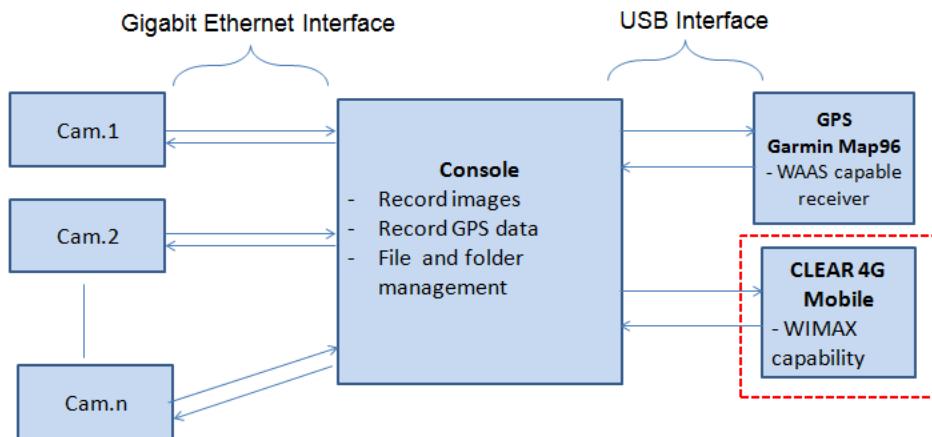


Figure 3.1 System design and architecture

3.1 SELECTION OF CAMERA

The quality of the images collected will depends on many factors. First, physical factors like airplane vibration, wind, altitude and speed. And optical factors like sunlight reflection and brightness variation. These factors need to be faced with the correct selection of cameras suitable to handle each of these problems. While considering different for cameras, possibility beyond visible spectrum should not be ignored. IR camera is one of such possibilities beyond visible spectrum which is widely used for surveillance. While in a case of ocean applications one thing to consider is that the IR cannot go beyond few millimeters of the water layer. This makes it unsuitable for aerial surveillance of sea animals. Focus is then on visible spectrum.

3.1.1 SLR cameras vs. industrial cameras vs. camcorders

Selection of camera is not a trivial task. Initially DSLR camera from Nikon and Canon brands were considered for their high resolution and high speed performance. There are some special requirements for our application for which DSLR cameras are not suitable. These requirements include:

- Continuous operation for 6 hours – DSLR cameras are not designed for continuous operation over 6 hours. With long continuous use the sensors are known to get hot and cause camera malfunction.
- Interface with PC for saving stored images in real time – DSLR cameras have minimal support for saving data directly to hard disk.

- Electronic shutter – with long continuous use, the mechanical shutters in DSLRs are known to break
- Camera weight – with the cameras mounted outside the aircraft and the aircraft moving at 120 miles per hour, minimizing the weight of the cameras is critical to minimize vibration and use light weight mounts.

Industrial cameras are designed for use in special application where user can programmatically control camera parameters and can work in rugged environmental settings for long period of time. Hence the use of a DSLR camera was ruled out.

Industrial Cameras come with two types of shutters namely

- Global Shutter
- Rolling shutter

3.1.2 Global vs. rolling shutter

CCD sensor cameras use electronic shutters which consist of two types of shutters, namely Global shutters and Rolling shutters. In sensors with rolling shutter, all pixels in one row of the imager collect light during exactly the same period of time, but the time light collection starts and ends is slightly different for each row. This creates rolling or skew effect which distorts the image with moving objects (as airplane in this application). While in sensors with Global Shutter, all pixels in imager are exposed simultaneously for same amount of time and hence don't have rolling or skew effect on the image. The example of rolling effect is shown in figure 3.2 given below [12]:

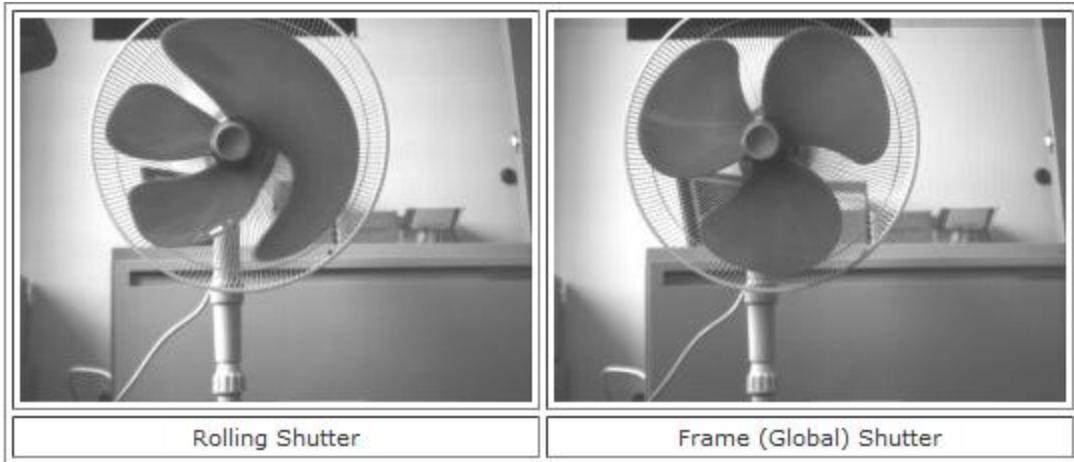


Figure 3.2 Rolling shutter vs. global shutter

Design Choice: We chose AVT GC2450C Industrial Camera. Following are the key features of AVT GC2450C camera:

- High resolution - 5 Megapixels (2448x2050)
- Fast – 15 frames per second
- Progressive Scan – Global Shutter
- Gigabit Ethernet interface
- Programmatic Control over camera parameters
- Designed to work for several hours of operation

3.2 SELECTION OF LENS

Based on the geometry shown in figure 2.1 and the sensor size of camera, lenses need to be selected. C-mount Pentax fixed focal length lenses were opted, which goes with the AVT GC2450C Industrial Camera. The angle of view of the lens is calculated based on the width of the strip covered by an individual camera and height of the flight. Based on the angle of view, focal length of the lens is calculated. Lenses with focal length of

25mm, 35mm and 50mm were chosen to cover the width of the strip with three cameras as defined in geometry. This configuration results in successive images to have 90 percent temporal overlap when recorded at 15 frames per second.

3.2.1 Manual vs. auto iris.

Optical lens come with two options for aperture control.

- Manual Iris Lens

The Manual Iris lens is the simplest type of iris control lens with manual adjustments to set the iris opening in a fixed position. The Manual Iris lens is convenient to use when it is not necessary to adjust the lens continuously for correct picture brightness and hence their application is where the surrounding light doesn't vary a lot.

- Auto Iris Lens

The Auto Iris Lens is a type of lens that can be electronically controlled. This type of lens allows maintaining the lighting level. Auto iris lenses are the best lens to use in any application as they are flexible and can adjust to changing light levels. For aerial surveillance application, light doesn't vary a lot.

Minimal variation in light due to different time of the day and position of sun is compensated by auto exposure mode of the camera and hence Manual Iris lens are suitable for this application. Auto Iris lens adds another parameter to be controlled programmatically which adds complexity to the system. The cost of manual iris lens is half of the auto iris lens which suits the proposed cost effective method of surveying.

3.2.2 Filters usage

The most commonly used filters for digital photography include polarizing (linear/circular), UV/haze, neutral density, graduated neutral density and warming/cooling or color filters. Few applications for each are listed below:

- Linear and Circular Polarizers - Reduce glare, improve saturation.
- Neutral density – Extend exposure time
- Graduated neutral density – control strong light gradients, reduce vignette.
- UV/haze – provide lens protection and improve clarity.
- Warming/cooling – change white balance

From the mentioned commonly used filters, circular polarizing filters are helpful for our application in a way, that it can reduce the glare created by the reflections from waves. Additional filters can be used, but filters do effectively introduce an additional piece of glass between your camera sensor and the subject. It also has the potential to reduce image quality. Hence we will restrict to use of only circular polarizing filters for our application.

Design Choice: Pentax CCTV lens with fixed focal length of 25mm, 35mm and 50mm were chosen to cover desired field of view as per geometry layout given in section 2.3. Manual iris with circular polarizing filters was added to address system requirements.

3.3 RECORDING SYSTEM

A recording console is integrated with the cameras and GPS devices in order to synchronize and coordinate them and collect data from these units. As the survey conducted is a data intensive process that also requires high computation for the processing of compressed images, the recording console should have a very good processing speed with large storage capacity (in range of TBs).

The key requirement for storage devices are reliability and recording speed. Reliable data storage is a mandatory requirement as the trip cannot be repeated. There are two main options for storage: namely solid state disk drives (SSD) and hard disk drives (HDD). Solid state drives are faster compared to hard disk drives when taking disk read speed into consideration. But when one considers disk write speed they are comparable. As of February 2011, the SSD costs around \$2/GB and the costs increase as the capacity increases while the HDD cost around \$0.10/GB. Moreover the SSD comes up to 2TBs in size while HDD comes in different size up to 3TBs. Hence the HDD proves to be a better option when the price as well as size is taken into consideration. Since we need an inexpensive solution that provides consistent and reliable data, option of solid state drive is ruled out. Two regular SATA, RAID 0 hard disk drives with a total capacity of 2 TB were selected which is fast and has enough space for all the data that are captured for the six hour trip.

Number of cameras and the storage requirements for console is also determined by Disk IO speed. Simple tool called PassMark Evaluation Test tool was used to measure the read and write speed for HDD [8]. The measured DISK IO speed for the console was 150MB/sec. The processor is also another consideration while choosing the data recorder.

To compress the images in real time the processor needs to be fast enough to process frames in real time. Intel has come up with quad core processor which has four cores for higher performance.

Design Choice: The system specifications of the PC chosen for aerial survey are

CPU Type: Intel Core i7 860 @ 2.80GHz

CPU Speed: 2793.7 MHz

Cache size: 256KB

O/S: Windows 7 (64-bit)

Total RAM: 4055.1 MB

Storage Memory: 2TB

3.4 SOFTWARE DESIGN

The camera software is designed to be robust and handle every event that could possibly disrupt the functioning of the system. Some factors affecting the software design are listed below:

- As the collector application runs using the camera SDK and is handling huge amount data at a fast rate for six continuous hours, there is a chance that the collector application may crash at some stage due to limits of processing capabilities or out of memory conditions.
- The camera may stop collecting data due to overheating and may restart. This should not affect the working of other cameras and the stopped camera should also start collecting data.
- The operating system may also crash, if camera threads are not handled properly.

- The power disruption should not affect the working of software. Essentially the camera and GPS software should record on power up.
- Storage management is also very important as the amount of data captured is huge (in order of TBs).
- Time synchronization of data from GPS and camera is also critical part of collecting data.

To address the issues of rebooting of system due to power fluctuations the camera and GPS software are designed as startup service and hence on rebooting of system the software starts recording data automatically. Additionally, the BIOS setting of the operating system is opted for BOOT ON POWER UP configuration so that the PC resumes the data collection even after power failure. To address issue of handling multiple cameras independently, threading was introduced. Hence each camera runs as an independent thread which is not affected by events such as other camera getting unplugged due to loose connection or overheating of sensor. The figure 3.3 shows a software sequence diagram for the camera software.

Another important feature of the camera software is file and folder management. The folder names are associated with time stamps so that if system restarts after system failure, the data will be saved in separate folder. The image files and the GPS data files are named using time stamps which uses system time. Hence the GPS and camera data can be time synchronized using the recorded timestamps.

The folder and file structure is as shown below:

```
..\\SurveyName[Timestamp]  
..\\CameraID[Timestamp]  
..\\ImageFiles[Frame counter....Timestamp]
```

Additionally, the configuration file (cfg) file was created for easy access to camera parameters and other settings such as jpeg quality.

3.4.1 Multithreaded application

The camera software works on the concept of threading which helps in handling multiple cameras with single code/application. Whenever the camera is plugged in, a new thread is created which initializes the camera and starts streaming and capturing frames with user defined settings from the configuration file. The software is designed in such a way that it can handle multiple cameras by creating different threads which work in parallel.

Whenever the camera is unplugged it closes the thread smoothly without disturbing other running threads.

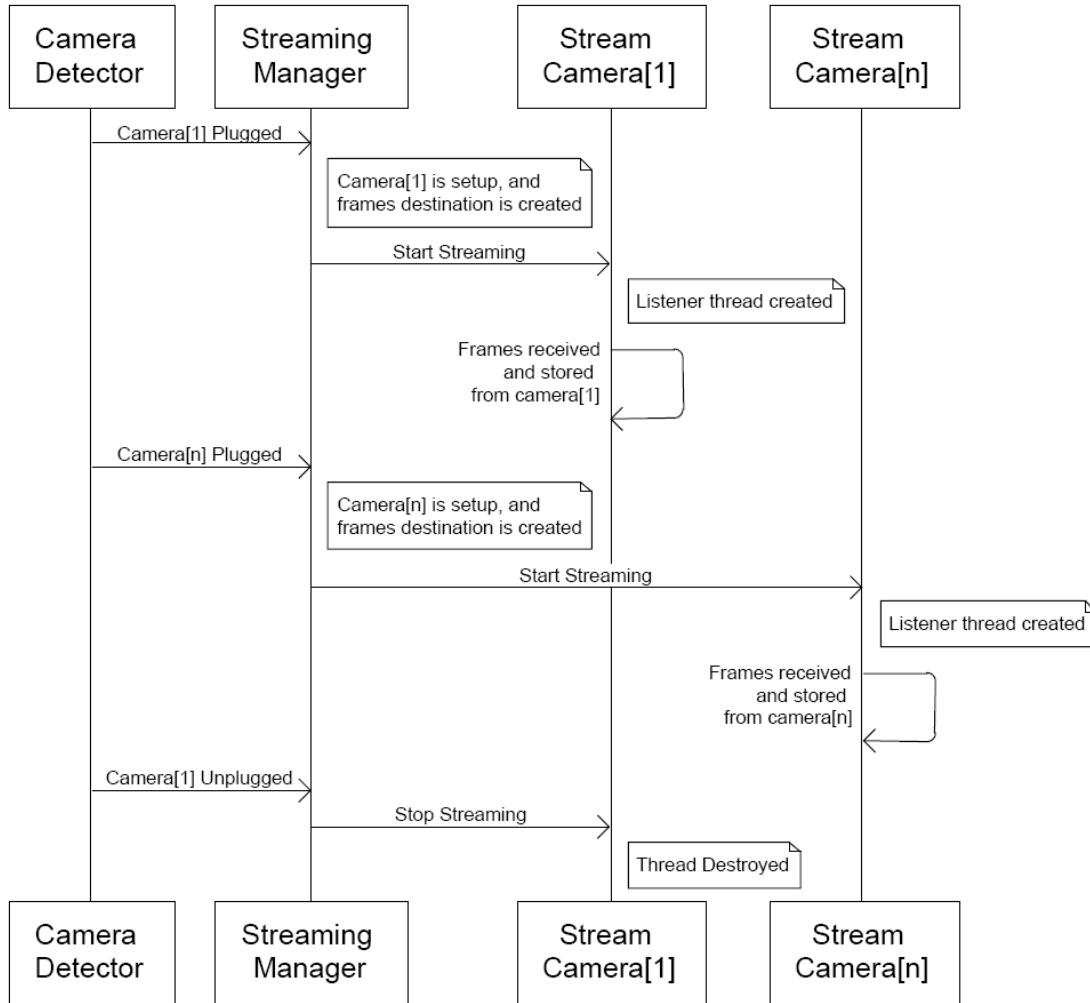


Figure 3.3 Software sequence diagram

Design Choice: Multithreaded application software with file and folder management was built to enable operator free run. Software application was made to run as startup service and hence will run autonomously when system is booted.

3.5 GPS

GPS is also an important part of Data Collection system to tag images with proper location co-ordinates which are necessary for cross verifying turtles found by biologists

and also for data post processing stages. We chose Garmin GPSMAP N96, which is Aviation GPS which has WAAS capable receiver. WAAS (Wide Area Augmentation System) is a system of satellites and ground stations that provide GPS signal corrections, giving a better position accuracy. WAAS capable receiver is considered the best for open land and marine applications. Update rate for this GPS is every second and the flight speed is 125 mph which means that the flight moves 50 meters every second. WAAS capability gives the accuracy of 16 meters laterally and 4 meters vertically hence update rate of one per second make ideal for tracing path and tagging the turtle sightings. Human observers also use handheld GPS to tag the turtle sightings with accuracy of seconds. GPS is manually programmed in such a way that the data recorded is stored in real time to PC through USB connection. For this application, Flight position and altitude is recorded every second by GPS and sent to PC where the session is recorded in a text file. The GPS is tagged with system time and hence is synchronized with the images captured by camera software.

Design Choice: Garmin GPSMAP 96 – Aviation GPS

3.6 CAMERA MOUNTING

The aircraft to be used for aerial survey is Cessna m-337 Skymaster which will fly at a speed of 100-110 knots (115-125mph) at 500-550 feet height. Only mounts available are bomb racks and wing mounts. There are two options to mount cameras onto the airplane. One option is inside the airplane and other option is outside the airplane. Cessna m-33y Skymaster has very small space inside (mostly for passengers and pilot) which makes it difficult to mount cameras inside. Moreover, the visibility also decreases because of glass

windows. Hence mounting the cameras outside is the only option. Gyro stabilizers which serve as camera mounts are used in airplanes and helicopters which work well to dampen vibrations. These stabilizers cost several thousand dollars and increase the cost of the system. Additionally, all gyro stabilizers on the market are designed for operating inside the aircraft and are bulky when the weight of the mounts is considered. We need to mount cameras outside the airplane to get clear view and proper angle rather than seeing through window glass. Hence wing mount was one of the options to mount camera.

Figure 3.3 shows the mounting of cameras on wings for our aerial surveillance project. Wing mounts are ideal for mounting the cameras as there is no engine present on wings to produce any vibration effects. To make sure cameras are properly placed at proper angle, pan and tilt mount is used to adjust angle of camera to capture the desired field of view. The mounts and the short exposure time help in eliminating the vibration effects on the recorded images.

Camera Enclosures were used to protect cameras from wind, dust and water particles while flying. Test trip was conducted to evaluate the performance of the overall system in terms of power usage and camera stability. At altitude of 500-550 feet the airplane was stable and hence the system was not exposed to heavy vibrations due to wind turbulence. The results showed that one of the two cameras was vibrating. In future, we expect to conduct more experiments to eliminate any kind of vibration effect on camera. As far as data recorder is concerned, the system was very stable and no issues were recorded.



Figure 3.3 Camera mounting on wing of aircraft

Design Choice: Camera enclosures and pan & tilt mount from APG manufacturer were used to protect camera.

3.7 POWER SUPPLY

Power supply on the Cessna Skymaster consists of two outlets of 12V and 24V each with maximum power drainage of 600W. The equipments needed to be powered up are storage console, cameras, and the GPS. AVT cameras use 3.5W and PC power requirement is 350W. GPS is powered up by AA batteries. Hence for 2 camera system Maximum power = $3.5 + 3.5 + 350 = 357W \approx 360W$.

The DC to AC inverter of 12V and 400W can be used to power up the entire Video Surveillance system. Experiments were conducted using the DC power source in a car to test the power requirements of the system. The system was powered up by using 400W DC to AC inverter from car's 12V DC outlet and was tested successfully. Similar experiment was conducted on airplane where the system was powered up by 12V DC outlet from airplane. As the cameras are mounted outside the plane (on wings), length of the cables is also important consideration. AVT cameras come with specialized power cables for cameras and hence before placing the order one need to make sure that the length of the power cable from adapter to camera need to be long enough to reach to wings of the plane. In this work, length of the wings and the placing of the system on the flight were recorded and accordingly the length of the Ethernet cables and power cables were chosen. One important thing to take into consideration is to check and verify internal electrical circuits of the flight as the flight models used for surveys are quite old and there is possibility of power leakage and non-functioning of components like power converters and sockets on plane. Good practice is to keep extra power inverter to run the system on plane.

Design Choice: 400W DC to AC inverter and power adapters for cameras.

3.8 CAMERA CALIBRATION EXPERIMENTS

The key requirement is that the system should be able to work without any assistance continuously for 6 hours of operation. Hence video acquisition system needs to be calibrated properly prior to survey. Key parameters to verify include camera parameters, overall software and hardware components of the system.

3.8.1 Geometry Verification

The cameras along with lenses are chosen in order to cover desired field of view. To verify the geometry, reduced scale setup was created as there was no arrangement to get elevation of 500 feet. Geometry verification experiment was carried out by taking pictures from top of Engineering building EE96 at Boca Raton campus of Florida Atlantic University. Using measurement tapes, field of view covered in the image captured was measured and verified with the available theoretical readings through lens and camera specifications. The schematic of experiment is shown in figure3.4. The sample of captured image is shown in figure 3.4. Pixel distribution estimation was also verified by obtaining the dimensions of stop sign seen in figure 3.4 and comparing with the pixels in captured image.

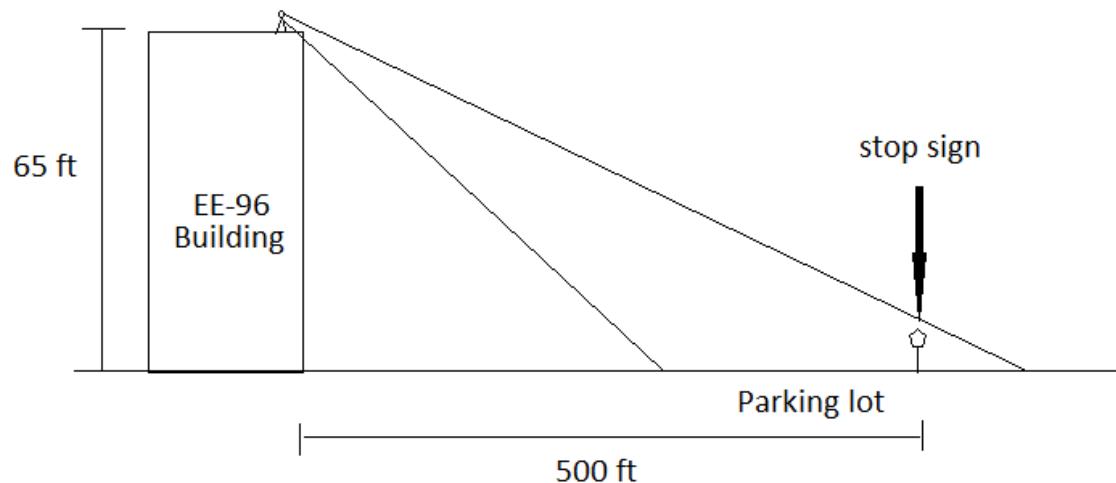


Figure 3.4 Schematic of geometry experiment



Figure 3.5 Snapshot of geometry experiment

3.8.2 Camera parameters verification

Camera parameters including exposure time, gain value and white balance value affect the quality of captured images. Hence these parameters have to be properly calibrated to get best set of images. Other parameters include lenses focus and aperture size which controls incoming light on sensor. A series of camera experiments were carried to evaluate following factors which affect image quality.

- Brightness/ contrast of image
- Image sharpness (check on blurring effect)
- Color – White balance
- Effect of Noise – gain balance

The AVT camera comes with auto controlled features such as exposure time, white balance and gain. The roof top experiments were carried out to test performance of different modes of camera operations. More the camera is exposed to light better will be brightness in image. But as the exposure times increase in motion photography (as in domain of aerial surveys), the motion blur is introduced in image affecting the image quality. The airplane moves at speed of 110 knots and with the field of view known, one can calculate the amount of motion blur at different exposure levels. The table-1 documents the behavior of motion blur at different exposure levels.

Table 1 Motion blur vs. exposure time

Motion blur vs. Exposure time at Airplane speed of 110 knots		
Exposure time (in microseconds)	Motion blur (in feet)	Motion blur (in pixels)
50	0.009	0.099
500	0.09	0.99
1000	0.18	1.98
2000	0.36	3.96
2500	0.45	4.95
3000	0.54	5.94

As the size of the target can be as small as 1 feet, we need exposure blur as low as possible. Hence cap is needed to control the exposure time. Motion blur of not more than 1 pixel was taken into consideration to control picture quality. Hence for the given exposure time with cap of 500 microseconds (derived from table3.1), the images were captured at different aperture level to check for brightness in captured images. Experiments were conducted for both lenses by fixing the exposure time manually and varying aperture size to get optimal aperture value. Next step was to use circular polarizing filter along with lenses and similar experiments were conducted to get optimal aperture value (with usage of filters). Underexposed images showed sensor artifacts which were eliminated by increase in aperture. After usage of filters, at same aperture sizes the images were underexposed to light and hence the gain was also taken into consideration.

With addition of gain, the noise level also increases. Experiments were carried out under controlled exposure time and aperture size to check for effect of different gain levels on the picture quality. The picture quality was tested by zooming into the captured image and observing the image at pixel level. Thus the cap on the gain level was also

introduced. The colors in the captured images were not sharp. Hence the experiment was carried out by exposing camera to colored object under auto white balance mode. The lack of the sharp colors was observed in captured images due to malfunctioning of the white balance. Set of experiments were carried out by varying the white balance red and white balance blue manually and taking readings of the color quality of resulting image. Optimal values of white balance red and blue were chosen based on the results from conducted experiments.

From series of calibration experiments carried out over period of time, the results derived are listed in table-2 given below.

Table 2 Calibration results

Camera and lenses parameters	Desired value or range
Exposure time and mode	Auto exposure mode with cap of 0.5 milli seconds
Aperture size for lenses	Little more than maximum aperture size
Gain value and mode	Auto gain mode with cap of 6 on scale of 0-24
White balance	Manual mode
Polarizing Filter	Necessary to avoid reflection from bright objects.
Focus of Lenses	Set to infinite (to capture distant objects)

3.8.3 Hardware verification

The outdoor experiments were carried out to check for any hardware failures during the operation. On-boat experiments and car experiments were part of the outdoor experiments. With the help of the team from Ocean Engineering Department, the system

was mounted on sea vessel and survey was carried out in turbulent ocean waters to check for any hardware failures as well as software operation. Similarly, car experiments were also carried out to check for any hardware issues.

Additionally, acquisition system was also tested on the airplane for short flight. The working of the camera mounting equipments and the power requirements of the system were practically evaluated. Issue of the power conversion from AC to DC power occurred during test. Different set of power inverters were used to test the system. Issue was successfully resolved which included the defect in internal electrical circuit of voltage converter on the airplane.

IV.SYSTEM MODELING

This chapter presents the modeling aspects in the design of the aerial surveillance system.

Figure 4.1 shows the field of view from the airplane and the area to be covered by camera(s). Factors such as the height and speed of the airplane, the strip location, and width and the size of the target –sea turtles – determine the components required for the data acquisition system. Consider a case of covering 150 meter wide strip at distance of 150 meters from line of transect, as used in aerial survey method for sea turtles in inshore North Carolina waters [4].

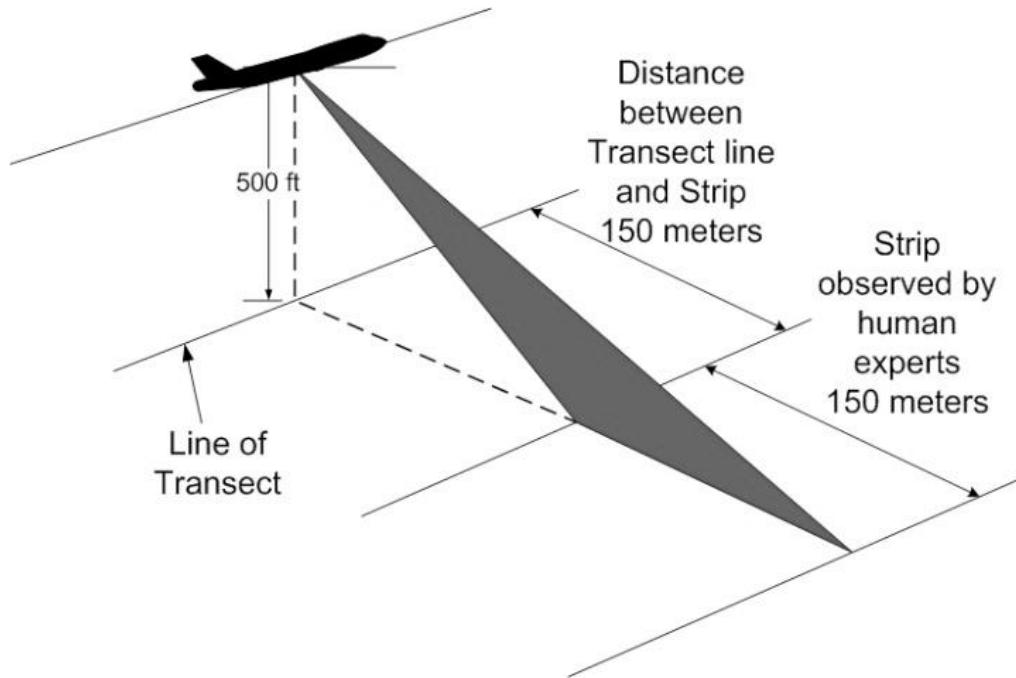


Figure 4.1 Sample geometry layout

First step in process of system modeling is camera selection. Multiple factors affect the process of camera selection. Factors such as size of target (here sea turtles), strip width and airplane are the factors affecting the camera selection process. The size of target varies from 1 foot long juvenile loggerhead turtles to leatherback turtles which grows as long as 6.6 feet. The camera should be able to capture images at high enough resolution to enable object detection algorithms detect targets in an image in the post processing stage. One foot object was considered as smallest target to be detected in an image. To detect one foot object in 150 meter wide strip is not a trivial task. It is highly desirable to have more than one camera to capture the target at high enough resolutions. Minimum of 10 pixels/feet was chosen to be enough to detect one feet object. For 90 meter strip the pixel distribution is shown in the table-3 given below

Table 3 Pixel distribution over 90 meter strip for different camera resolutions

Strip Width	Avg. Pixels/ft for 2K resolution	Avg. Pixels/ft for 2.5K resolution	Avg. Pixels/ft for 3K resolution	Avg. Pixels/ft for 4K resolution
90 m	8	11	13	16

Hence two cameras with 2.5K resolution are sufficient to cover the 150 meter wide strip. 30 meter spatial overlapping between field of views of two cameras was taken into consideration to avoid loss of data to error in orientation of camera. The figure given below explains the distribution of field of view for two cameras to capture 150 meter wide strip.

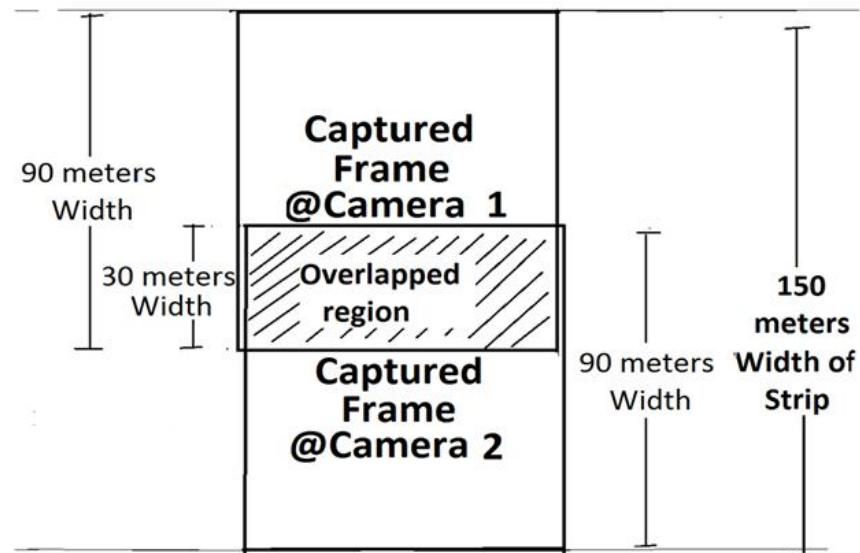


Figure 4.2 Field of view for two cameras and spatial overlapping

Once camera is selected the next step is to get appropriate optical lenses which would help to cover the field of view accurately. Key considerations while selecting lenses are the angle of view, focal length and aperture. Based on the altitude of airplane (500ft), distance of the strip from transect line (150m) and field of view for two cameras, angle of view for the two lenses can be calculated. For application where field of view is already defined, fixed focal length lenses with adjustable aperture is suited. Next step would be calculation of angle of view for a given geometry. Figure 4.3 gives a layout of geometry and angle of view for individual cameras. Angle of view can be calculated using the formulae given in equation (1) and (2).

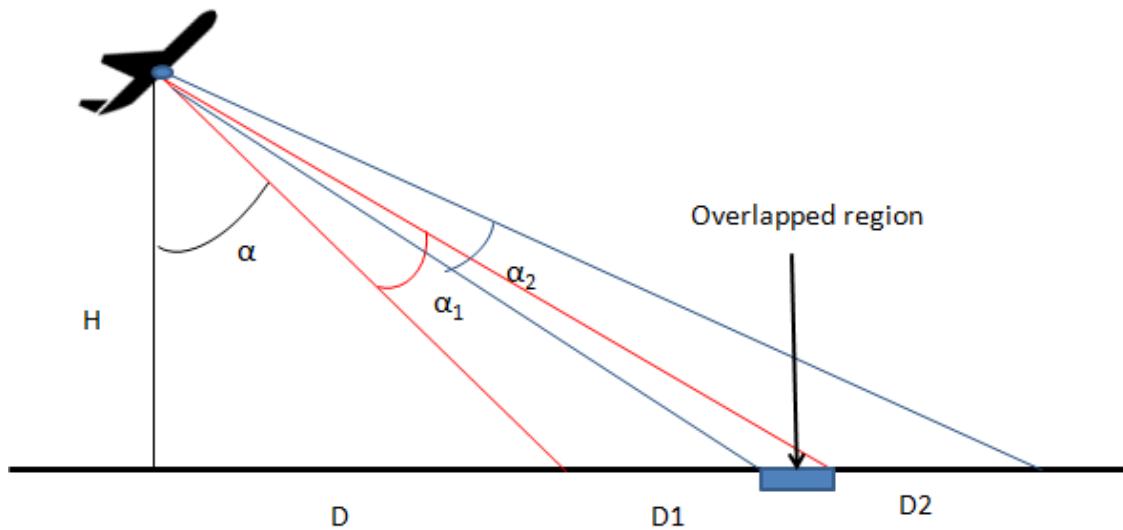


Figure 4.3 Angle of view and geometry layout

$$\alpha_1 = \tan^{-1} ((D+D1)/H) - \tan^{-1} (D/H) \quad (1)$$

$$\alpha_2 = \tan^{-1} ((D+D1+D2)/H) - \tan^{-1} ((D+D1)/H) \quad (2)$$

Focal Length can be calculated from angle of view using the formula given in equation (3). Focal length and angle of view are inversely related and their relation is shown in figure 4.4. For the given problem, the focal length and the angle of view calculated is listed in table-4.

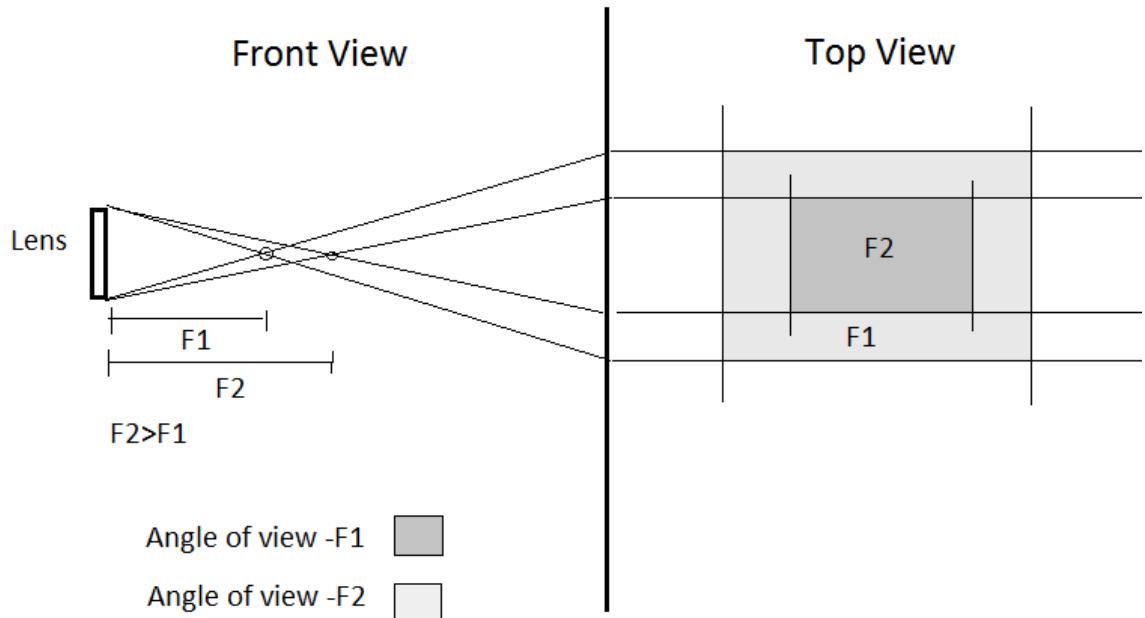


Figure 4.4 Focal length and Angle of view

$$\text{Focal length (in mm)} = (2 * \tan(\alpha/2))/d \quad (3)$$

Where,

d = sensor size (horizontal, vertical or diagonal)

α = angle of view (horizontal, vertical or diagonal)

Table 4 Focal length and angle of view calculations

Focal length and angle of view calculations				
Camera Number	Distance of strip from line of transect (in meters)	Width of strip (in meters)	Angle of view for Camera/Lenses	Selection of Lenses(in mm)
2	210	90	9	50
1	150	90	13	35

Other than getting image data, biologist requirements include getting position of the target identified or detected. As the survey carried out on the sea away from land using airplane moving at high speed (100-110 knots), regular terrestrial GPS won't give accurate positions. Hence aviation GPS which are designed for airplanes is necessary for getting accurate data at fast rate. Garmin Aviation GPS MAP96 is low cost aviation GPS which can be used, update rate is also high (one time every second). GPS is interfaced with recording console to log data real time and also to time synchronize with frames captured.

The flight duration is 5-6 hours and hence huge amount of data is expected. For two cameras system running at 15 fps for 6 hours duration as many as 650,000 images is captured. Hence sufficient storage memory is pre requisite for configuring PC. Additionally, these cameras are generally mounted outside the airplane, possibly on wings. Hence enclosures designed for rugged environmental conditions are needed to protect camera from moisture, dust particles and wind thrusts.

V.DATA STORAGE

5.1 SAVING RAW IMAGES

The AVT camera SDK does come with a function to save the RGB frame buffer to image file in tiff format. As the sensor resolution is 5 MP, the size of the RGB color image would be three times the sensor resolution. Hence each image file will be 15 Megabytes. Table-5 shows the storage space requirements at different frame rates for two camera system. For capture rate of 15 fps (maximum frame rate), 9 Terabyte of storage space is required and makes the system heavy and expensive. Hence compressing images in real time is quite essential to make camera run at maximum frame rate.

Table 5 Storage requirements for saving tiff format

Analysis of Storage requirements for 2 camera configuration (in GB)	
Duration @ fps	TIFF format
6 hrs @3 fps	1817
6 hrs@ 6 fps	3634
6 hrs @ 9fps	5452
6 hrs @ 12fps	7269
6 hrs @ 15fps	9086

Compression of images is one of the essential parts of the data storage due to huge amount of data collection during survey operation of 6 hours. The compression scheme is shown in the figure5.1. AVT camera gives output in form of bayer image (WxH) with RGGB pattern as seen in figure 5.1. This bayer image after demosaicing gives R, G and

B data which after up sampling gives three color channels, each one of WxH dimensions. With the help of JPEG encoding these channels can be compressed to form YUV and RGB images. Alternately, the bayer image can be directly compressed using lossless data compression algorithm into compressed binary data.

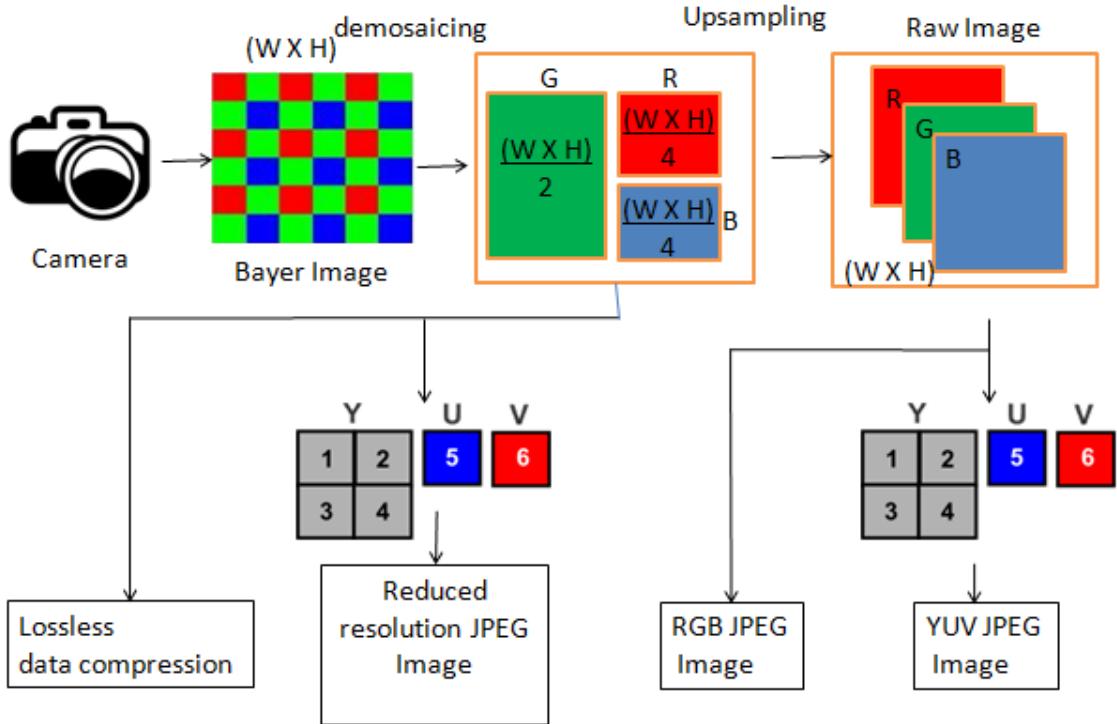


Figure 5.1 Compression scheme

5.2 JPEG IMPLEMENTATION

Compressing images may be lossy or lossless. Lossless image compression is used for archival purposes where one wants to avoid compression artifacts. It is mostly used in medical imaging and technical drawings. Lossy image compression is used in photography of natural images where minor loss of fidelity is acceptable. In applications where the size of the compressed image is also critical, lossy image compression is

desirable to use. One of the most commonly used method lossy image compression is JPEG

5.2.1 JPEG implementation using OPENCV

OPENCV (Open source Computer Vision) is a library of varied functions mainly aimed at real time computer vision. It has a BSD license (free for commercial or research use). OPENCV libraries were originally written in C but now have a full C++ interface and all new development is in C++ [13]. OPENCV comes with simple functions to create and save image buffers to different formats among which JPEG format is widely used. JPEG implementation with different compression factors were achieved using OPENCV library. The compression achieved for different compression factors is given in table-7. Two cameras were interfaced with PC and were made to run in continuous capture mode. The analysis showed that the camera was able to capture at maximum rate of 3 frames per second as each time compression took 0.371 seconds to save single frame. The performance analysis results are shown in table-6.

Table 6 Performance analysis – OPENCV implementation

Time for individual steps in 'save as jpg' release mode	
Process	Time (in secs)/Frame rate
creating rgb image from bayer buffer	0.062
rgb to yuv image	0.098
cVsavelImage	0.211
Total time	0.371/2.7

In order not to miss even a single event (turtle coming up to the ocean surface for breathing) we want a camera to capture at maximum available frame rate which is 15 frames per second. Hence OPENCV implementation is inadequate for this application.

Table 7 Storage requirements (JPEG format)

Storage requirements (JPEG format) for 2 camera configuration			
Duration @fps	Compression factor	Filesize	Total storage (in GB)
6 hrs @3 fps	100	2500	309
6 hrs@ 6 fps	100	2500	618
6 hrs @ 15fps	100	2500	1545
6 hrs @3 fps	99	2400	297
6 hrs@ 6 fps	99	2400	593
6 hrs @ 15fps	99	2400	1483
6 hrs @3 fps	95	1625	201
6 hrs@ 6 fps	95	1625	402
6 hrs @ 15fps	95	1625	1004
6 hrs @3 fps	85	1200	148
6 hrs@ 6 fps	85	1200	297
6 hrs @ 15fps	85	1200	742

5.2.2 JPEG implementation using INTEL IPP

The other option for JPEG implementation is using Intel IPP. Intel Integrated Performance Primitives (Intel IPP) is an extensive library of highly optimized software functions for multimedia, data processing, and communications applications on multicore processors. Intel IPP offers many optimized functions covering frequently used fundamental algorithms which also includes JPEG algorithm. Intel IPP functions are designed to deliver performance beyond what optimized compilers alone can deliver, by matching the function algorithms to low-level optimizations based on the processor's available features such as Streaming SIMD Extensions (SSE) and other optimized instruction sets [9]. The JPEG encoder application was optimized to work on multicore processor by utilizing all four threads for processing data. The JPEG implementation was divided into three step process. Firstly, the buffer was converted to R, G and B channels and then RGB image was converted to YUV format and then compressed to YUV JPEG image. Alternately, RGB image obtained in second step was compressed to RGB jpeg. Performance analysis was carried out by measuring the execution time for compressing each frame at different jpeg quality factors and the results are listed in table-8 and table-9. Maximum frame rate of 15 was not achieved for any of the compression quality factors due to large size of the buffer (15MB). PSNR was also used as performance measurement tool for JPEG analysis. The YUV JPEG hence stored was compared with up sampled raw image as mentioned in figure 5.1 and PSNR was calculated based on the difference of two.

Table 8 Performance analysis of YUV JPEG compression

JPEG Quality	PSNR	Compressed file size(in KB)/ Bits per pixel	Execution time (ms)/Frame rate (fps)
100	56.53	4160/6.79	0.110/9
97	56.35	2420/3.95	0.099/10
95	56.32	1845/3.01	0.097/10
90	56.15	1110/1.81	0.083/12
85	56.03	800/1.30	0.076/13
80	55.89	655/1.06	0.067/14
75	55.62	550/0.89	0.065/15

Table 9 Performance analysis of RGB JPEG compression

JPEG Quality	PSNR	Compressed file size(in KB)/ Bits per pixel	Execution time (ms)/Frame rate (fps)
100	56.69	8825/14.40	0.170/6
97	56.63	6300/10.28	0.160/6
95	56.61	5100/8.32	0.155/6
90	56.55	3490/5.69	0.140/7
85	56.48	2745/4.48	0.125/8
80	56.35	2290/3.73	0.112/9
75	56.31	1980/3.23	0.105/10

Alternatively, the R, G and B channels available from demosaicing were compressed to half and quarter resolution as shown in compression scheme given in figure 5.2. In first case of half resolution jpeg, the height of R and B channels were up sampled from H/2 to H resulting into all three channels with $(W/2) \times (H)$ dimensions. Resolution was reduced from 2448X2050 to 1224X2050. The resultant channels were compressed as RGB jpeg and YUV jpeg. In second case, the height of the green sample was reduced to half from original height resulting into all three channels with $(H/2) \times (W/2)$ dimension. Hence the resultant data to be compressed to RGB and YUV JPEG with quarter resolution to the original data. Resolution was reduced from 2448X2050 to 1224X1025. The JPEG compression scheme is shown in figure 5.2. INTEL IPP functions were used for up sampling and down sampling.

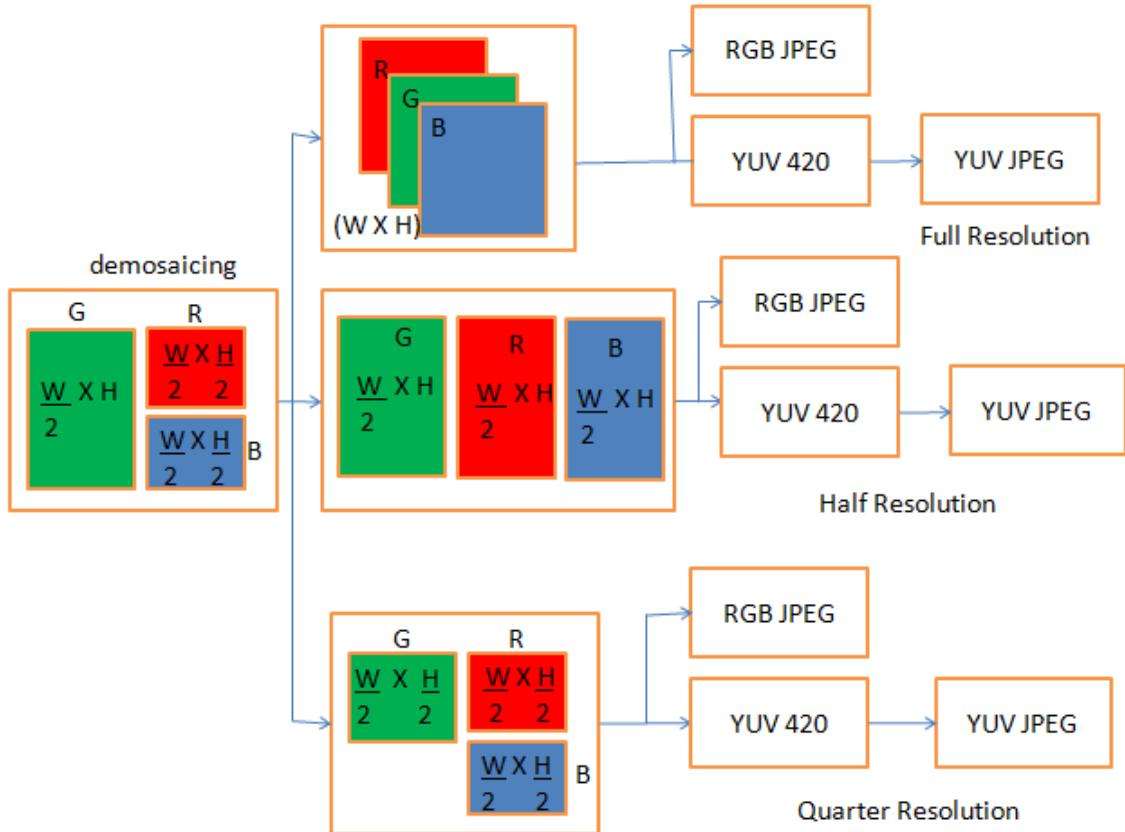


Figure 5.2 JPEG compression scheme

The performance analysis for all four cases is shown in tables-10 to 13. Once again, PSNR was used as tool for comparison for different JPEG quality and compression factor was the other measurement tool. As we have compressed R, G and B data as YUV jpeg image, getting back the image at original resolution for PSNR calculation requires additional tasks. YUV and RGB jpeg files were decoded from jpeg image the individual channels were separated and scaled to original size (2448x2050) for PSNR calculation. MATLAB and IMAGE MAGICK [15] were used for decoding the original image and resizing to desired scale.

Table 10 Performance analysis for half resolution RGB JPEG

JPEG Quality	PSNR	Compressed file size(in KB)/ Bits per pixel	Execution time (seconds)/Frame rate (fps)
100	56.05	5120/8.35	0.081/12
97	55.94	3735/6.10	0.066/15
95	55.88	3105/5.06	0.063/16
90	55.74	2240/3.65	0.059/17
85	55.65	1760/2.87	0.054/18
80	55.57	1460/2.38	0.049/20
75	55.53	1240/2.02	0.046/22

Table 11 Performance analysis for half resolution YUV JPEG

JPEG Quality	PSNR	Compressed file size(in KB)/ Bits per pixel	Execution time (ms)/Frame rate (fps)
100	55.73	2170/3.54	0.056/18
97	55.58	1300/2.12	0.046/22
95	55.55	1010/1.69	0.043/23
90	55.32	565/0.92	0.037/27
85	55.26	385/0.63	0.034/29
80	55.17	300/0.49	0.033/30
75	54.95	240/0.40	0.030/33

Table 12 Performance analysis for quarter resolution RGB JPEG

JPEG Quality	PSNR	Compressed file size(in KB)/ Bits per pixel	Execution time (ms)/Frame rate (fps)
100	55.65	2880/4.70	0.050/20
97	55.53	2055/3.35	0.044/22
95	55.49	1720/2.80	0.043/23
90	55.32	1225/1.99	0.039/26
85	55.23	930/1.51	0.036/28
80	55.14	735/1.19	0.032/31
75	55.08	585/0.95	0.030/33

Table 13 Performance analysis for quarter resolution YUV JPEG

JPEG Quality	PSNR	Compressed file size(in KB)/ Bits per pixel	Execution time (ms)/Frame rate (fps)
100	55.21	1075/1.75	0.033/30
97	54.92	635/1.03	0.029/34
95	54.91	500/0.81	0.028/36
90	54.65	180/0.30	0.024/42
85	54.48	165/0.27	0.021/48
80	54.34	135/0.22	0.020/50
75	54.28	110/0.18	0.019/52

Hence six different methods were used to compress raw image as jpeg image file with different resolutions. For analysis all the compressed image files were compared with original raw image (BMP format). PSNR vs. Bits per pixel graph for jpeg quality of all six methods were plotted in a graph shown in figure5.3. Table-14 shows comparison between the execution time/ frame rate for each of the compression methods. Based on the memory space and frame rate, one can select the compression method for their application. Quarter RGB and YUV jpeg are more suitable to use for quick viewing purposes.

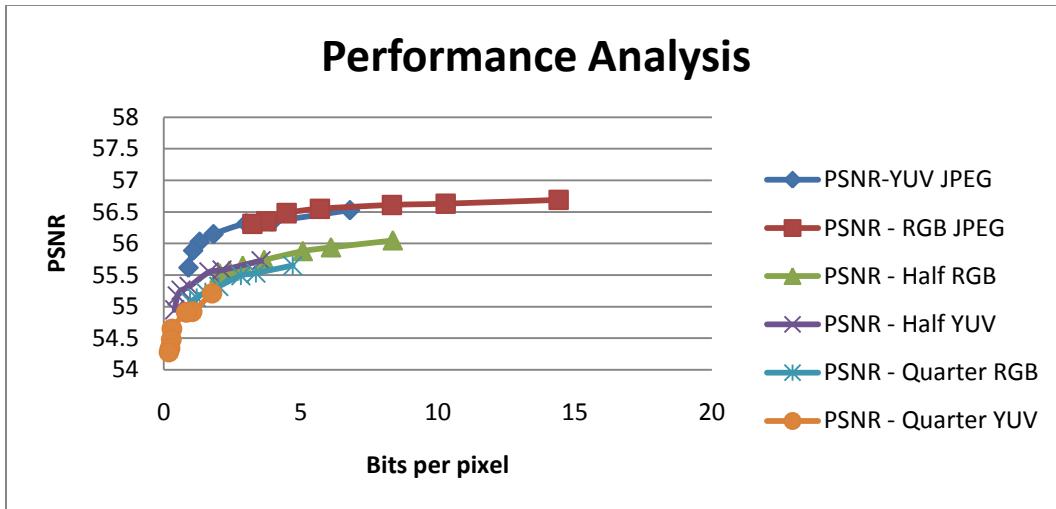


Figure 5.3 Performance analysis for JPEG compression

Table 14 Comparison between different JPEG compression methods

Compression method	Avg. PSNR	Avg. Bits per pixel/file size (in KBs)	Avg. execution time(in seconds)/frame rate (fps)
Full RGB	56.55	7.12/3577	0.140/7.14
Full YUV	56.15	2.69/1351	0.084/11.90
Half RGB	55.74	4.25/2135	0.059/16.95
Half YUV	55.32	1.43/716	0.037/27.07
Quarter RGB	55.32	2.35/1445	0.039/25.64
Quarter YUV	54.65	0.66/328	0.024/41.66

5.3 LOSSLESS COMPRESSION OF BAYER IMAGE

In most digital cameras, Bayer CFA images are captured and demosaicing is generally carried out before compression. The compression first approach is non-conventional approach where compression of pixel buffer is carried out before demosaicing. Lossless compression of bayer data before demosaicing is one approach to avoid loss of any possible data. An efficient prediction based lossless compression scheme for bayer images (King-Hong et al, 2008) can be used for bayer compression. Though better compression can be achieved using this scheme, it cannot be used in our case as it is more compute intensive and hence may not be ideal to use in real time data compression. Intel IPP data compression module was used in the evaluation of lossless compression methods in real time.

5.3.1 Bayer image compression using INTEL IPP

The Ipp_compress sample illustrates ways of implementing different effective lossless data compression solutions by using API of Intel® Integrated Performance Primitives Data Compression domain. The universal primitives implement the most popular and effective algorithms of lossless data compression without supporting any compatibility with the existing format of the compressed data [9].

The universal primitives are designed to be treated as building blocks for creating algorithmic chains that can be used as solutions for different lossless data compression issues. The set of the universal primitives used by the Ipp_compress sample includes

primitives for Huffman encoding/decoding, Burrows-Wheeler Transform (BWT), General Interval Transform (GIT) and Ziv-Lempel-Storer-Szymanski dictionary based compression algorithm (LZSS) [9]. The part of the Ipp_compress sample that illustrates the use of the universal primitives consists of modules. Each module corresponds to one of the compression algorithm implemented. Each module contains a top-level function for compression and decompression using the corresponding compression algorithm.

Sample binary data files captured from the camera were taken to test each of the methods based on compression ratio and execution time. Execution time was taken in to consideration; as matter of fact that the compression was carried out in real time with camera capturing at high frame rate. The results are shown in table-15. As interpreted from table-15, Huffman compression outperformed other algorithms in terms of execution time and also compression ratio is comparative better under given execution time. BWT and GZIP gave better compression but at cost of higher execution time.

Table 15 Performance analysis – execution time and compression ratio

Compression Method	Compressed file size	Compression ratio	Bits per pixel	Time taken (in seconds)
BWT	3458	1 : 1.42	5.64	0.351
GIT	3532	1 : 1.38	5.76	0.926
LZSS	5020	1 : 0.98	8.19	0.163
GZIP	3897	1 : 1.26	6.36	0.253
Huffman	3925	1 : 1.25	6.40	0.031

The data compression in application domain where it is needed to be carried out in real time, execution time is major factor to be taken into account. Hence Huffman and RLE algorithms come out very useful as they are not computationally intensive and also they are efficient in compressing images with less variation as in case of printers and scanners where most area is white. One way of implementing RLE along with Huffman algorithm is to modify BWT in IPP-data compression module. Comparison between Huffman and modified BWT algorithm by running both of them on image samples is shown in table-16. As observed from the table, Huffman outperforms the combined method (Huffman and RLE) when compression ratio is taken into consideration. To further enhance the performance, preprocessing of data is necessary before using Huffman algorithm for compression. For analysis, both the methods of compression (Huffman and Huffman+RLE) were used for performance analysis.

Table 16 Huffman vs. Huffman +RLE

Compression method	Original file size (in KBs)	Compressed file size (in KBs)	Compression ratio	Time taken (in seconds)
Huffman + RLE	4901	3925	80.09	0.044
Huffman	4901	3850	78.55	0.031

5.3.2 Preprocessing of bayer data

The Bayer data to be compressed is hence combination of R, G and B channels together in a RGGB Bayer pattern. There will higher correlation between elements of individual

channels compared to combined Bayer grid. First step of preprocessing will be separating R, G and B channels. The individual channels are then compressed to form three binary files for single frame. The experimental results show improvements in compression ratio with marginal increase in execution time. This can be observed from results shown in table 5.11. As we are aware of the application domain, more improvements are expected. Three preprocessing methods are proposed which can enhance performance of the compression algorithms. The data compression scheme is shown in figure

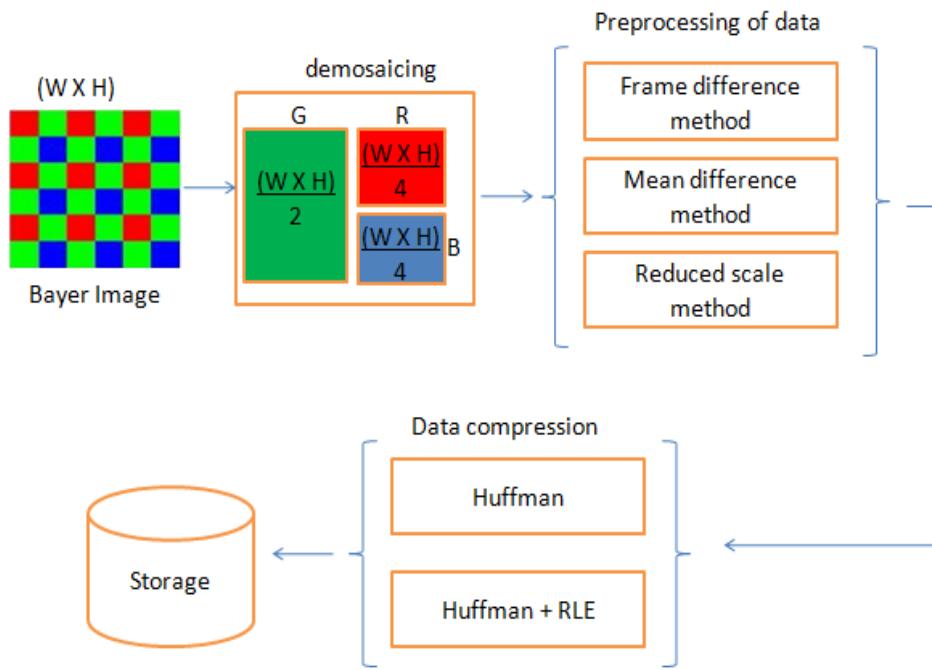


Figure 5.4 Data compression scheme

5.3.2.1 Frame difference method

The proposed frame difference method can be useful in domain where the images are captured in continuous mode as in case of marine aerial surveys. There is slight variation

in consecutive frames as the surveys are carried out in calm weather conditions. As the name suggests, the pixel wise difference in consecutive frames is taken in real time for three color channels and compressed instead of compressing original image. Slight variations occur as the moving plane capture images in continuous mode and hence better compression is expected. Set of 10 images were used to experimental basis and then the resulting frame difference data was compressed using Huffman algorithm. The results are shown in table-17. This method is pure lossless as original frame can be obtained in offline processing.

Table 17 Performance analysis – frame difference method

Compression method	Compression ratio	Bits per pixel	Execution time (in seconds)/frame rate (fps)
Huffman	1 : 1.347	5.94	0.036/27
Huffman+RLE	1 : 1.311	6.10	0.055/18

5.3.2.2 Mean difference method

The images captured during marine aerial surveys are dominantly blue in color. Hence the blue intensity is high compared to Green and Red channels. Histograms of three colors channels were observed for sample ocean images, one of which is shown in figure 5.5. Concluding from the histograms in figure 5.5, all the three channels have only one dominant slope. Similarly observations were also noted in other sample oceans images.

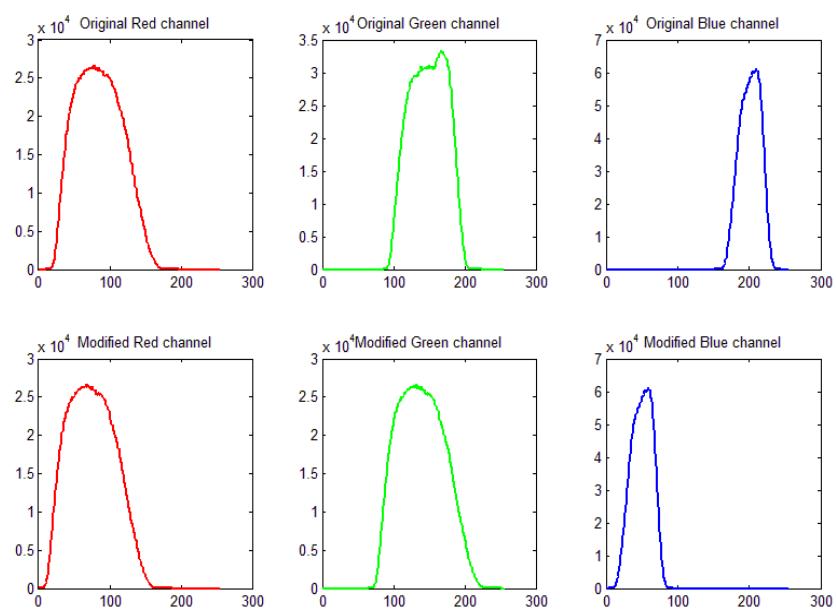


Figure 5.5 Histograms of R, G and B channels

The proposed mean difference method utilizes the dominant slope in histograms. Firstly, the mean of the pixel values are taken for individual color channel of captured image. In second step, the whole frame (pixel wise) is deduced from the mean value. The minimum value shift is used as third step to eliminate negative numbers. In final step the resulting data obtained in second step is compressed using Huffman and Huffman+RLE. For each frame, mean value and shift value are calculated and saved in separate text file which is useful in offline processing when the decompression is carried out to obtain original frame. The mathematical representation for this method is shown below:

$$A = \sum_{\substack{0 \leq i \leq m \\ 0 < j < n}} a(i, j) - \frac{1}{(m * n)} \sum_{\substack{0 \leq i \leq m \\ 0 < j < n}} a(i, j)$$

$$\bar{A} = A + \min \left\{ \sum_{\substack{0 \leq i \leq m \\ 0 < j < n}} a(i, j) \right\}$$

Alternatively, the mean value and shift values are calculated for 10 or 100 consecutive frames and saved to reduce the computation in contrast of calculating mean for individual frame as the ocean surface is flat with minimal change in scenery. This method leads to lossless compression. The experimental analysis was carried out on the same set of 10 images and compression results are documented in table-18. The flow of the steps is explained using pseudo code given below.

For each R, G , B channels

```
{
    For each element  $a_{ij}$ ,
         $a_{ij}^- = a_{ij} - \text{mean}(R);$ 
         $a_{ij}^+ = a_{ij}^- + |\min(R)|;$ 
}
```

Table 18 Performance analysis – mean difference method

Compression method	Compressi on ratio	Bits per pixel	Execution time (in seconds)/frame rate (fps)
Huffman	1 : 1.298	6.17	0.048/20
Huffman+RLE	1 : 1.276	6.27	0.078/12

5.3.2.3 Reduced scale method

In contrast to above two preprocessing methods, this method leads to near lossless compression. This is due to fact that the bit depth value of a pixel is reduced to 128 from

256. First step involves dividing each pixel values of the three color channels by two. In second step, they are compressed using Huffman algorithm. In offline processing, the compressed data is decompressed and the values are multiplied by 2 to regain original bit depth of 256. Far better efficiency in compression ratio is expected compared to other methods. The loss of data occurs in first step of preprocessing the data because of odd integers division by two and result being an integer. Hence this method can be counted as near lossless compression method. In worst case scenario when all pixels have odd values then the PSNR would be 48.13 dB. Best case scenario when all pixels have even values will result in the PSNR value of infinity (lossless). The experimental results over the same set of 10 images are documented in table-19.

Table 19 Performance analysis – reduced scale method

Compression method	Compression ratio	Bits per pixel	Execution time (in seconds)/frame rate (fps)
Huffman	1 : 1.523	5.25	0.033/28
Huffman+RLE	1 : 1.451	5.51	0.058/17

All the above mentioned three methods are not computationally intensive and hence the resulting increase in execution time is marginal as compared to compression without any preprocessing as can be observed in their performance analysis tables. In all three methods of preprocessing Huffman algorithm outperformed Huffman+RLE algorithm. Hence Huffman was chosen as the method for compression. Now as the compression ratio is known for the three methods, PC storage requirements for the aerial survey

operation can be estimated. The resulting storage requirement for two 2.5K resolution camera configuration is documented in table-20 given below.

Table 20 Storage requirement for different compression methods

Storage requirements for two 2.5K resolution camera configuration (in GBs)			
Compression method	6hrs duration @ 10 fps	6hrs duration @ 12 fps	6hrs duration @ 15 fps
Uncompressed	2019	2423	3029
Huffman	1598	1918	2398
RGB separate	1561	1873	2341
Frame difference	1499	1798	2248
Mean difference	1557	1869	2336
Reduced scale	1325	1590	1988

Hence for the PC with 2 TB memory space, the camera can run at maximum frame rate (15 fps) if one uses reduced scale compression method. Based on the available memory, desired compression method can be implemented at different frame rate by following the chart shown in table-20.

VI.CONCLUSION AND FUTURE WORK

6.1 CONCLUSION

This thesis presented the design and development for video acquisition for aerial survey. The requirements were systematically studied and components were selected. Extensive modeling and experimentation was carried out in finalizing the system design. The system requirements as described in section 2.4 are addressed by proposed thesis which is listed in table-21. It can be concluded that YUV half resolution jpeg implementation is comparable to lossless compression scheme with reduced scale implementation when comparing frame rate though jpeg implementation gives higher bits per pixel compared to lossless compression method.

Table 21 System requirements addressed

System requirements	Proposed solution
Sea turtle identification and classification	High Resolution Industrial cameras selected as outcome of system modeling can capture target with desired resolution
Robust system	Boot on power up configuration and application running as startup services make sure that system is robust enough to handle power interruptions.
Autonomous operation	Camera software is designed to run as startup and standalone application. Properly calibrated camera parameters make the system autonomous.
Offline Video reviews	File and folder management with time stamps and frame counters help in offline video reviews. Moreover the GPS tags help biologists to crosscheck their sightings with system data.
Camera protection	Camera enclosures with pan and tilt mount makes sure that the cameras are protected against the outside environment factors like wind, debris and moisture particles.
Minimize system size and weight	The lossy and lossless compression schemes presented in this thesis help in minimizing system size by compressing the collected data in real time.

6.2 FUTURE WORK

Currently the aerial surveillance project status is on verge of carrying out flight surveys. Hence the future work will include running the system in real time to record data. The compression algorithms can be fully tested during the flight surveys. Initial experiments to test the feasibility of the solution were conducted and it proved to be very solid. With the progress of the current project, we expect to gain valuable experience after carrying more number of surveys in calibrating the system to meet the biologist requirements. With minor modifications, the thesis can be also useful for aerial surveillance of other sea animals or even monitoring sea vessel traffic.

Future work also includes implementing the WIMAX coverage to monitor the surveys offline by sending the reduced scale images over network to base station. Improvements will be required once the system is used for surveys based on feedback from the biologists once they observe the recorded data offline.

BIBLIOGRAPHY

- [1] Buckland, S.T. et al. (1993). 'Distance Sampling, Estimating Abundance of Biological Populations'
- [2] Caughley, G. (1977), 'Sampling in Aerial Survey', *The Journal of Wildlife Management* 41(4), 605--615.
- [3] Cardona, L. et al. (2005), 'Western Mediterranean immature loggerhead turtles: habitat use in spring and summer assessed through satellite tracking and aerial surveys', *Marine Biology* 147(3), 583-591.
- [4] Chester, A. et al.(1995), 'Aerial surveys for sea turtles in North Carolina inshore waters', *Fishery Bulletin* 93(2), 254-261.
- [5] Elko, N. A. et al. (2005), 'Quantifying the Rapid Evolution of a Nourishment Project with Video Imagery', *Journal of Coastal Research* 21(4), 633-645.
- [6] Hobbs, R.C., et al. (2000). 'Beluga, Delphinapterus leucas, group sizes in Cook Inlet, Alaska, based on observer counts and aerial video'. *Marine Fisheries Review*. 62(3):46-59.
- [7] *Usage of filters* <cambridgeincolour.com/tutorials/camera-lens-filters.htm> (last accessed on 10/28/2010)
- [8] *Passmark Tool* <passmark.com> (last accessed on 05/12/2011)
- [9] *Intel IPP documentation* <software.intel.com/en-us/articles/intel-integrated-performance-primitives-documentation/> (last accessed on 07/13/2011)
- [10] *Glossary of photography terms* <photographytips.com/page.cfm/1587> (last accessed on 07/13/2011)
- [11] *Digital Image formats* <wfu.edu/~matthews/misc/graphics/format/format.html> (last accessed on 07/01/11)
- [12] *Rolling shutter effect* <ptgrey.com/support/kb/index.asp?a=4&q=115> (last accessed on 07/15/2011)

- [13] *OPENCV documentation* <opencv.itseez.com/> (last accessed on 04/27/2011)
- [14] King-Hong Chung; Yuk-Hee Chan, "A Lossless Compression Scheme for Bayer Color Filter Array Images," *Image Processing, IEEE Transactions on* , vol.17, no.2, pp.134-144, Feb. 2008
- [15] *Convert, edit and compose images* <imagemagick.org/script/index.php> (last accessed on 07/27/11)