NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

AN ADAPTIVE METHOD FOR THE ENHANCED FUSION OF LOW-LIGHT VISIBLE AND UNCOOLED THERMAL INFRARED IMAGERY

by

James W. Scrofani

June 1997

Advisor: Charles W. Therrien

Approved for public release; Distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June, 1997	Master's Thesis

4. TITLE AND SUBTITLE **AN ADAPTIVE METHOD FOR THE ENHANCED FUSION OF LOW-LIGHT VISIBLE AND UNCOOLED THERMAL INFRARED IMAGERY**

5. FUNDING NUMBERS

6. AUTHORS Scrofani, James, W.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT(*maximum 200 words*)

Night vision sensors, such as image-intensifier (II) tubes in night vision goggles and forward looking infrared sensors (FLIR) are routinely used by U.S. naval personnel for night operations. The quality of imagery from these devices however, can be extremely poor. Since these sensors exploit different regions of the electromagnetic spectrum, the information they provide is often complimentary, and therefore, improvements are possible with the enhancement and subsequent fusion of this information into a single presentation. Such processing can maximize scene content by incorporating information from both images as well as increase contrast and dynamic range. This thesis introduces a new algorithm, which produces such an enhanced/fused image. It performs adaptive enhancement of both the low-light visible (II) and thermal infrared imagery (IR) inputs, followed by a data fusion for combining the two images into a composite image. The methodology for visual testing of the algorithm for comparison of fused and original II and IR imagery is also presented and a discussion of the results is included. Tests confirmed that the fusion algorithm resulted in significant improvement over either single-band image.

14. SUBJECT TERMS Enhancement, Fusion, Imagery, Infrared, Low-light Visible, Night Vision, Sensor Fusion, FLIR, NVG, Peli-Lim Algorithm

15. NUMBER OF PAGES 102

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18 298-102

AN ADAPTIVE METHOD FOR THE ENHANCED FUSION OF LOW-LIGHT VISIBLE AND UNCOOLED THERMAL INFRARED IMAGERY

James William Scrofani
Lieutenant, United States Navy
B.S.Ch.E., University of Florida, 1987
M.B.A., Brenau University, 1994

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
June 1997

ABSTRACT

Night vision sensors, such as image-intensifier (II) tubes in night vision goggles and forward looking infrared sensors (FLIR) are routinely used by U.S. naval personnel for night operations. The quality of imagery from these devices however, can be extremely poor. Since these sensors exploit different regions of the electromagnetic spectrum, the information they provide is often complimentary, and therefore, improvements are possible with the enhancement and subsequent fusion of this information into a single presentation. Such processing can maximize scene content by incorporating information from both images as well as increase contrast and dynamic range. This thesis introduces a new algorithm, which produces such an enhanced/fused image. It performs adaptive enhancement of both the low-light visible (II) and thermal infrared imagery (IR) inputs, followed by a data fusion for combining the two images into a composite image. The methodology for visual testing of the algorithm for comparison of fused and original II and IR imagery is also presented and a discussion of the results is included. Tests confirmed that the fusion algorithm resulted in significant improvement over either single-band image.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Foremost, the author would like to acknowledge the Lord, Jesus Christ, for his gifts of reasoning and creativity that were required to complete such an endeavor.

Next, my wife, Lori, and children, Sydni and C.J., for their love, patience and sacrifice, and Professors Charles W. Therrien and William K. Krebs for their wisdom, guidance and steadfast support in the completion of this thesis.

Finally, the U.S. Army Night Vision and Electronic Sensors Directorate, Fort Belvoir, VA for use of their imagery in this thesis and, the financial support of the Office of Naval Research (ONR), Life Sciences Division, for allowing the purchase of the equipment used in this thesis. This work was performed under Contract N0001497WR30078.

I. BACKGROUND

A. INTRODUCTION

"Surprise is a vital ingredient in conducting successful warfare. As early as 500 B.C., the Chinese general Sun Tzu recognized this simple fact in his oft-quoted treatise on the art of war. Throughout history, commanders have employed the darkness of night to gain surprise and to grasp the initiative from the hands of the enemy." Despite the difficulties associated with conducting such operations, history has revealed that, " 'darkness is a double-edged weapon,' and like terrain, 'it favors the one who best uses it and hinders the one who does not.' " Furthermore, one former-Soviet general and historian has noted, " 'troops should be equally capable of operations both during the day and at night' and that night operations have an 'urgent significance in modern warfare [Ref. 9].' " This thesis seeks to improve the night operations capability of the military, by improvements in the imagery produced by current night vision sensors. In particular, we address the problem of how two sources of nighttime scenic information can be enhanced and combined to produce an image superior to either.

B. PURPOSE

Current night vision sensors, such as image intensifier (II) tubes in night vision goggles and forward looking infrared sensors (FLIR) are routinely used by U.S. naval personnel for night operations. The quality of imagery from these devices however, can be extremely poor, suffering from poor contrast, limited dynamic range, graininess and many other reported problems. These deficiencies often lead to confusion of textures, the inability to segment them and visual illusions, resulting in disorientation, aborted missions, and lost aircraft and personnel [Ref. 10, 11]. Since these sensors exploit different regions of the electromagnetic spectrum, the information they provide is often complimentary, and therefore, improvements are possible with

the enhancement and subsequent fusion of this information into a single presentation. Such processing can maximize scene content by incorporating information from both input images as well as increase contrast and dynamic range. This thesis introduces a new algorithm, which performs adaptive enhancement of both low-light visible (II) and thermal infrared imagery (IR) inputs, followed by a data fusion technique for combining the two images into a composite image. The goal is to develop an enhancement/fusion algorithm that consistently produces a final image that is superior to either of the original images, for a wide range of reflectivity and emissivity conditions. Utility of this development includes improvements in night piloting, both navigation and targeting, man overboard detection, firefighting, special forces operations as well as civilian night driving, law enforcement and assistance for the visually impaired.

C. HISTORY OF NIGHT VISION SYSTEMS

The high priority afforded to facilitate night operations has resulted in the military's development and deployment of various night vision devices (NVDs). Such devices are able to exploit the visible and infrared (IR) energy content of a nighttime scene, enhancing visibility and even producing information previously invisible to the naked eye. Of particular interest to this thesis, are NVD applications in USN/USMC aircraft; hence a brief history of significant developments follows [Ref. 1, 12].

1. NIGHTBIRD

In the late 1970s the United Kingdom (U.K.) Royal Air Force, in conjunction with contractor support at Royal Air Force Establishement, Farnborough, England, explored the initial concept of using NVDs to provide an inexpensive, passive navigation and attack system under night conditions. This program, which was called NIGHTBIRD, formed the basis for current USN/USMC Night System concepts. The program's intent was to demonstrate the feasibility of displaying imagery on a raster heads-up display (HUD) to enable low altitude night pilotage. The system initially

used image-intensified (II) imagery; however this was later replaced by FLIR imagery because of its ability to operate independently of scene illumination. Final developments incorporated the imagery of night vision goggles (NVG), a navigation FLIR (NAVFLIR) receiver projected onto a wide field of view (FOV) HUD, and a moving map display. Additionally, the project demonstrated the feasibility of sensor fusion, a concept in which the combination of data from different sensors is used to form a more complete view of the scene.

2. CHEAPNIGHT

CHEAPNIGHT was the first USN/USMC night vision program, initiated at Naval Warfare Center, China Lake, CA. in 1984. This program was implemented to test NIGHTBIRD technology and assess its applicability to USN/USMC aviation platforms. The U.K. systems evaluated included a raster HUD, a pod-mounted NAVFLIR, second and third generation NVGs, and a moving map. Operational testing resulted in poof-of-concept of a passive Night Attack system.

3. QUICKNIGHT

The CHEAPNIGHT program was followed by QUICKNIGHT at the Naval Air Test Center in Patuxent River, MD. This program examined the feasibility of performing a quick install of third generation NVGs into the A-6E to give the platform passive Night Attack capability. The program further evaluated both the CATS EYES and Aviator's Night Vision Imaging System (ANVIS) NVGs. Additionally, low altitude comfort levels in low light conditions were assessed and the A-6E's targeting FLIR was tested as a NAVFLIR. The program concluded that the use of NVGs could give the A-6E limited passive Night Attack capability, and full capability could be achieved with a wide FOV NAVFLIR and HUD combined with NVGs and a moving map. Testing also resulted in the selection of CATS EYES as the NVG of choice for A-6E and other fixed-wing aircraft.

4. FLEETNIGHT

Following QUICKNIGHT was FLEETNIGHT, a fleet evaluation conducted in 1986 which included both east and west coast A-6E and F/A-18 squadrons. Selected aircraft were modified with NVG compatible lighting and several crews were trained to conduct night operations using CAT EYES NVGs. The results of this evaluation supported the concept of NVGs and showed that they were very effective as a passive complementary sensor to the radar in navigation and targeting applications.

5. REALNIGHT

The REALNIGHT program (1986-87) was developed to continue examination of the full Night Attack concept and the uses of its various components. These evaluations were performed at Naval Air Test Center in Patuxent River, MD, using an A-6E test bed equipped with wide FOV HUDs, CATS EYES NVG, NAVFLIR, touch screen displays, and a digital color map unit. The program tested operational and integration issues, and further explored the concept of sensor fusion.

6. AV-8B Night System

The AV-8B Night System program began flight testing in 1987 and was completed in July of 1988. Testing results showed that the Night System, which includes the NAVFLIR, CATS EYES and a moving map, gave the AV-8B an enhanced and effective low-level capability under night visual conditions.

7. F/A-18 Night System

Flight testing began on the F/A-18 Night System upgrade in 1989 and was completed in 1991. Testing indicated that the Night System, which includes the NAVFLIR, CATS EYES and digital moving map, enhanced F/A-18s night vision combat effectiveness; therefore, all subsequent F/A-18s have been equipped with night vision capability.

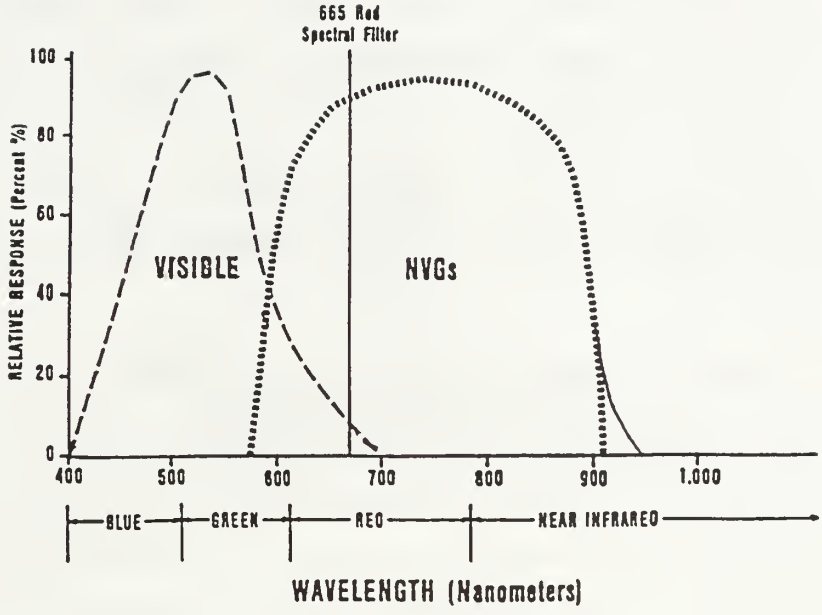

Figure 1. Spectral Response of a Typical Night Vision Goggle System [Ref. 1]

D. DISCUSSION OF SENSORS

1. Night Vision Goggles

Night vision goggles (NVGs) are passive image intensifiers which operate in the red and near-IR regions of the electromagnetic spectrum (Figure 1). The image intensifier tube produces a bright monochromatic (green) electro-optical image of a scene in which light level is too low for normal human vision. A typical NVG assembly consists of an objective lens, photocathode, microchannel plate, phosphor screen and combiner eyepiece assembly (Figure 2).

The image produced in a NVG assembly is based on the amount of light present in the scene, or the illuminance, and the amount of light reflected from objects in that scene, the luminance. This reflected light enters the goggles and is focused by

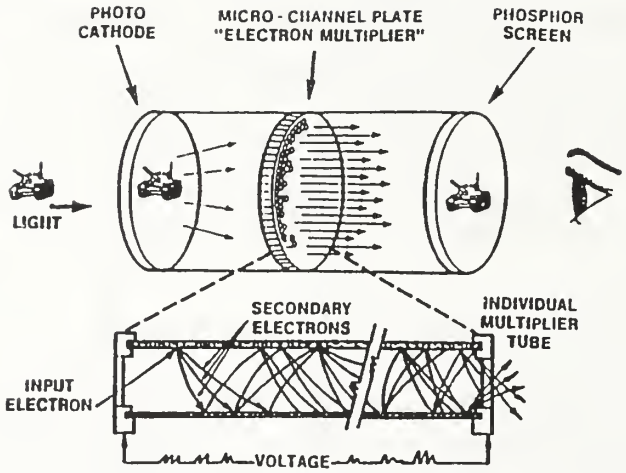

Figure 2. A Typical Night Vision Goggle Assembly [Ref. 1]

the objective lens onto the photocathode. The photocathode, which is responsive to both visible and IR radiation, converts the incident light to electrical energy.

Photons of light striking the photocathode cause a release of electrons in an amount proportional to the light incident on the photocathode. The released electrons are then accelerated away from the photocathode surface by an externally applied electric field. These accelerated electrons are then channeled through the microchannel plate, a very thin wafer of tiny glass tubes coated with a material that promotes secondary electron emissions. For each electron entering the microchannel plate, 1000 or more exit and are accelerated toward a phosphor screen. Incident electrons on the phosphor screen result in light emission.

The phosphor used in Generation 2 and 3 image intensifier tubes is referred to as P20 and emits a yellow-green light (560 nm) which matches the peak sensitivity of the photopic (day) human eye. Additionally this phosphor has a fast decay time, appropriate for aviation applications where high speeds require fast visual updates.

In the final part of the NVG assembly a combiner lens (not shown) carries the intensified image from the phosphor screen to the eye. The combiner lens returns the image to its natural orientation, rescales the image 1:1 and correctly registers the image.

2. Forward-Looking Infrared (FLIR) Sensors

A forward-looking infrared sensor (FLIR) is a device which detects the self-radiating and reflected infrared (IR) energy from objects in a scene and converts this energy into a visible presentation. Generally, IR energy is generated from the heating of an object, which increases molecular vibrational energy, causing an increase in molecular energy state. The subsequent return to a normal energy state results in the emission of IR radiation. The spectral distribution of IR energy is depicted in Figure 3.

The ability of a material to emit IR energy compared to a blackbody at the same temperature is known as its *emissivity* and indicates to what degree this heating will result in the emission of IR radiation. Emissivity is a function of both the type and surface finish of the material. Table I lists values for some common materials.

IR sources can be classified as either thermal or selective radiators. Thermal radiators output a wide spectrum of energy with a maximum radiant energy at some particular frequency, while selective radiators release energy concentrated about a narrow band of frequencies (e.g., laser emission). Examples of thermal radiators include, the sun, the hot metal of a jet engine tail pipe, aerodynamically heated surfaces, motorized vehicles, human personnel, and terrain. Most objects of interest are thermal radiators and emit maximum radiant energy in the 8-12 micron range, corresponding to the detection range of many IR devices.

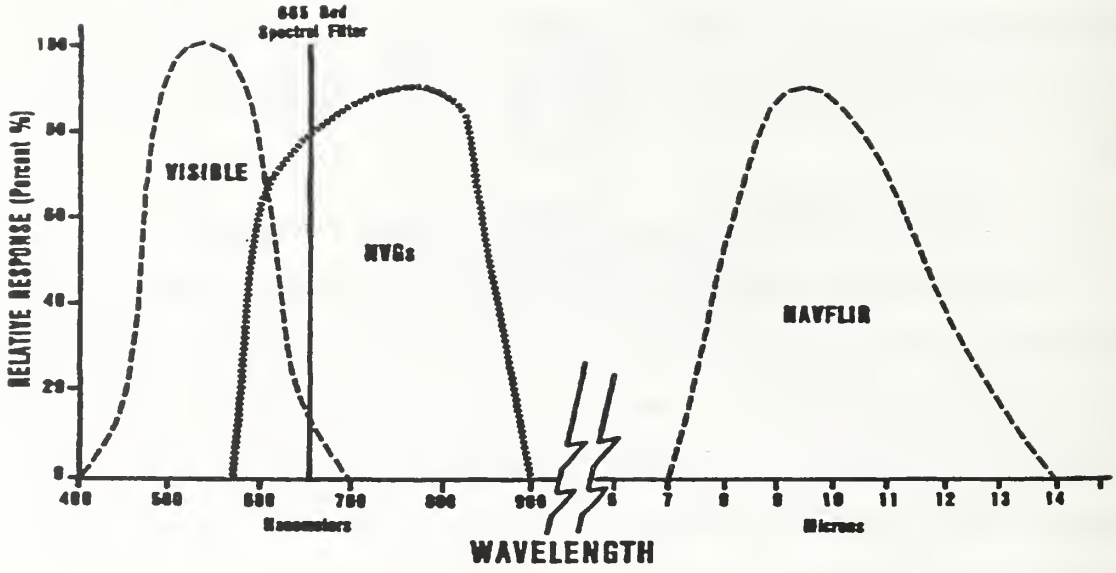

Figure 3. Spectral Distribution of Infrared Radiation [Ref. 1]

A typical navigational FLIR (NAVFLIR) device is comprised of two major systems; the sensor and the cockpit display systems. The sensor system is of most interest here and is described below. This system is comprised of the sensor head/sensor unit which includes the IR window, IR telescope, scanning assembly, detector array and cooling system (see Figure 4).

material	emissivity
highly polished silver	0.02
highly polished aluminum	0.08
polished copper	0.15
aluminum paint	0.55
polished brass	0.60
oxidized steel	0.70
bronze paint	0.80
gypsum	0.90
rough red brick	0.93
white lacquer	0.95
green or gray paint	0.95
lamp black	0.95
water	0.96

Table I. Emissivity of Some Common Materials [Ref. 1]

Electromagnetic energy incident to the FLIR interacts first with the IR window. The IR window preferentially passes IR energy in the 8-12 micron range to the IR telescope. Since glass completely absorbs radiation in this band, the window is composed of germanium with a high efficiency carbon coating for durability.

The IR telescope, which is located directly behind the IR window, focuses the thermal energy onto the motor drive scanning assembly. The magnification level of the telescope is selected to match the heads-up display field of view so that 1:1 registration with the true scene is achieved. The telescopic lenses are also made from germanium.

The IR telescope transmits the thermal energy to the scanning assembly by a series of mirrors and lenses. The scanning assembly consists of a motor driven scanner which opto-mechanically scans the detector array across the thermal scene. A 2:1, 60 Hz interlaced scanning process is used resulting in a refresh rate (30 Hz) and field of view (FOV) suitable for standard commercial TV 525 line format.

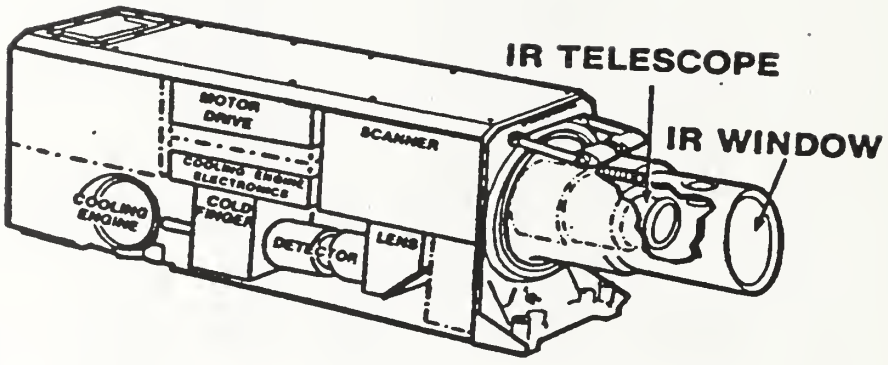

Figure 4. NAVFLIR Sensor Head/Sensor Unit [Ref. 1]

The detector array is a quantum detector, composed of mercury-cadmium-telluride. This material is sensitive to radiation wavelengths in the 8-12 micron range. When subject to radiation of this wavelength, a dramatic increase in conductivity and hence an increase in electrical current occurs in the material. The detector thus converts incident thermal energy into a proportional electrical signal.

Due to the inherent thermal energy of the detector a cryogenic cooling system is required to reduce the electrical current generated by ambient conditions. To adequately minimize this current, the detector is maintained at a constant temperature of approximately -193 degrees C.

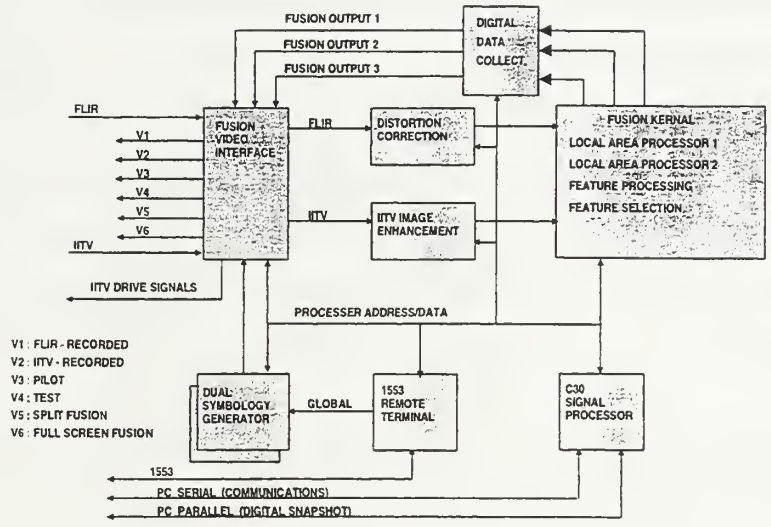

Figure 5. Army/TI Image Fusion Testbed Block Diagram [Ref. 2]

E. RELATED WORK

1. Image Fusion Program

The Army in conjunction with Texas Instruments Corporation (TI) has an ongoing program investigating the night pilotage benefits of fusing imagery from infrared and image intensified sensors [Ref. 2, 13]. Known as the Aviation Applied Technology Directorate's (AATD) Image Fusion Program, this program has recently demonstrated significant night pilotage benefits from the fusion of FLIR and image-intensified (II) imagery. Evaluation pilots have demonstrated overwhelming preference for the fused output.

The impetus for these studies was based on several Government and industry studies that evaluated the relative merits of image intensified and thermal imagery as they pertained to helicopter night pilotage [Ref. 2, 13]. The results of these studies revealed that each of the sensors performed optimally under different conditions and environments and that most pilots preferred to have both sensors available. The stud-

ies further indicated that a complementary relationship existed between the sensors and that an ideal pilotage system would incorporate both sensors.

Based on these studies TI developed a system to dynamically combine imagery from both sensors and display them in a single presentation. A series of unsolicited proposals and demonstration flights then resulted in the Army developing the TI Image Fusion Program. Reference [Ref. 2] includes a full discussion of the program; however, a basic system discussion and overview of the proprietary image fusion process is excerpted below and a system block diagram is depicted in Figure 5.

> The primary goal of the Image Fusion processing is to provide the highest quality scene information at each pixel in the resultant fused image. In order to accomplish this goal, it is necessary for the processing to be tightly coupled with the individual sensors. As mentioned in the previous sections, care is taken to register, optimize and normalize the individual sensor videos prior to the fusion process. Resultant sensor signal to noise and other image quality metrics are estimated as a function of individual sensor gains and post processing statistics.
>
> The fusion Kernel function, which performs the core Image Fusion algorithms, receives distortion corrected FLIR video and enhanced II video. It separates each video signal into components, based on local area criteria. The Fusion Kernel further processes these components from each sensor using a process which preserves maximum detail in the resultant fused images. The Fusion Video Interface board combines the resultant fused digital video with symbology and converts it to RS-170 composite video.

Apparently, after careful preprocessing (registration, enhancement, noise filtering), a decomposition of each image (II and IR) is performed. The resultant components are subject to processing (fusion kernel) to preserve maximum detail in the fused image.

2. Biological Models

References [Ref. 14, 3] propose a method to combine low-light visible and thermal IR imagery, which provides a true color night vision capability. This method is based on "biological models of color vision and visible-IR fusion." The color vision model attempts to model the color processing observed in the retina of humans and monkeys, while the visible-IR fusion method is developed from study of the fusion of

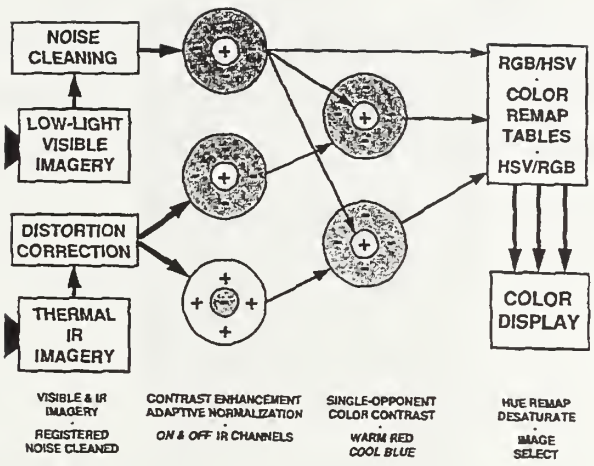

Figure 6. Neurocomputational Model of Proposed Color Night Vision System [Ref. 3]

thermal and visible imagery observed in rattlesnakes and pythons. Both biological models are incorporated into a set of neurodynamic equations which are solved using what is known as a feedforward center-surround shunting neural network. The development of the neurodynamic theory is complex and is discussed in the cited references.

Figure 6 shows a block diagram of the color night vision system. Inputs to the system are low-level light (II) from a Gen III intensified charge-coupled device (CCD) (0.6-0.9 μm) and long-wave infrared (LWIR) imagery from a Texas Instruments thermal imager (7.5-13 μm). Nearly registered images are produced by each sensor. The II imagery is median filtered for noise removal and registration distor-

tion present in the LWIR imagery is removed via distortion correction computations. Center-surround shunting neural networks are used first within bands for contrast enhancement and normalization as well as for developing ON and OFF channels of IR. Additional center-surround networks are then used between bands to create single-opponent color-contrast (gray-scale fused) images. Following this, the images are sharpened (not pictured) to create two double-opponent color-contrast images. The double-opponent images and enhanced visible images are then mapped to the red-green-blue (RGB) color domain for display, or remapped to the hue-saturation-value (HSV) domain for a tailored display, i.e., more natural color scheme.

3. NRL Color Fusion

The Naval Research Laboratory, in conjunction with the Naval Postgraduate School, is conducting research in the color fusion of imagery from image-intensified charge-coupled devices (IICCD) and infrared (IR) sensors. The program, termed the NITE Hawk project, has the objective of providing dual-band color night vision, by using II and IR sensors integrated onto an aircraft pod, and outputing the resultant fused imagery to multifunctional, navigation or helmet-mounted displays.

The sensor suite integrates an IICD into the gimbal assembly of the sensor head of a Lockheed-Martin NITE Hawk IR pod [Ref. 4], simultaneously providing II and IR imagery. This pod is manufactured for use on F/A-18 Hornet aircraft and is pictured in Figure 7.

The fusion of the two sensors is achieved through an adaptive statistical processing algorithm described in Figure 8. Band 1 (L_1) and Band 2 (L_2) information, II and IR pixel intensities, are statistically decomposed into orthogonal components L_1' and L_2'. The principal component direction L_1' represents high correlation between bands; i.e., both the II and IR images have similar intensities at a given pixel location, and are represented by varying grayscale intensities. The orthogonal component L_2' accounts for uncorrelated pixel intensities; i.e., the II and IR images have different intensities at a given pixel location, and are represented by varying color opponent

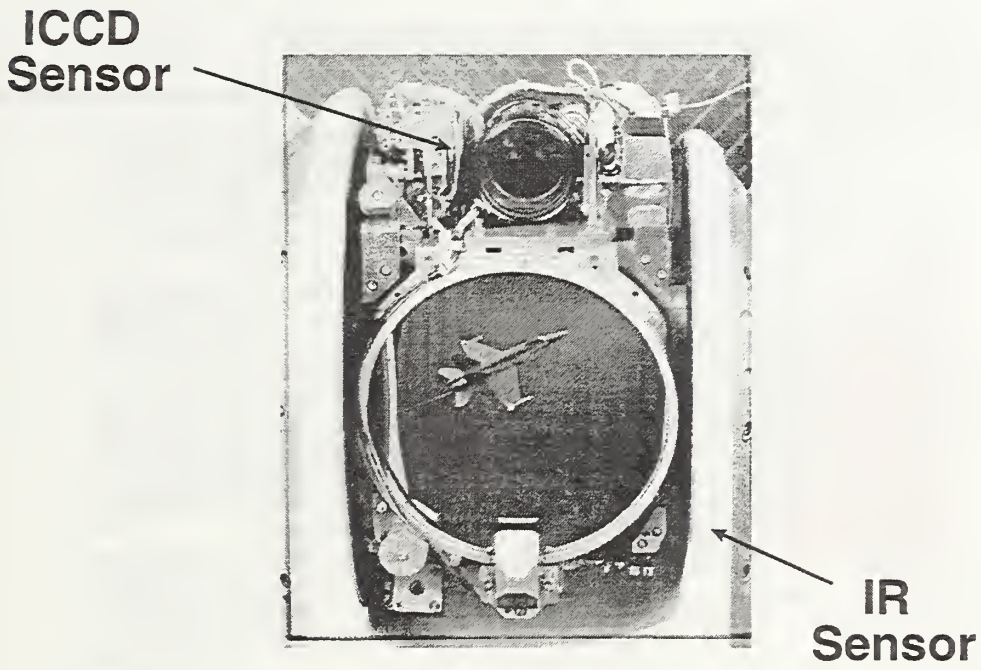

Figure 7. Integrated IICCD/IR Sensor Suite for F/A-18 Use [Ref. 4]

intensities (red/cyan). Therefore fused image pixel intensity is assigned based on the proximity of II and IR pixel intensity pairs to the principal component axis. II-IR pairs that are close to L_1' will result in a grayscale pixel intensity, while those that are distant will result in some combination of the color opponent intensities. Figure 9 shows the lookup table or colormap that assigns pixel intensity pairs to fused intensity output.

Figure 10 shows an example of this color fusion on an II-IR image-pair. Notice that features exclusive to a single band are represented by red or cyan while features prevalent in both are represented by varying grayscale intensity.

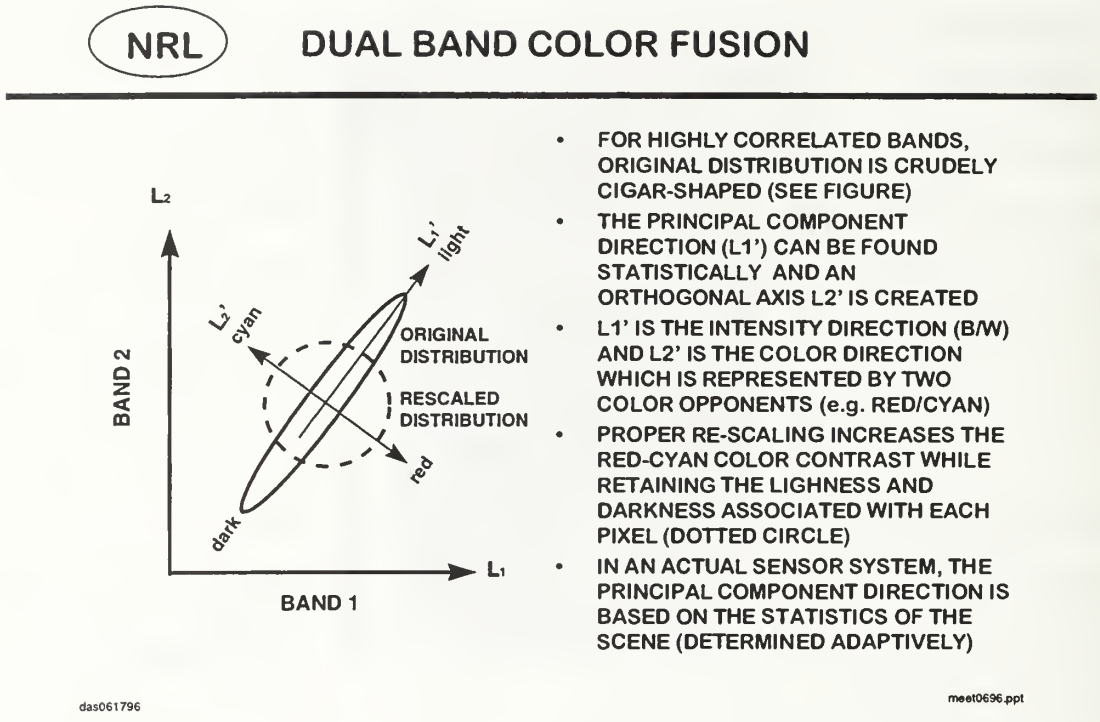

Figure 8. NRL Adaptive Statistical Processing Algorithm for Dual-Band Color Fusion [Ref. 5]

4. Wavelet-Based Fusion

Reference [Ref. 6] introduces a multi-sensor fusion technique based on the wavelet transform. This algorithm, which is depicted in Figure 11, performs pixel-level fusion on the input images and produces an output image that affords improved human visual perception of a given scene.

In the first stage of the algorithm, the wavelet transform is computed for each of the input images. The images are decomposed into low-high, high-low and high-high bands at different scales, where the transform coefficients with largest absolute value generally correspond to sharper brightness changes (and thus the "salient features" in the images). To extract the dominant features at each scale, at the next stage, a

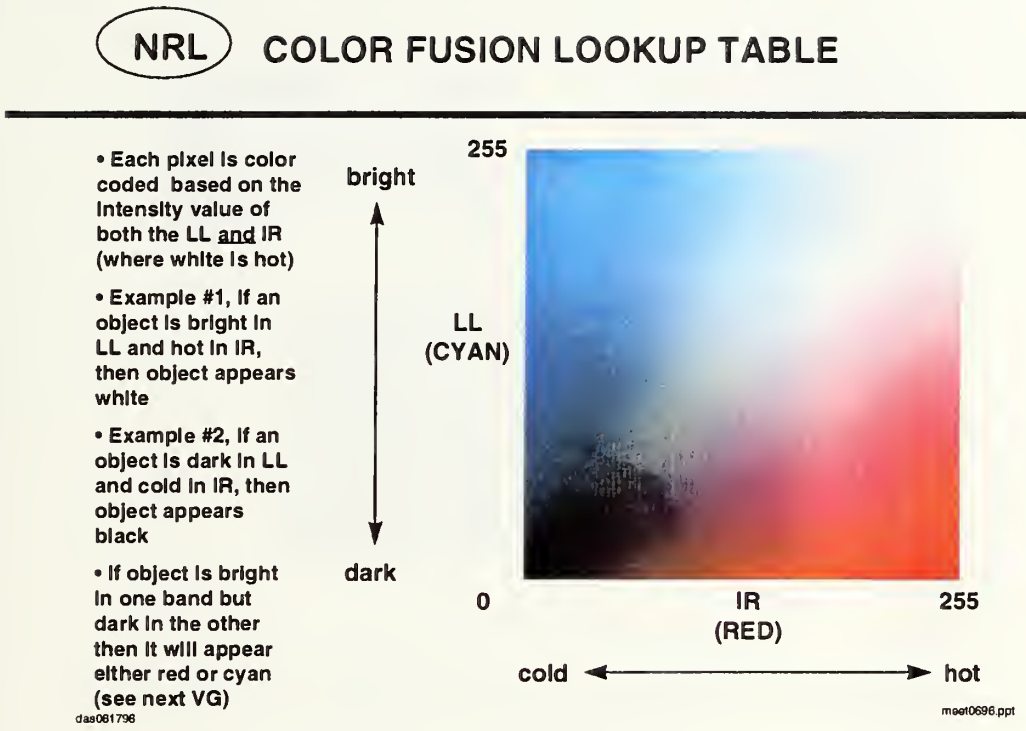

Figure 9. NRL Dual-Band Color Fusion Lookup Table [Ref. 5]

comparison of wavelet coefficients is performed and the one with the higher absolute value is retained. Finally, the inverse wavelet transform is computed on the retained coefficients, and the output image is produced.

This algorithm does not suffer from the contrast reduction prevalent with the direct fusion methods nor the blocking artifacts often associated with Laplacian pyramid based fusion techniques.

F. OUTLINE OF THESIS

This thesis presents a monochrome enhancement/fusion algorithm. This algorithm seeks to maximize the "scene content" of an enhanced/fused output image given low-light visible (II) and infrared (IR) input images. The remainder of this

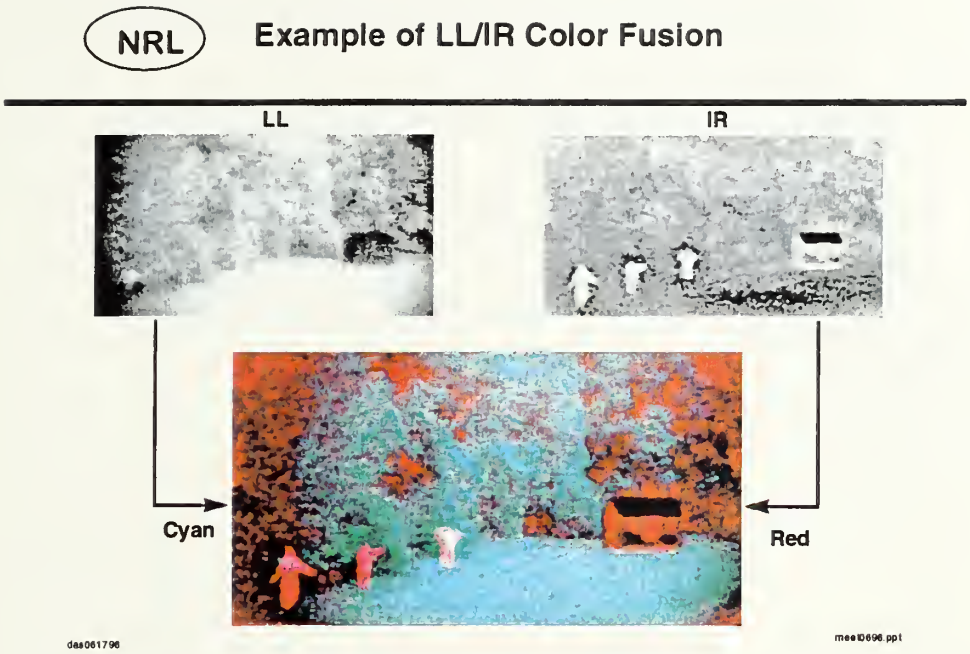

Figure 10. NRL Dual-Band Color Fusion Example [Ref. 5]

thesis is organized as follows. Chapter II describes the enhancement/fusion algorithm. This chapter begins with a discussion of the Peli-Lim algorithm [Ref. 7, 8], which is the basis for the enhancement/fusion in our algorithm. The chapter then presents a general discussion of the enhancement/fusion algorithm, followed by a detailed description of the algorithm, in the context of application to a particular image-pair. The final results of application to an image-pair are then shown and discussed. Chapter III discusses the methodology, analysis, and results of visual testing with human subjects performed on the enhanced/fused output of the algorithm. Finally, Chapter IV provides conclusions and a discussion of findings.

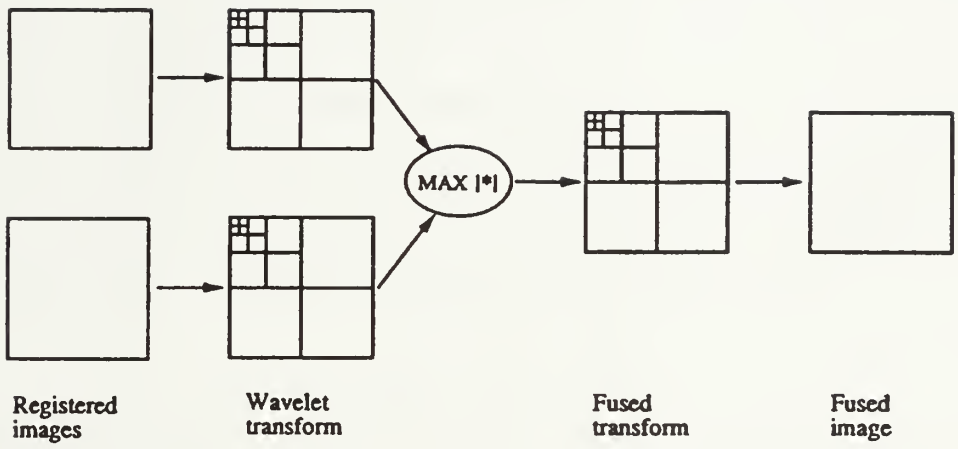

Figure 11. Block Diagram of the Multi-Sensor Wavelet Fusion Algorithm [Ref. 6]

II. ENHANCEMENT/FUSION ALGORITHM

A. PELI-LIM ALGORITHM

Traditional methods of image enhancement, such as histogram-based gray scale transformation or various filtering techniques, manipulate the entire image, i.e., they are spatially invariant. In certain applications, it is desirable to modify only particular regions of the image. The method described in [Ref. 7, 8], known as the Peli-Lim algorithm, allows variation of local contrast and local luminance mean as local characteristics of the image vary. As an example, if the strategy is to bring out detail in dark regions of an image, the algorithm affords this capability without affecting regions of the image which are not dark.

A block diagram of the Peli-Lim algorithm is given in Figure 12. The unprocessed image is denoted by $f(n_1, n_2)$, where n_1 and n_2 represent pixel indices within the image. The local luminance mean, $f_L(n_1, n_2)$, is obtained by passing the original image through a simple low-pass FIR filter whose output is given by:

$$f_L(n_1, n_2) = \frac{1}{(2N_1+1)(2N_2+1)} \sum_{k=-N_1}^{N_1} \sum_{l=-N_2}^{N_2} f(n_1 - k, n_2 - l) \qquad \text{(II.1)}$$

The parameters N_1 and N_2 control the neighborhood of the averaging operation. With small values, averages are highly dependent on close neighbors while with larger values they incorporate the influence of distant neighbors and result in more blurring. The local contrast, denoted by $f_H(n_1, n_2)$, is obtained from removing the local luminance mean component from the original image

$$f_H(n_1, n_2) = f(n_1, n_2) - f_L(n_1, n_2) \qquad \text{(II.2)}$$

The resulting image f_H contains just high spatial frequency components.

In the Peli-Lim algorithm, the two components f_L and f_H are modified separately then recombined. The modification of these image components is based on the local luminance mean. For example, if the strategy is to bring out detail in dark regions of an image, dark regions are identified by observing where the local luminance

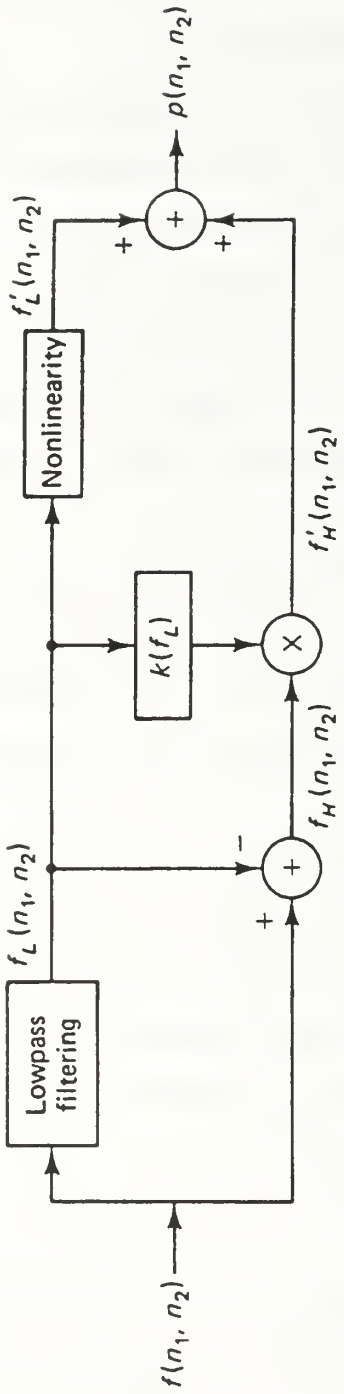

Figure 12. Block Diagram of the Peli-Lim Algorithm [Ref. 7]

mean has low values, and local contrast is increased in these regions. The modification occurs by scalar multiplication of the high-pass image by a gain factor $K(f_L)$, derived from the local luminance mean (see Figure 12). For $K(f_L) > 1$ contrast is increased within the image, while for $K(f_L) < 1$ contrast is decreased.

The local luminance mean, f_L , is modified by application of a generally non-linear transformation. This modification restores the dynamic range of the resultant image to that of the original. The final processed image is formed by the combination of the modified components,

$$p(n_1, n_2) = f'_L(n_1, n_2) + f'_H(n_1, n_2) \tag{II.3}$$

The Peli-Lim process has been used in several application with various gain and local luminance transformation curves such as those shown in Figures 13, 14, and 15. For example, in general enhancement of an image, it may be desirable to increase the local contrast of an entire image which may be of inferior visual quality due to under- or over-exposure during imaging. In this case, selection of $K(f_L)$, is independent of the local luminance mean. The enhancement is applied over the entire image. The local luminance mean is modified by a non-linearity chosen to restore appropriate dynamic range. Figure 13 depicts a set of gain and local luminance transformation curves which are suited to this application.

A second application is the enhancement of images degraded by cloud cover. Regions of an image covered by clouds exhibit an increase in local luminance mean and a decrease in local contrast, both of which vary with the amount of cloud cover present. In this case it is desirable to detect regions where local luminance mean is high and increase the local contrast in these regions. The local luminance mean is modified by a non-linearity, chosen as before, to restore the dynamic range. Figure 14 depicts a set of gain and local luminance transformation curves which are suited to this application.

A final example is the enhancement of images degraded by shadow regions. Regions of an image which are underexposed or have shaded regions exhibit decreased

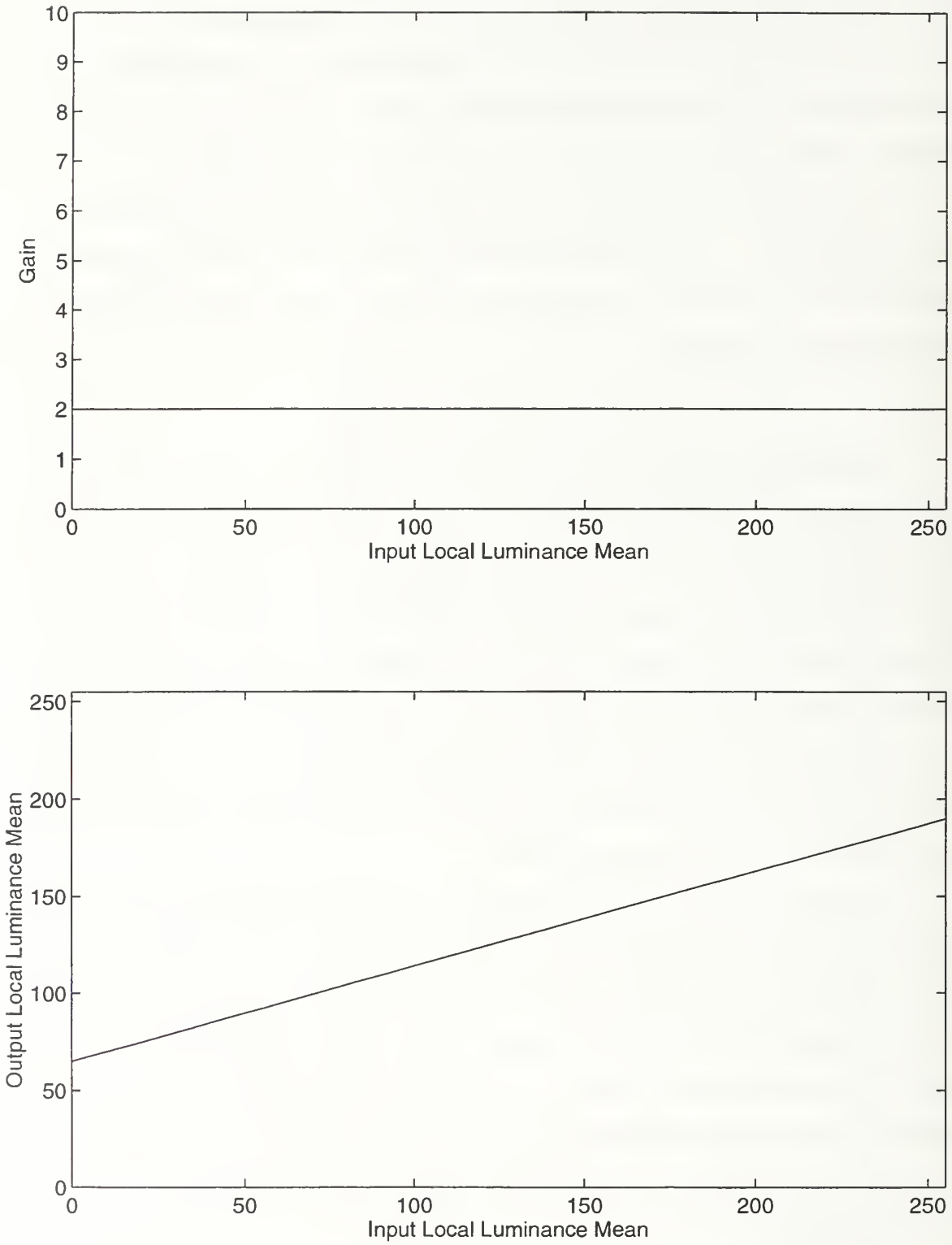

Figure 13. Gain (K) and Local Luminance Transformation (NL) Curves for Enhancing a General Image [Ref. 8]

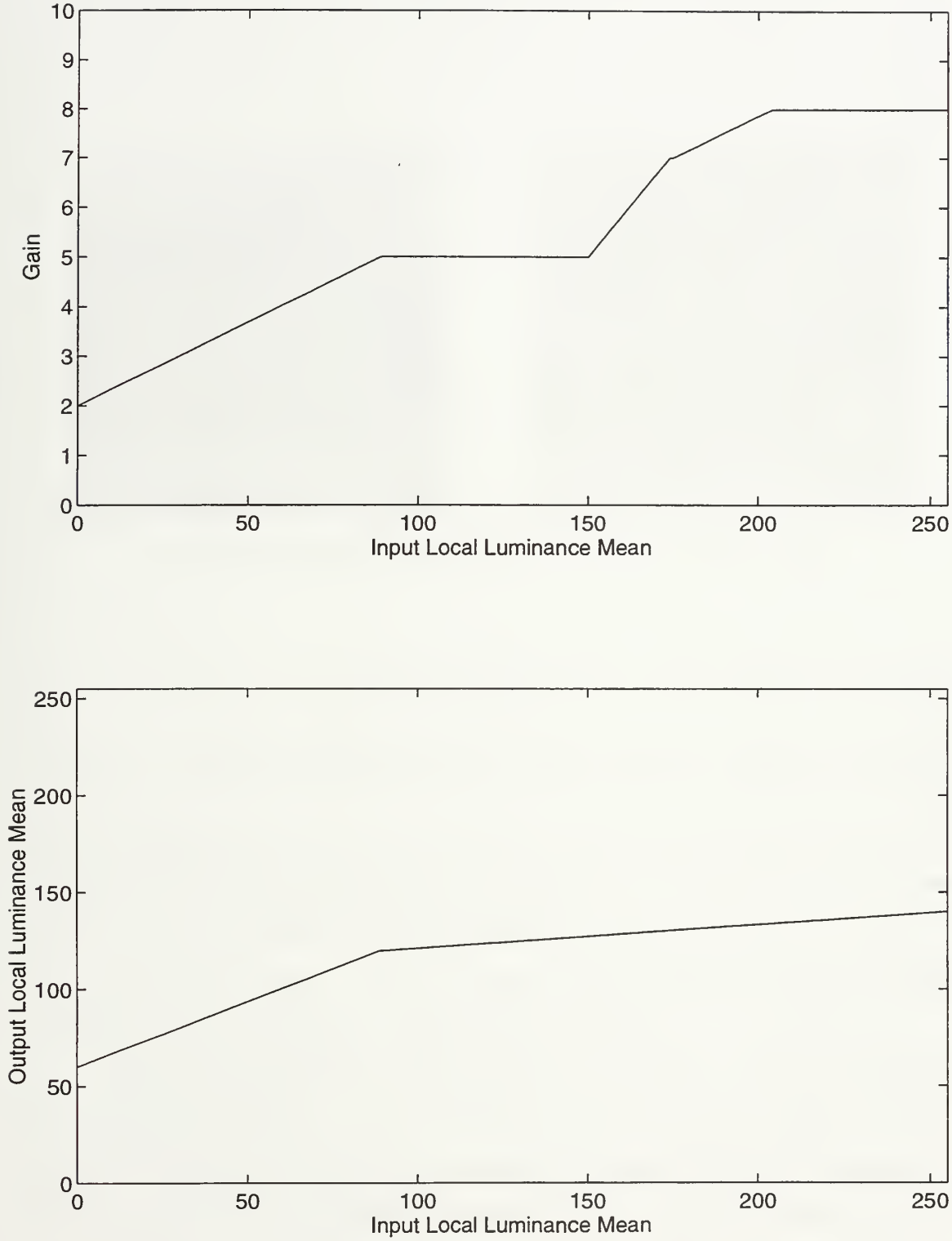

Figure 14. Gain (K) and Local Luminance Transformation (NL) Curves for Enhancement of an Image Degraded by Cloud Cover [Ref. 8]

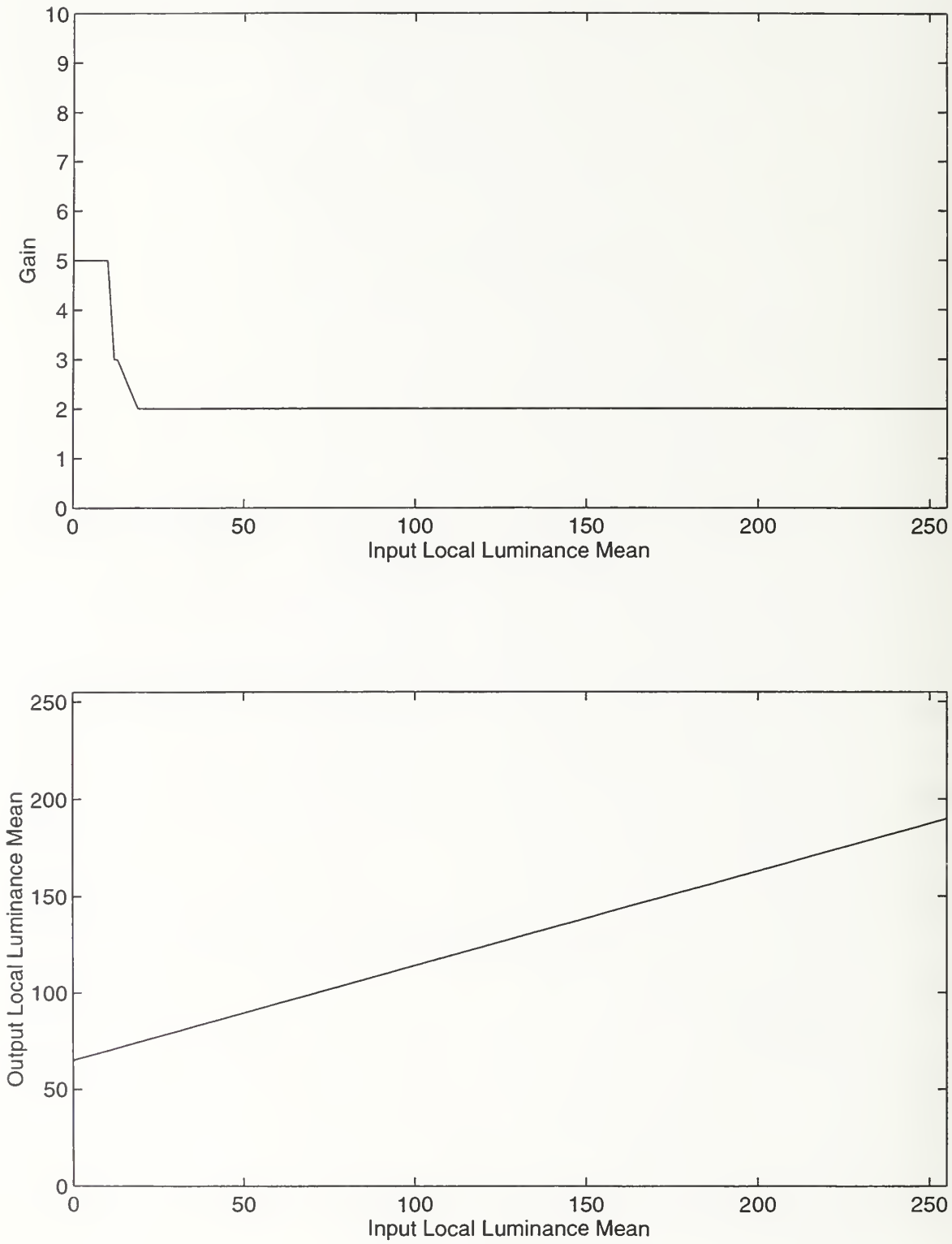

Figure 15. Gain (K) and Local Luminance Transformation (NL) Curves for Enhancement of an Image Degraded by Shadow Regions [Ref. 8]

Original Image

Enhanced Image

Figure 16. Original and Enhanced Image Using the Peli-Lim Algorithm [Ref. 8]

local luminance and decreased local contrast. In this case it is desirable to detect regions where local luminance mean is low and increase the local contrast in these regions. The local luminance mean is modified by a function, chosen as before, to restore the dynamic range. Figure 15 depicts a set of gain and local luminance transformation curves which are suited to this application.

Additional applications for this enhancement procedure include the enhancement of images degraded by varying amounts of smoke cover, fog or haze in different regions of an image and local luminance mean equalization for image segmentation [Ref. 8]. In each case suitable transformation curves area chosen to match the application.

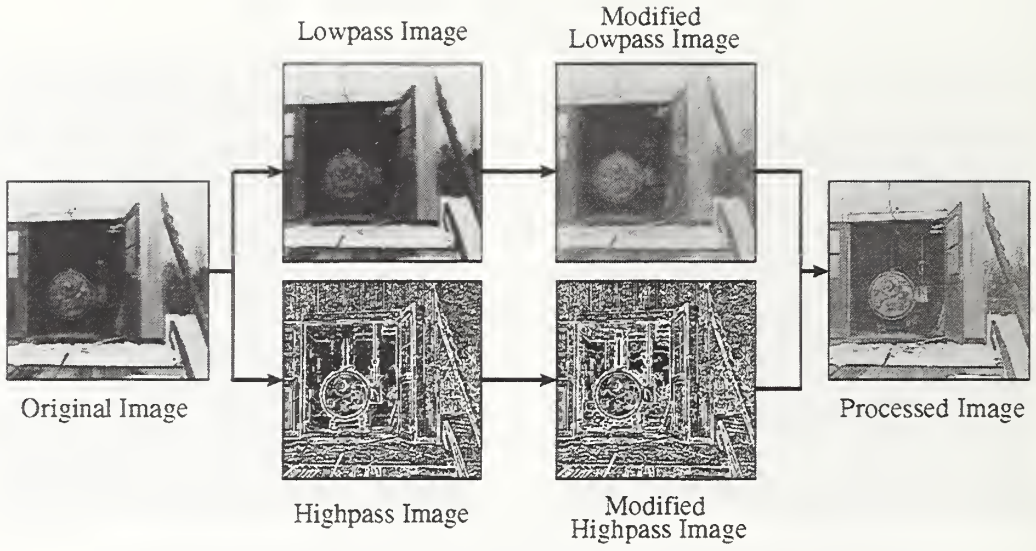

Figure 17. Image Depicting the Peli-Lim Decomposition and Enhancement Process

Figure 16 shows the results of this enhancement procedure on an image of a boiler room (size of 512 × 512 pixels) degraded by large shadow regions. The image is the same one used in [Ref. 7, 8] and the processing is similar to what is described there. The image on the left is the unprocessed image, the one on the right is the processed image. The parameters used for this processing were $N_1 = 3$, $N_2 = 3$, for the low-pass averaging and the curves used are those of Figure 15 . Note that the processed image reveals significant details in the previously shaded regions. Details in the back of the boiler room are visible as are some details within the trees outside. Figure 17 illustrates how the various image components are formed, modified, and recombined to achieve the final result.

The major advantages of this enhancement procedure are that it is spatially adaptive to particular regions within an image; it can be tailored to a particular class of images, e.g., images with shadow regions; and it is both conceptually and computationally simple. The algorithm is intuitive and requires only algebraic computation. Additionally, the algorithm is robust, in that for a particular class of images, the same gain and local luminance transformation curves work well, even with some variation within the class.

B. ENHANCEMENT/FUSION ALGORITHM

1. Overview of Algorithm

The algorithm proposed in this thesis, which produces a fused and enhanced monochrome image, seeks to maximize "scene content" in the output image. The goal is to incorporate all information available from each sensor and to optimally combine this information into a single image presentation. The algorithm first performs adaptive modification of the local contrast and local luminance mean for enhancement of both the low-light visible (II) and thermal infrared (IR) imagery; this is followed by a data fusion technique which compares corresponding local energies between the two images, then scales them based on their contribution to image detail. Finally, the modified image components are recombined to produced the enhanced/fused image. Figure 18 shows a block diagram of the algorithm.

The spatially adaptive enhancement and fusion is based on a modified version of the Peli-Lim algorithm. In this stage, the raw visible and IR data are each separated into spatial high- and low-pass components (f_H, f_L, g_H, g_L). For each of these data types, enhancement to the high-pass portion is achieved my multiplying by a gain factor that depends on the local luminance mean through the function $K_i(\cdot)$. The low-pass component is passed through a generally nonlinear luminance transformation $NL_i(\cdot)$ whose purpose is to reduce the dynamic range so that when this component is recombined with the enhanced high-pass component, saturation will not occur.

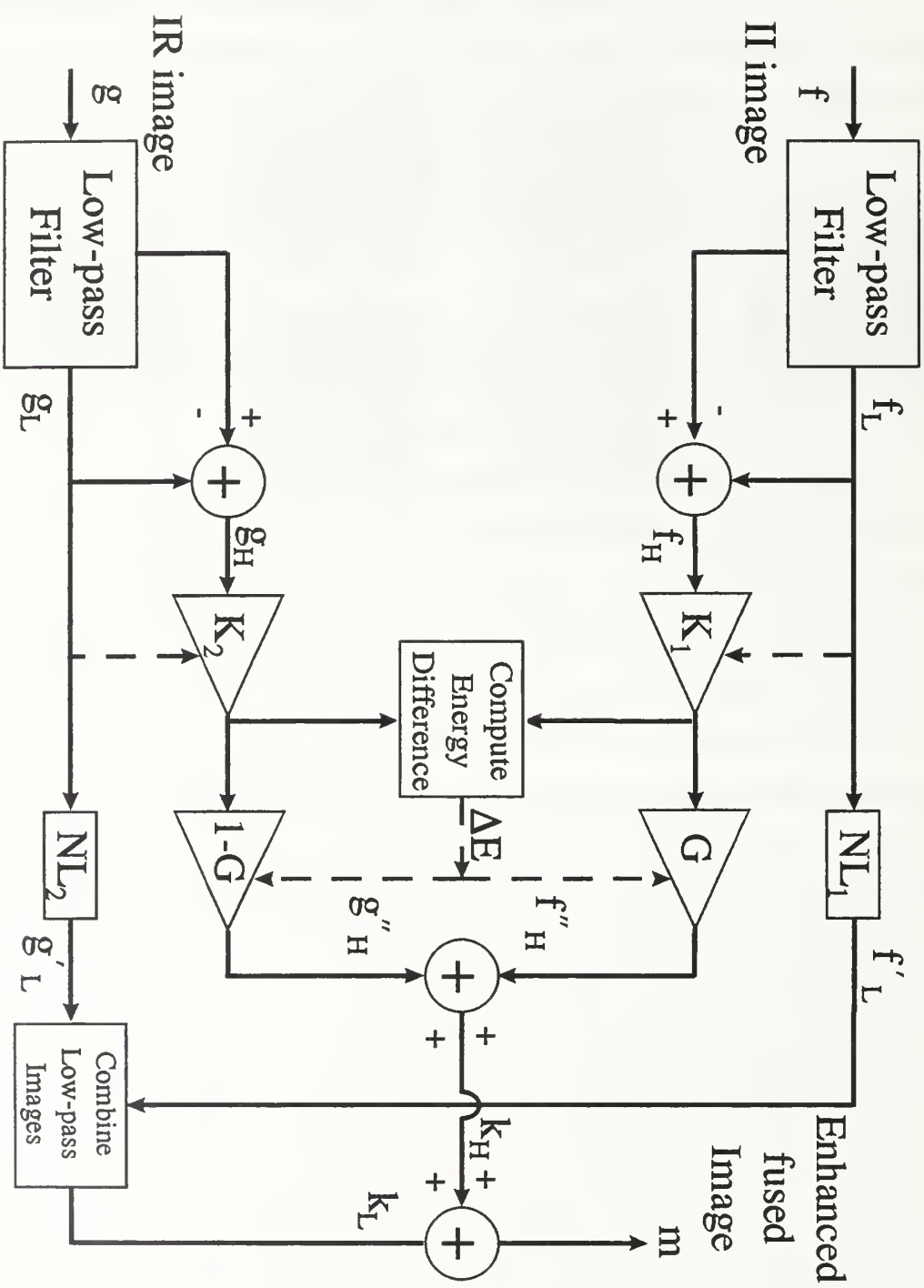

Figure 18. A Block Diagram of the Enhancement/Fusion Algorithm

Different functions K_i and NL_i are used for the low-light visible (II) and the IR data, with each function specifically tailored to address the enhancement problems pertinent to that type of data. Although the functions K_i and NL_i do not change, contrast and other enhancement is spatially adaptive, depending on the luminance characteristics in the local area.

For the next stage (fusion), the enhanced high-pass components of the two data types, which contain most of the detail, are compared through an energy computation, and a weighting G and $(1 - G)$ with $0 \leq G \leq 1$ is given to each of these components based on a normalized difference of local energies:

$$E_i(n_1, n_2) = \frac{1}{(2N_1 + 1)(2N_2 + 1)} \sum_{k=-N_1}^{N_1} \sum_{l=-N_2}^{N_2} f^2(n_1 - k, n_2 - l) \qquad \text{(II.4)}$$

The energy comparison and weighting insures that the high-pass component (II or IR) that contains the most detail will be most heavily weighted in the fusion.

In the final step, the four modified components (f'_L, f''_H, g'_L, g''_H) are recombined to produce the enhanced, fused image. The low-pass components are first combined to one composite image and the weighted high-pass components are added in. In combining the two low-pass components, both linear and nonlinear functions have been used to map intensity values in a two-dimensional space (II, IR) onto a one-dimensional space (fused intensity). For the latter, nonlinear optimization algorithms have been used to determine the mapping.

2. Mapping Considerations

Sammon [Ref. 15] describes a nonlinear mapping algorithm which preserves structure when mapping points from an L-dimensional subspace to one of a lower dimension. Structure is preserved by fitting the N points in the lower-dimensional subspace such that their intersample distances approximate, as closely as possible, the intersample distances of the N points in the L-dimensional subspace.

We are given N vectors in an L-space designated X_i, $i = 1, 2, 3, ..., N$ and corresponding to these we define a set of N vectors in a l-space (of lower dimension

l) designated Y_i, $i = 1, 2, 3, ..., N$. The distance between the vectors X_i and X_j in the L-space is defined as $d^*_{ij} \equiv dist[X_i, X_j]$ and the distance between the corresponding vectors in the l-space is defined as $d_{ij} \equiv dist[Y_i, Y_j]$, where any appropriate distance metric is used.[1] Now an initial l-space configuration is chosen either randomly or based on some *a priori* knowledge about the data and is denoted by:

$$Y_1 = \begin{bmatrix} y_{11} \\ \vdots \\ y_{1l} \end{bmatrix} \quad Y_2 = \begin{bmatrix} y_{21} \\ \vdots \\ y_{2l} \end{bmatrix} \quad \cdots \quad Y_N = \begin{bmatrix} y_{N1} \\ \vdots \\ y_{Nl} \end{bmatrix}$$

Based on this l-space configuration, intersample distances, d_{ij} are computed and compared to the original L-space distances, d^*_{ij}. The squared error $\mathcal{E}$ between the two distances represents a measure of how well structure is preserved between dimensions and is computed as follows:

$$\mathcal{E} = \frac{1}{\sum_{i<j}^{N} d^*_{ij}} \sum_{i<j}^{N} \frac{[d^*_{ij} - d_{ij}]^2}{d^*_{ij}} \tag{II.5}$$

By minimizing $\mathcal{E}$ using an appropriate optimization routine, the l-space configuration which best preserves intersample distance, and hence structure, can be determined.

3. Algorithm Details

The following discussion provides details of the algorithm proposed in this thesis. All elements of the algorithm, which include, enhancement, combination of low-pass components, and complete image fusion are presented. Five strategies to

---

[1] A distance metric function $dist[X_i, X_j]$ satisfies the following conditions, [Ref. 16]

$$dist[X_i, X_j] = \begin{cases} > 0 \\ = 0 \end{cases}$$

$$dist[X_i, X_j] = dist[X_j, X_i]$$

$$dist[X_i, X_j] + dist[X_j, X_k] \geq dist[X_i, X_k].$$

II Image IR Image

Figure 19. Unprocessed Image-Intensified and Infrared Images of Scene 1

achieve combination of low-pass components are offered, two linear and three nonlinear methods. To facilitate this discussion, the application is described with reference to a particular image-pair. This image-pair, referred to as Scene 1, is shown in Figure 19.

a. Enhancement

As previously discussed, the first stage of the algorithm (see Figure 18) results in the decomposition of each image into its high- and low-pass spectral components. For the image-pairs considered in this thesis, the low-pass averaging filter parameters used in Equation II.1 were taken to be $N_1 = 5, N_2 = 5$.

After decomposition, each image is modified by application of a set the functions K_i and NL_i. The selection of these functions is based on the particular

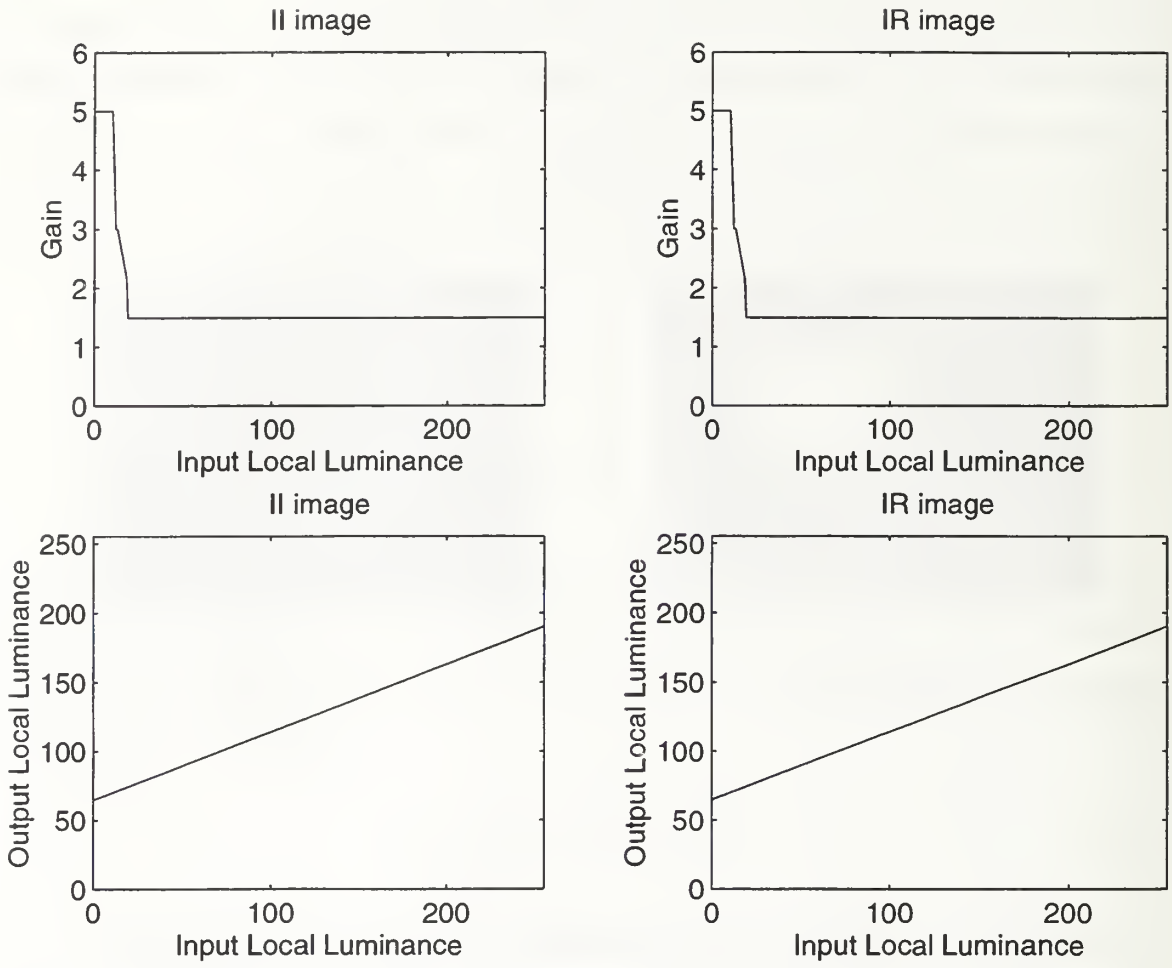

Figure 20. Transformation Curves (K and NL) Used in the Enhancement of Scene 1

luminance characteristics of the given image and the desired contrast enhancement. For example, for a given image-pair where both the II and IR images contain large shadow regions (as is often the case in our application), it is desirable to increase gain (K) in regions where the local luminance mean is low. Figure 20 shows selection of such a curve for the processing of Scene 1. Observe that the gain coefficient is chosen relatively large ($K = 5$) for low luminance intensities, and constant at $K = 1.5$ for other values of the luminance intensity. This results in the two-fold effect of increasing contrast in shadow regions and enhancing overall contrast of the image.

Without compensation, upon recombination of the high and low-pass components, these modifications in contrast would result in intensity values in excess of the original image's dynamic range . This requires the application of a luminance transformation NL to the luminance image. Figure 20 also shows the luminance transformation curves used to restore the dynamic range in both the II and the IR images for Scene 1.

b. Fusion of High-pass Components

The second stage of the algorithm involves the fusion of the enhanced high- and low-pass components. At this point in the processing, the energy associated with each high-pass image is compared and the pixel intensities associated with each are scaled accordingly. Since the high-pass images contain the "details" in the images, the image (II or IR) where the most detail is present is weighted most heavily in the processed image. The difference in energies, calculated from Equation II.4, is computed and normalized to obtain

$$\Delta E(n_1, n_2) = \frac{E_1(n_1, n_2) - E_2(n_1, n_2)}{\mid E_1(n_1, n_2) - E_2(n_1, n_2) \mid_{max}}, \tag{II.6}$$

where the maximum is over all pixels in the image. Thus the energy difference at any pixel $\Delta E(n_1, n_2)$ satisfies $-1 \leq \Delta E \leq 1$. This normalized energy difference, ΔE, is then used to compute the scaling factor from the function $G(\Delta E)$ shown in Figure 21. The function $G(\Delta E)$ could also be made nonlinear. Examples when a nonlinear relationship might be considered include cases when it is desirable to weight the energy contribution due to a particular sensor more heavily than the other or the situation when the range of ΔE is limited and an increase in dynamic range is desired.

If E_1 represents the energy associated with the II image and E_2 represents that of the IR image, a value of $\Delta E(n_1, n_2) = 1$ at a given pixel (n_1, n_2) indicates that the largest amount of energy associated with that particular pixel was due to the II image and the scaling factor is assigned the value $G(1) = 1.0$. Likewise $\Delta E(n_1, n_2) = -1$ indicates that the largest amount of energy associated with that

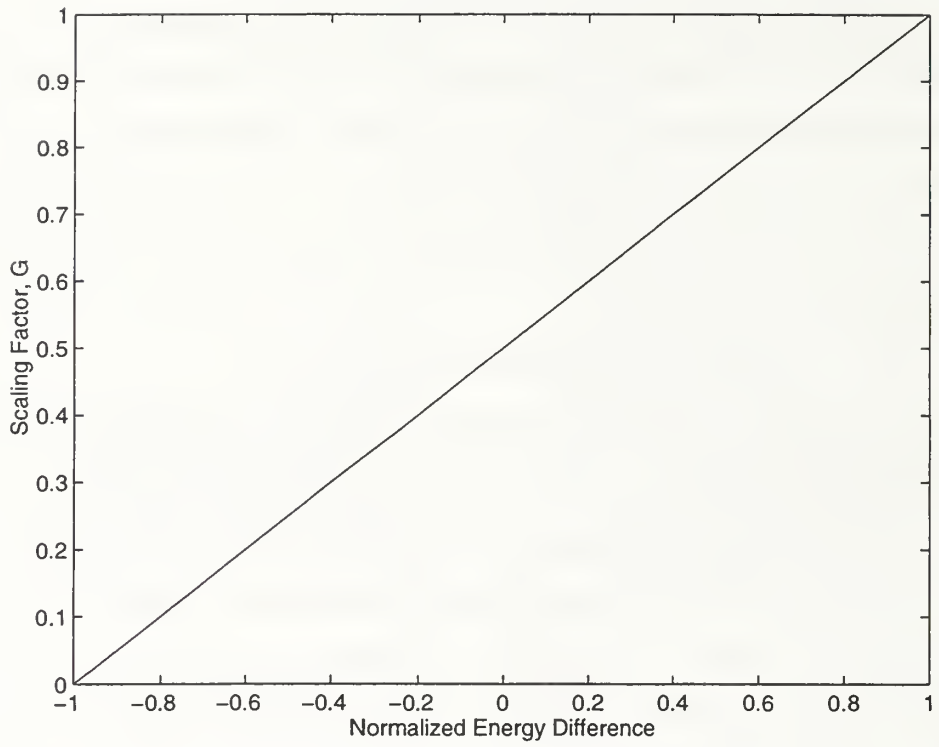

Figure 21. Function $G(\Delta E)$ Providing the Energy Scaling Factor as a Function of Normalized Energy Difference

particular pixel was due to the IR image and the scaling factor is assigned the value $G(-1) = 0$. For $\Delta E(n_1, n_2) = 0$, the energy associated with that particular pixel was equally distributed between the two images and the scaling factor is assigned the value $G(0) = 0.5$.

As shown in Figure 18, the contrast-enhanced high-pass II image is scaled by $G(\Delta E)$ and the IR image is scaled by $1 - G(\Delta E)$. The resultant images are combined to form the final high-pass image, that is

$$k_H(n_1, n_2) = f''_H(n_1, n_2) + g''_H(n_1, n_2) \tag{II.7}$$

where

$$f''_H(n_1, n_2) = f'_H(n_1, n_2)G(n_1, n_2) \tag{II.8}$$

and

$$g''_H(n_1, n_2) = g'_H(n_1, n_2)(1 - G(n_1, n_2)) \tag{II.9}$$

Method	Mapping Function
Direct Linear Mapping	$k_L(n_1, n_2) = [f'_L(n_1, n_2) + g'_L(n_1, n_2)]/2$
Weighted Linear Mapping	$k_L(n_1, n_2) = af'_L(n_1, n_2) + (1 - a)g'_L(n_1, n_2)$
Nonlinear–1 Mapping	$H(I, R) = (a_1 + a_2 I + a_3 I^2)(a_4 + a_5 R + a_6 R^2)$
Nonlinear–2 Mapping	$H(I, R) = (a_1 + a_2 I + a_3 I^2)(a_4 + a_5 R)$
Nonlinear–3 Mapping	$H(I, R) = (a_1 + a_2 I)(a_3 + a_4 R + a_5 R^2)$

Table II. Methods Used for Combining Modified Low-Pass Images

c. Combination of Low-pass Components

The modified low-pass images, f'_L and g'_L, are combined using a particular function that maps intensity values in a two-dimensional space (II, IR intensities) onto a one-dimensional space, the fused low-pass intensity domain. Five methods were tested in this thesis. Two of the methods use a linear mapping function to accomplish this, while the others use a nonlinear mapping function based on the Sammon mapping criterion previously discussed in this chapter. These five methods are listed in Table II and are discussed separately below.

(1) Direct Linear Mapping Algorithm. The first method, referred to as direct linear mapping, is simply the average of the two modified low-pass images:

$$k_L(n_1, n_2) = \frac{f'_L(n_1, n_2) + g'_L(n_1, n_2)}{2} \tag{II.10}$$

This method is the simplest to implement and requires no image-dependent parameters.

(2) Weighted Linear Mapping Algorithm. The second linear mapping, referred to as weighted linear mapping, is the linear combination of the two modified low-pass components:

$$k_L(n_1, n_2) = af'_L(n_1, n_2) + (1 - a)g'_L(n_1, n_2) \tag{II.11}$$

In applying this method the values for a evaluated were from $a = 0$ to $a = 1$ in steps of 0.1. Application of this mapping for Scene 1 resulted in choosing $a = 0.2$.

(3) Nonlinear Mapping Algorithms. The three remaining mapping methods are based on the nonlinear mapping theory discussed earlier. That is, we consider a pixel at a location (n_1, n_2) in the II image and assume that the II image intensity at this location is represented by the variable I. Similarly, the image intensity at the same location (n_1, n_2) in the IR image is represented by the variable R. These two values can be represented by a point in the (2-dimensional) I-R plane (see Figure 22). Then combining the images is equivalent to mapping this point in the plane to a one-dimensional space or line (again see Figure 22). The points in the I-R space are the X_i in the Sammon mapping discussion and the mapped values are the Y_i.

Polynomial forms with certain constraints on their coefficients, were chosen for the nonlinear mapping functions. Specifically, the three nonlinear mapping functions considered are:

$$H(I,R) = (a_1 + a_2 I + a_3 I^2)(a_4 + a_5 R + a_6 R^2) \tag{II.12}$$

$$H(I,R) = (a_1 + a_2 I + a_3 I^2)(a_4 + a_5 R) \tag{II.13}$$

$$H(I,R) = (a_1 + a_2 I)(a_3 + a_4 R + a_5 R^2), \tag{II.14}$$

where I and R represent pixel intensity in the modified low-pass II and IR images, respectively and $H(I,R)$ represents the combined low-pass pixel intensity (k_L). In this method the coefficients of the appropriate nonlinear function are chosen that best preserve intersample distances according to the Sammon mapping criterion II.5. Recall that the Sammon mapping criterion attempts to preserve intersample distances and thus preserve structure when mapping from a higher to a lower-dimensional space.

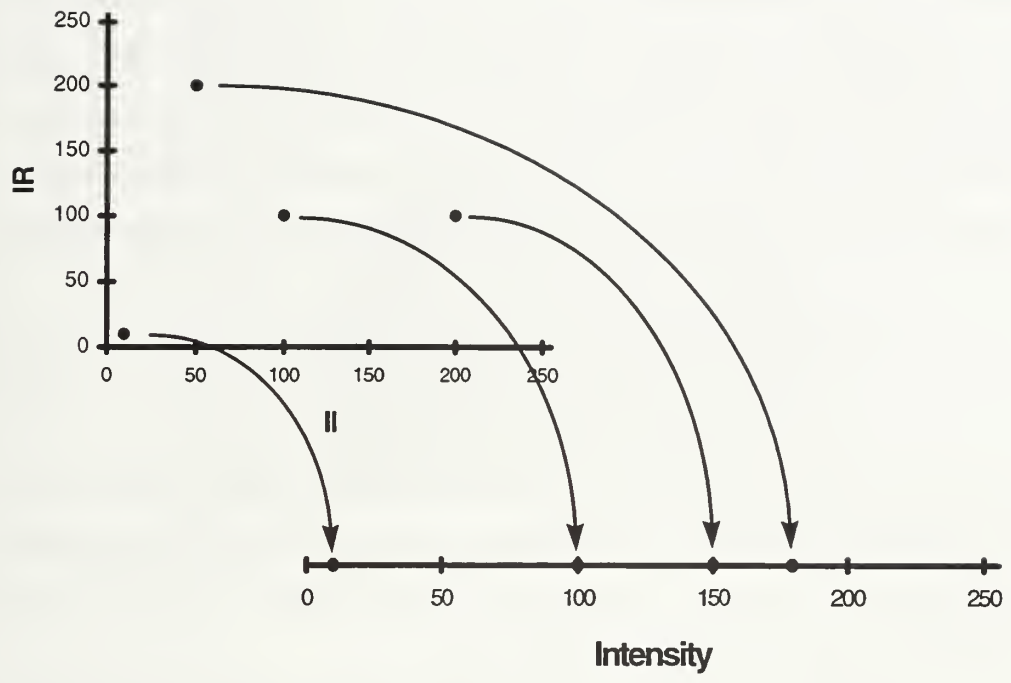

Figure 22. Mapping from the I-R Domain to the Fused Intensity Domain

To determine the coefficients we formulate the following constrained optimization problem:

$$\text{minimize } f(\underline{a}) \quad \text{subject to}: \begin{cases} H(0,0) = 0 \\ H(255,255) = 255 \\ \frac{\partial H(I,R)}{\partial I} \geq 0 \\ \frac{\partial H(I,R)}{\partial R} \geq 0 \end{cases} \tag{II.15}$$

where f(a) is,

$$f(\underline{a}) = \frac{1}{\sum_{i<j}^{N} d_{ij}^{*}(I,R)} \sum_{i<j}^{N} \frac{[d_{ij}^{*}(I,R) - d_{ij}(\underline{a})]^2}{d_{ij}^{*}(I,R)}. \tag{II.16}$$

and $d_{ij}(\underline{a})$ are the intersample distances in the fused intensity domain, which are implicitly a function of the parameter vector $\underline{a}$. The first two constraints are chosen to

preserve the dynamic range of the image, while the second two ensure monotonicity of the resultant nonlinear function. The monotonicity criterion guarantees a unique mapping from two- to one-dimensional space, i.e., no two (I,R) pairs can map to the same fused intensity point. Packaged optimization software [Ref. 17], based on the sequential quadratic programming method (SQP) was used to solve the constrained nonlinear optimization problem. In performing the optimization the Euclidean distance is used as the distance metric, that is $d^*_{ij}(I,R)$ and $d_{ij}(\underline{\mathbf{a}})$ are computed from:

$$d^*_{ij}(I,R) = [(I(i) - I(j))^2 + (R(i) - R(j))^2]^{\frac{1}{2}} \tag{II.17}$$

$$d_{ij}(\underline{\mathbf{a}}) = \mid H(I(i),R(i);\underline{\mathbf{a}}) - H(I(j),R(j);\underline{\mathbf{a}}) \mid \tag{II.18}$$

where $(I(i),R(i))$ represent the coordinates of the ith point in Figure 22 and we have indicated the explicit dependence of the mapping function H on the parameter $\underline{\mathbf{a}}$. Note that the d^*_{ij} are computed only once, prior to optimization, since the I and R image values do not change during the optimization.

Since the images evaluated have $640 \times 480 = 307,200$ pixels, there would be $\binom{307200}{2} = 4.7 \times 10^{10}$ intersample distances, and it is not computationally feasible to attempt the optimization using all of the points. Thus a suboptimal method is used to find the coefficients of the nonlinear mapping. The goal of this method is to represent the image data in the I-R plane (see Figure 23), by some smaller, yet representative set, that will make the optimization computationally feasible. This is achieved by using some representative I-R "centers" instead of the complete set of points in the I-R plane. A clustering algorithm known as the K-means algorithm [Ref. 16] was used to extract an appropriate number of centers representative of the structure of the original data, ($N_c = 25$ for all image-pairs considered in this thesis). Figure 23 shows the I-R plane and the resultant centers identified by the K-means algorithm for Scene 1.

In applying the clustering and optimization techniques different values of the parameter $\underline{\mathbf{a}}$ would be obtained for each pair of images. Table III shows

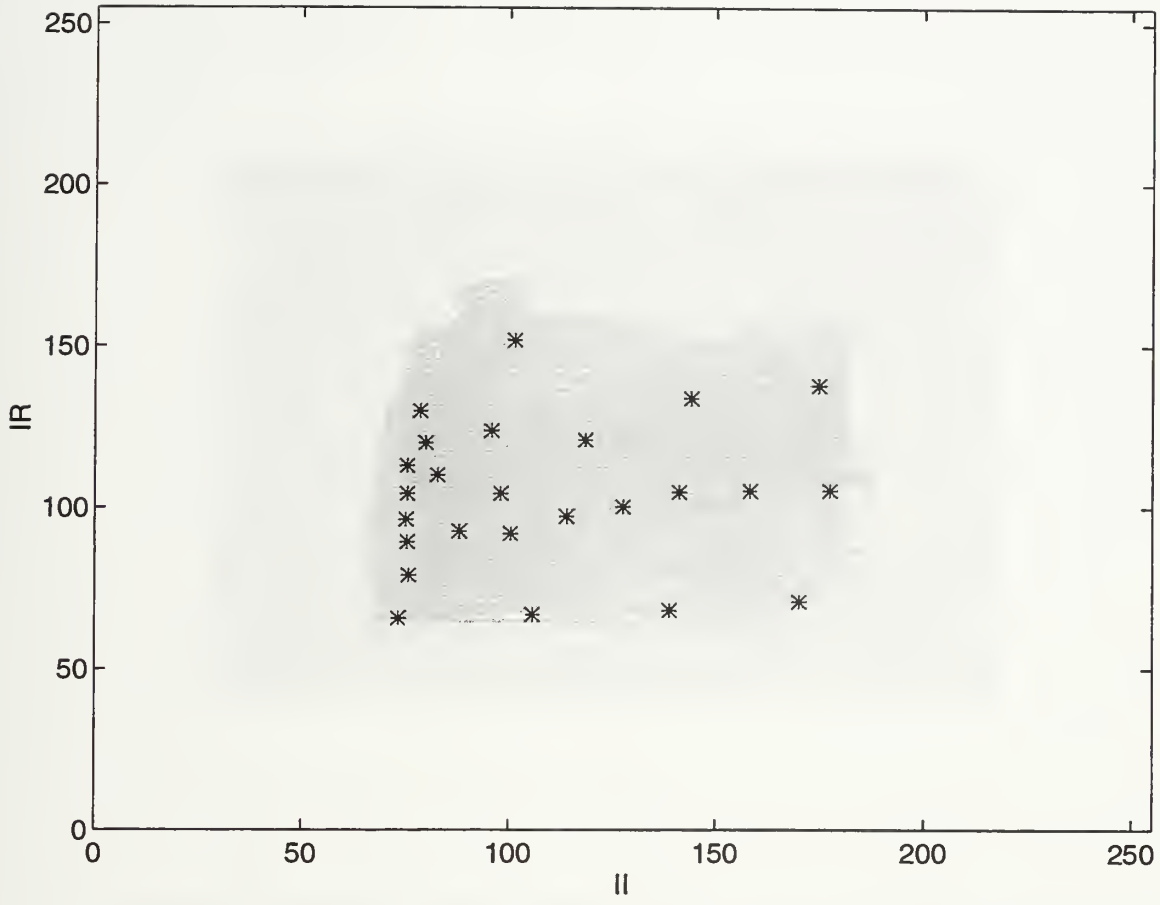

Figure 23. Modified Low-Pass I-R Plane with Associated Centers for Scene 1

the values obtained by optimization for Scene 1. The hope is that the values obtained by considering typical image examples would be sufficiently robust to produce good results over a larger class of images without reoptimization. We discuss how well this was realized in Section D.

Algorithm	a1	a2	a3	a4	a5	a6
Nonlinear–1	0.7637	0.0000	0.0000	0.0000	1.9256	-0.0024
Nonlinear–2	0.0000	4.0715	-0.0074	0.4586	0.0000	–
Nonlinear–3	0.6700	0.0000	0.0000	0.8587	0.0025	–

Table III. Optimization Coefficients for Nonlinear Mappings (Scene 1)

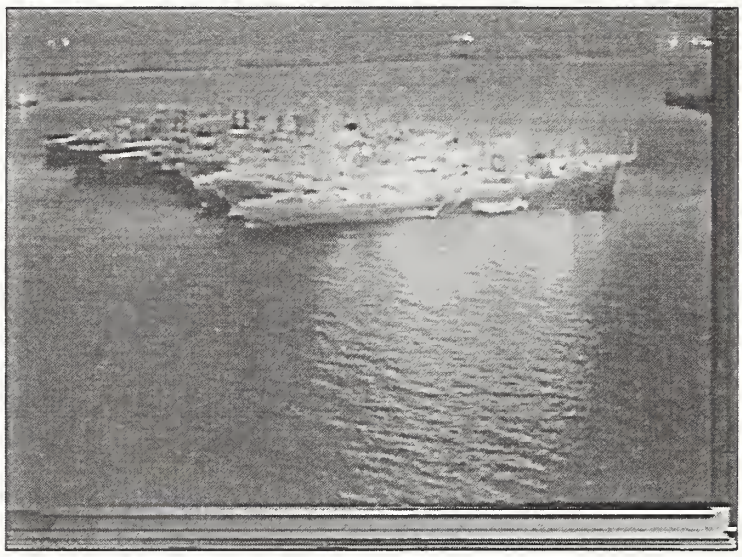

Figure 24. Algorithm Results Using the Direct Linear Mapping (Scene 1)

d. Complete Image Fusion

The final fused enhanced image m is obtained by adding the fused high-pass image k_H and the combined low-pass image k_L (derived according to one of the methods of Table II). The final fused image is thus given by

$$m(n_1, n_2) = k_H(n_1, n_2) + k_L(n_1, n_2). \tag{II.19}$$

C. ENHANCEMENT/FUSION RESULTS

In this section, the results of processing Scene 1 using the various methods for combining low-pass images are described. The results of the direct linear mapping, II.10, are shown in Figure 24, while the results of the weighted linear mapping, II.11,

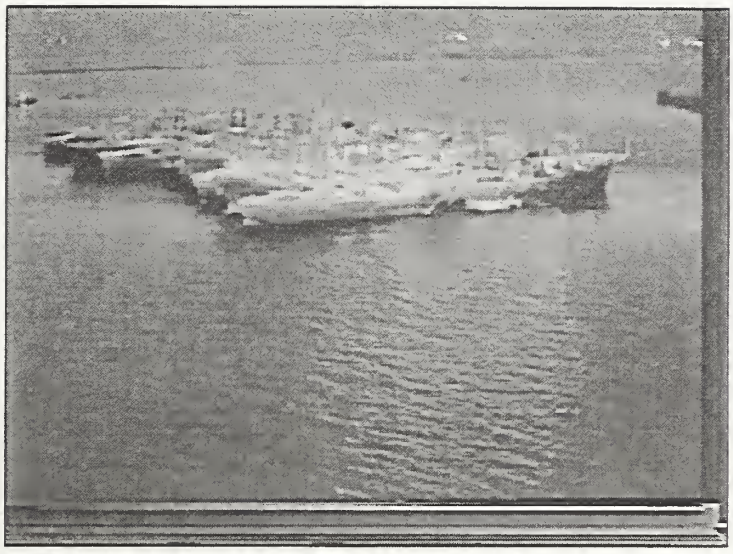

Figure 25. Algorithm Results Using the Weighted Linear Mapping (Scene 1)

are shown in Figure 25. The results of the three nonlinear mapping methods, II.12, II.13 and II.14, are given in Figures 26, 27 and 28, respectively.

To assess the performance of the enhancement/fusion algorithm, a visual comparison was made between the images processed according to the various methods and the original II and IR images. The performance criteria used included contrast, edge sharpness, clarity of details and among the processed images, scene content. Scene content refers to a qualitative measure of the extent that a combined image portrays "all the information" contained in both the II and IR images.

For Scene 1 it was judged that the best result was achieved from application of the nonlinear-1 mapping. This result is shown in Figure 26. The contrast of the image is superior to all others, the edges are sharpest, revealing mast, superstructure,

Figure 26. Algorithm Results Using the Nonlinear–1 Mapping (Scene 1) (This was judged to be the best result)

and hull-form details not evident in either original image. In addition, lighting details from the II image are portrayed while artifacts due to saturation ("blooming") are minimized. The other images shown in Figures 24, 25, 27, and 28, suffer from decreased contrast, and remnants from saturation in the II image; but in every case, the processed images offer an improvement to either original.

A set of five other image-pairs (Scenes 2 through 6) were processed as part of this thesis research. The results of processing are given in Appendix A.

Figure 27. Algorithm Results Using the Nonlinear–2 Mapping (Scene 1)

D. GENERAL ENHANCEMENT/FUSION RESULTS

The images processed in this thesis, Scenes 1-6, are generally similar (poor contrast, limited dynamic range, nighttime scenes) and can therefore be considered to be a particular "class" of images. As discussed in Section 3, it is desirable to consider a typical image within such a class, perform optimization on it and compute $\underline{a}$ so that it may then be used in the processing of all other images within the class. By using $\underline{a}$ from such a typical image, it would not be necessary to perform optimization for the remaining images, and therefore a significant savings in processing could be achieved.

The coefficients $\underline{a}$ derived from the nonlinear–3 mapping for Scene 2 were used as the class standard. Other sets of coefficients derived from other scenes were tried

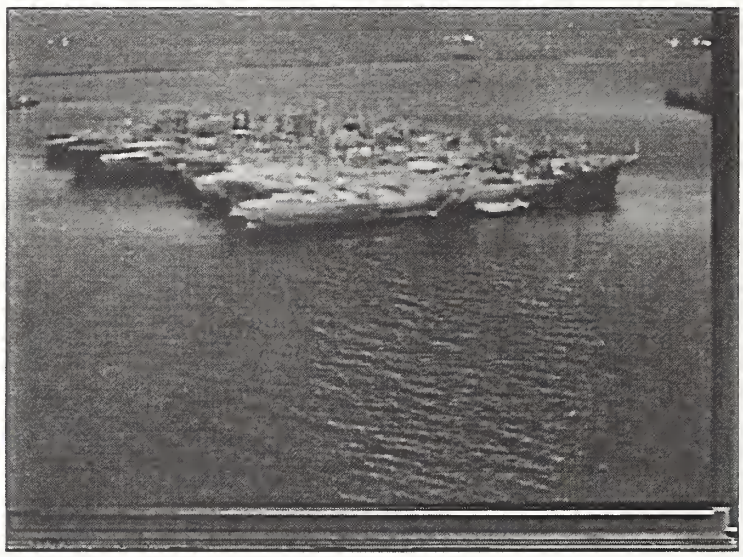

Figure 28. Algorithm Results Using the Nonlinear–3 Mapping (Scene 1)

on all image-pairs, but this set produced the best results. The nonlinear–3 mapping was chosen because of its preference in visual testing discussed in Chapter III. Figure 29 shows the result of using these coefficients on the image-pair of Scene 1. Results for all other images are included in Appendix B. The processing of Scene 1 by this method results in an image that we judge to be visually superior to either (II or IR) original image (based on criteria discussed in Section C), but poorer than any method using optimized coefficients.

The next chapter presents the methodology and results of human testing of the enhancement/fusion algorithm. Particularly, we examine the performance of each of the methods among themselves and as compared to the original single-band images.

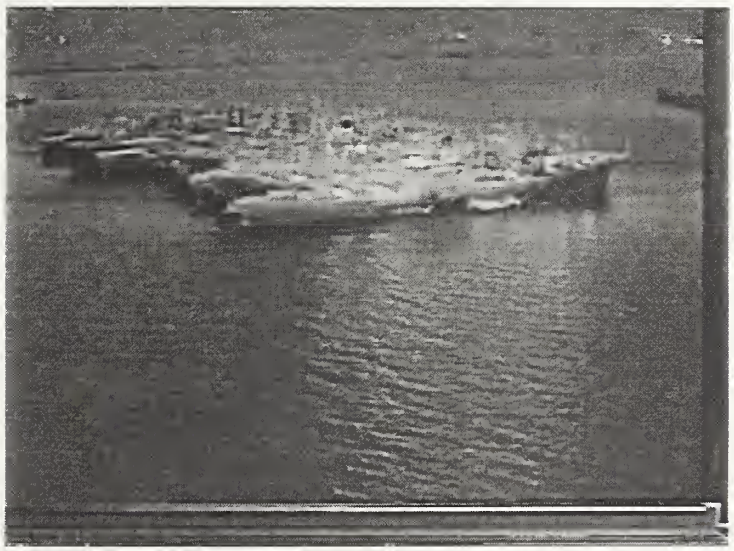

Figure 29. Algorithm Results Using General Coefficients Derived from Scene 2 (Scene 1)

III. TESTING RESULTS

A. INTRODUCTION

Any system whose goal is the improvement or enhancement of human sensory perception (e.g., visual, auditory, tactile) requires frequent interaction with the intended user throughout its design. Theoretical figures of merit and other engineering computations, are useful in quantifying such a systems performance or effectiveness, but are often inadequate in predicting human response. Therefore it is critical to perform human testing during system development.

To evaluate the performance of the various forms of the enhancement/fusion algorithm previously discussed, human perception testing was performed. The goal of this testing was to determine if monochrome fusion was an improvement over either of the single-band (II or IR) presentations, and if so, which of the five fusion methods was preferred.

B. PARTICIPANTS

Nine male military officers ranging in age from 30 to 39 years old with a mean age of 33.7 (2.8σ) volunteered for testing. All subjects signed informed consents and were briefed on the ethical conduct of subject participation specified in the Protection of Human Subjects, SECNAV Instruction 3900.39B. All subjects had normal or corrected to normal visual acuity (20/20). All subjects, except two, had military aviation experience, with an average of 1387 hours of flight time in their primary aircraft. All subjects, except one, had experience with either NVG or FLIR systems, and six had experience with both. The average NVG experience was 106 (102σ) hours and the FLIR experience was 92 (165σ) hours.

C. EQUIPMENT

A Sun SPARCstation 20 microprocessor workstation equipped with a 24-bit Parallax video card using a 20-inch RGB monitor (30.3-degree by 24.2-degree viewable area) with a resolution of 1280 by 1024 pixels (0.31 mm dot pitch) and frame rate of 76 Hz were used to present the stimuli. The participants viewed the screen from approximately 60 cm. The test room was darkened and a backlight was positioned on the floor behind the monitor to minimize the luminance differential between the monitor and environment.

D. IMAGES

Five night-time still image-pairs were collected from the U.S. Army Advanced Helicopter Pilotage program [Ref. 2, 13]. These image-pairs were obtained using a low-light visible Gen III image intensifier tube (0.6-0.9 μm) and a first generation FLIR display (8-12 μm). The five scenes were imaged in varying night-time illumination conditions and included both ocean surfaces as well as differing land terrains. This variation in lighting and terrain provided a wide range of reflectivity and emissivity conditions for analysis. All five image-pairs had 640×480 pixel resolution. The image-pairs were not spatially registered and required minor translational adjustments of no more than 20 pixels for proper registration. No geometric distortion was evident. A sixth image-pair was used with 276×508 pixel resolution and required no registration. Table IV list the scenes and a brief descriptor.

E. PROCEDURES

A two-alternative forced-choice procedure was used for the comparison of images. For each trial, two different representations of the same scene were presented to the subject in sequential order. At the beginning of each trial a fixation cross (0.67-degree) was displayed in the center screen. The subject initiated the first trial by clicking the left button of a trackball controller, 30 msec after this, the fixation

Scene	Descriptor
1	Nested ships
2	Winding road
3	Distant tower
4	Distant tower (near-view)
5	Fence on the horizon
6	Beach scene

Table IV. Image Scenes and Brief Descriptor

cross was extinguished, followed by presentation of the first image. The first image was displayed for 3000 msec, followed by a 30 msec blank, then the second image was displayed for 3000 msec. The subject responded "A," indicating preference of the first image displayed, by clicking the left button on the trackball controller or responded "B," indicating preference of the second image displayed, by clicking the right button on the trackball controller. The next trial began 100 msec after the response to the previous trial. A subject could repeat a trial without penalty. The test images were evaluated in six distinct blocks (a block for each scene), each block consisting of $\binom{7}{2} = 42$ image-pair trials, which accounts for all permutations of the seven image types (sensors) considered in this thesis. Therefore each subject observed 6×42 for a total of 252 image-pair trials.

Participants were presented all pairwise combinations of the images shown in Table V for each scene. This set included the original II and IR images of each scene and the images derived from their fusion. For all (nine) participants and all (six) scenes, each ordered image-pair was presented 54 times. For example, the combination of presenting the direct linear mapping derived image (image 2) followed by the weighted linear mapping derived image (image 6) occurred 6 times per participant, once per scene. Therefore the ordered-pair 2-6, was encountered 6 times per participant for the 9 participants or $6 \times 9 = 54$ times.

Identifier	Image Type/Algorithm
0	Original II
1	Original IR
2	Direct Linear Mapping
3	Weighted Linear Mapping
4	Nonlinear–1 Mapping
5	Nonlinear–2 Mapping
6	Nonlinear–3 Mapping

Table V. Image Types and their Identifiers Used in Visual Testing

For each image-pair, the evaluation criterion was, "Between the two, which image conveys the most information about the scene?" The intent of this criterion was to determine which sensor fusion algorithm provided the most scene content for a given scene.

F. ANALYSIS

1. Method of Discordances

To determine fusion algorithm performance from visual testing results, a method is used which analyzes image orderings by counting discordances [Ref. 18, 19]. The ordering with the fewest number of discordances is considered the best. From this analysis, a ranking of images and thus algorithms, most to least preferred, is determined.

Table VI shows the results of the visual testing. The table indicates the number of times the image presented first was preferred over the image presented second. Row elements represent the first image, while column elements represent the second image. Entries in the table represent the number of times the first image was chosen over the second image. For example, when image 2 was presented before image 6, image 2 was chosen 17 times, and when image 6 was presented before image 2, image 6 was chosen 19 times.

		2nd Image						
		0	**1**	**2**	**3**	**4**	**5**	**6**
1st	**0**	NA	16	11	14	17	16	12
I	**1**	37	NA	14	14	10	13	10
m	**2**	42	39	NA	19	21	30	17
a	**3**	42	42	18	NA	18	28	13
g	**4**	40	44	16	18	NA	29	13
e	**5**	38	36	23	17	11	NA	14
	6	39	40	19	26	23	28	NA

Table VI. Image Preferences Based on Visual Testing

The ordering of the images refers to the ranking of each image from most preferred to least. For example, the ordering 6-0-2-1-5-3-4 indicates that image 6 is most preferred, image 0 is the next preferred, and so on. Given seven images, there are $7! = 5040$, possible orderings to consider. For a particular ordering, the number of discordances is the number of times a particular image is preferred contrary to the order specified. For example if the ordering is 6-0-2-1-5-3-4, any observation where image 0 was preferred over image 6 would be recorded as a discordance.

For image-pair (0-1), of the 54 presentations when image 0 was presented first, it was preferred 16 times and therefore image 1 was preferred 38 times $(54 - 16)$. When image 1 was presented first, of the 54 presentations, it was preferred 37 times and therefore image 0 was preferred 17 times $(54 - 37)$. This pair would contribute $38+37 = 75$ discordances for those orderings where image 0 was ordered preferentially over image 1, and $16 + 17 = 33$ discordances for those orderings where image 1 was ordered preferentially over image 0. Based on this analysis, over the complete set of orderings, Table VII indicates the ranking of algorithms. The preferred ranking is 6-2-3-4-5-1-0, indicating that overall, the nonlinear–3 mapping algorithm performed the best, followed by the direct linear mapping algorithm, weighted linear mapping algorithm, nonlinear–1 mapping algorithm, nonlinear–2 mapping algorithm, original IR image, and finally the original II image.

6	2	3	4	5	1	0
6	2	4	3	5	1	0
6	3	2	4	5	1	0
2	6	3	4	5	1	0
2	6	4	3	5	1	0
6	4	2	3	5	1	0
6	3	4	2	5	1	0
6	4	3	2	5	1	0

Table VII. Preferred Ordering Based on Method of Discordances

2. Ordering Effects

Another consideration in pairwise analysis is the effect of ordering. For example, if image-pair (0-1) is shown and then image-pair (1-0), it is expected that there should be no variation in preference, i.e., ordering should not effect the decision made. Therefore, by examining ordering effects, we can measure the certainty viewers have for a particular choice (confidence). For decisions of high certainty, order differences should have a minimal impact on choice, while for decisions of low certainty, order differences may evoke large variations in choice.

A nonparametric test, sign test [Ref. 20], was used to test the direction of the differences between two sets of data. Specifically, did subjects prefer one interval compared to another interval regardless of image or sensor type. For a given image-pair, the number of times the second interval was preferred over the first was examined. Preference for the second interval was indicated by a plus sign (+), while preference for the first interval was indicated by a minus sign (−). No preference was indicated by a zero (0). The comparison results are shown in Table VIII.

To test the order effect, the probability of committing a Type I error (rejecting the null hypothesis, H_0, when it is, in fact, true) will be set at 0.025. If there was an

Order	2nd Image Preferred	Order	1st Image Preferred	Sign Test
0-1	38	1-0	37	+
0-2	43	2-0	42	+
0-3	40	3-0	42	−
0-4	37	4-0	40	−
0-5	38	5-0	38	0
0-6	42	6-0	39	+
1-2	40	2-1	39	+
1-3	40	3-1	42	−
1-4	44	4-1	44	0
1-5	41	5-1	36	+
1-6	44	6-1	40	+
2-3	35	3-2	18	+
2-4	33	4-2	16	+
2-5	24	5-2	23	+
2-6	37	6-2	19	+
3-4	36	4-3	18	+
3-5	26	5-3	17	+
3-6	41	6-3	26	+
4-5	25	5-4	11	+
4-6	41	6-4	23	+
5-6	40	6-5	28	+

Table VIII. Ordering Effects on Visual Testing

order effect, the alternative hypothesis (H_1) will be accepted. Accordingly, setting a stringent level of significance at 0.025 will assure that the rejection of the null hypothesis was due to the independent variable rather than due to chance alone.

The null hypothesis H_0, is the probability that there is no order effect, i.e., each sensor has the same probability of being chosen in either interval, and is represented by

$$H_0 : P[X_i > Y_i] = P[X_i < Y_i] = 1/2 \,, \tag{III.1}$$

where X_i is subject's preference for interval one and Y_i is subject's preference for interval two. When the null hypothesis is true, half the pairs will yield a positive sign and the other half will yield a negative sign. H_0 will be rejected (H_1), if too few differences of sign occur, implying that subject's prefer interval two over interval one regardless of sensor or image type. The probability associated with the occurrence of a particular number of pluses and minuses can be calculated from the binomial

distribution as

$$P[Y_i \geq x] = \sum_{i=x}^{N} \binom{N}{i} p^i q^{N-i} , \qquad \text{(III.2)}$$

where N is the total number of pairs and x is the total number of instances where the second interval is prefererred over the first (number of pluses). Note that two orders were in agreement (represented by zero) and therefore the sample size was reduced from $N = 21$ to $N = 19$ and x the number of pluses is 16. From the binomial distribution, the probability of observing 16 or more pluses has a one-tailed probability, when H_0 is true, of 0.0022. Since 0.0022 is less than 0.025, H_0 is rejected and H_1 is accepted. Thus we can conclude subjects preferred interval two regardless of sensor characteristics or scene texture compared to interval one.

Based on this analysis, we can make the following remarks. The II images (image 0) are never preferred over other images. IR images (image 1) are only preferred over II images (image 0). Direct mapping images are preferred over nonlinear–1 mapping images. All other image comparisons resulted in low confidence and no additional conclusive remarks can be made.

IV. CONCLUSIONS

This thesis has introduced a new algorithm which performs adaptive enhancement and fusion of low-light visible and infrared imagery. Variants of the algorithm were used to process six nighttime scenes and the composite images were tested along with the originals to determine which image displayed "maximum scene content". Based on the results of this study, the following conclusions were reached:

The enhancement/fusion algorithm result (all variants) is superior to either single-band image (II and IR). This is based on the results of human testing where the method of discordances revealed that the preferred images, in order of best to worst, were derived from the nonlinear–3 mapping, direct linear mapping, weighted linear mapping, nonlinear–1 mapping, nonlinear–2 mapping, original IR, original II. From ordering effects analysis the data again shows that all other mappings were preferred over the single-band imagery (II and IR).

The best variant of the enhancement/fusion algorithm is not clear. As stated before, all variants are preferred over single-band imagery, but the results of the discordance analysis and ordering effects analysis are inconclusive. The discordance method points to nonlinear–3 mapping as the most preferred variant; however the ordering analysis shows low certainty when nonlinear–3 mapping is compared to all other variants. This lack of certainty is exhibited for all comparisons among variants. Therefore, it appears that the variants perform comparably.

Given that the visual testing shows that performance differences are minor, the next consideration in comparing variants is their computational efficiency. The direct linear mapping, which averages modified low-pass pixels from each image, is the least computationally intensive and is therefore recommended. However weighted linear mapping is only slightly more computationally intensive and provides additional flexibility. This may also be a good choice.

Although a number of variations were implemented, and perceptual testing was conducted, a number of additional issues could be considered for further evaluation of the algorithm.

During enhancement, when examining a class of similar images, a class specific set of gain and local luminance transformation curves can be used. While this may degrade performance on some scenes, the use of more generic transformation curves on an entire class of images is ultimately necessary in a practical implementation of the algorithm.

To conduct human testing, six scenes were analyzed using nine participants. For future studies, greater benefit would be achieved by increasing both the number of scenes processed and presented as well as the number of subjects tested.

Ordering effects are an important consideration in pairwise comparisons of images. The testing revealed that in almost every case, that algorithm variant preference was a function of presentation order. The viewer's assessment of a particular image-pair was different depending on the order the pair was presented. To preclude such anomalies, future pairwise testing procedure should not use sequential presentation of images, but instead use simultaneous presentation of the images.

APPENDIX A. ENHANCEMENT/FUSION RESULTS

The following images result from the application of the enhancement/fusion algorithm presented in this thesis. The image labels (m0–m4) refer to the particular method used in the algorithm.

Identifier	Method
m0	Direct Linear Mapping
m1	Weighted Linear Mapping
m2	Nonlinear–1 Mapping
m3	Nonlinear–2 Mapping
m4	Nonlinear–3 Mapping

Table IX. Image Identifiers and Associated Algorithm Methods

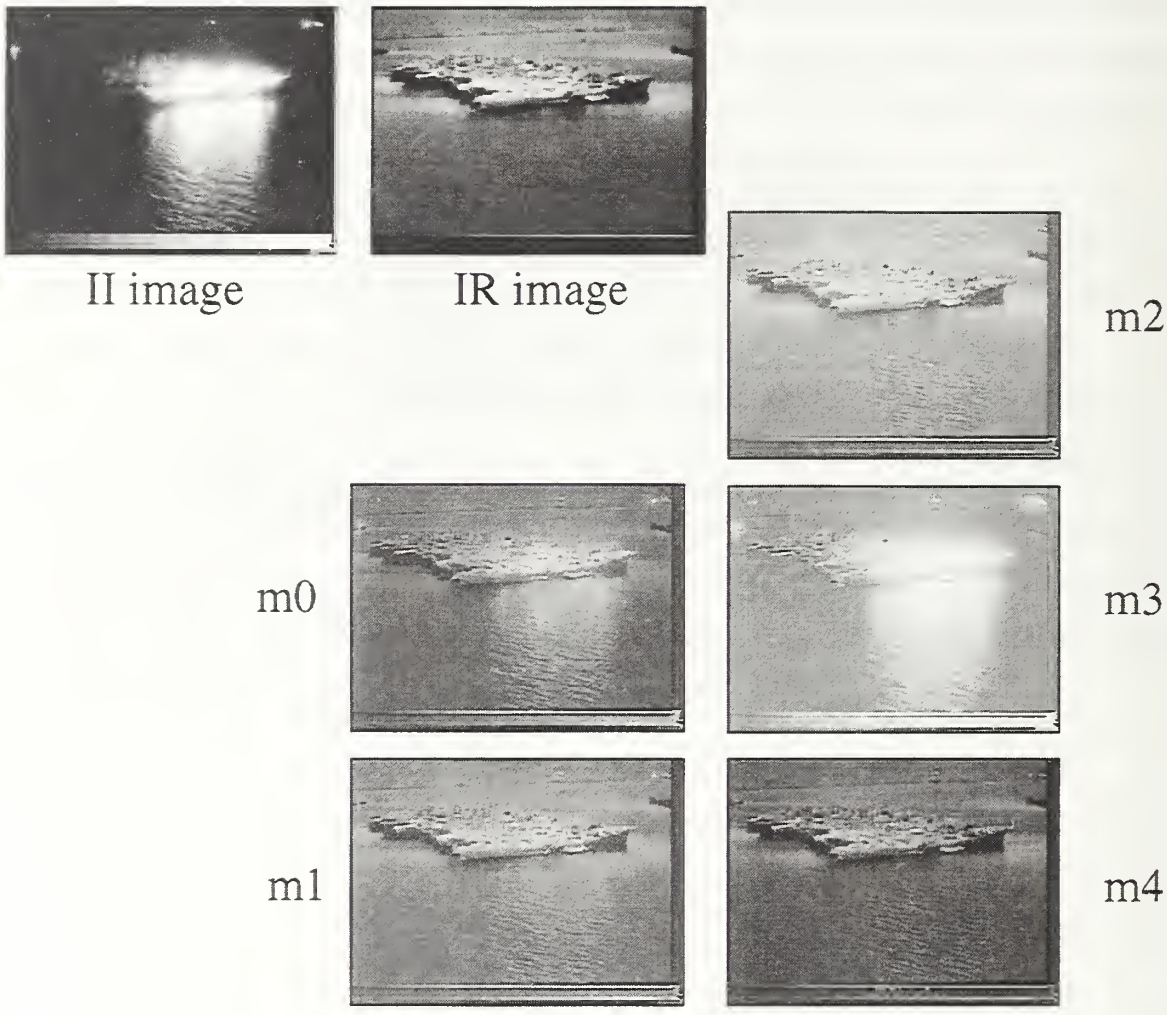

Figure 30. Set of Processed Images for Scene 1

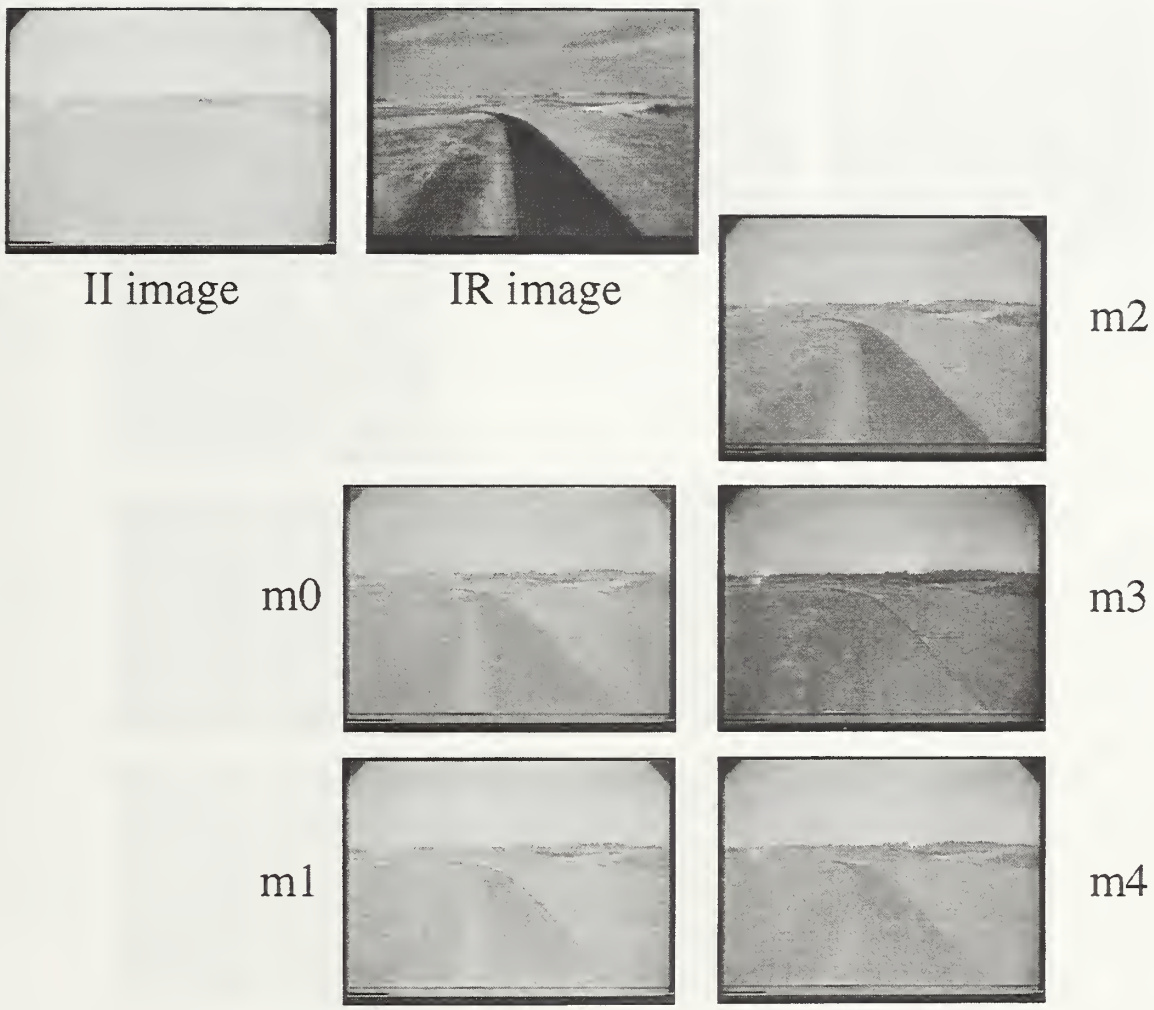

Figure 31. Set of Processed Images for Scene 2

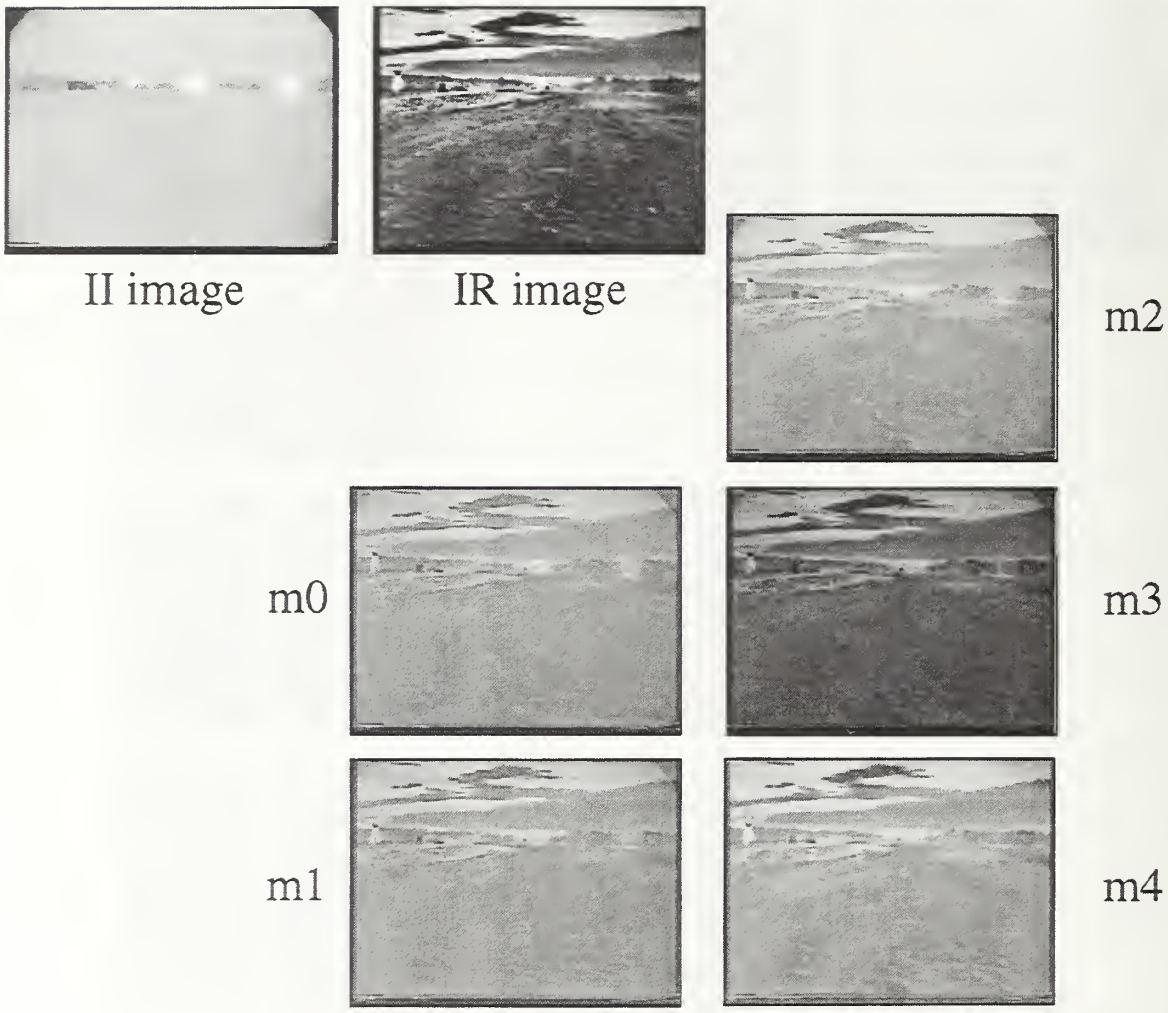

Figure 32. Set of Processed Images for Scene 3

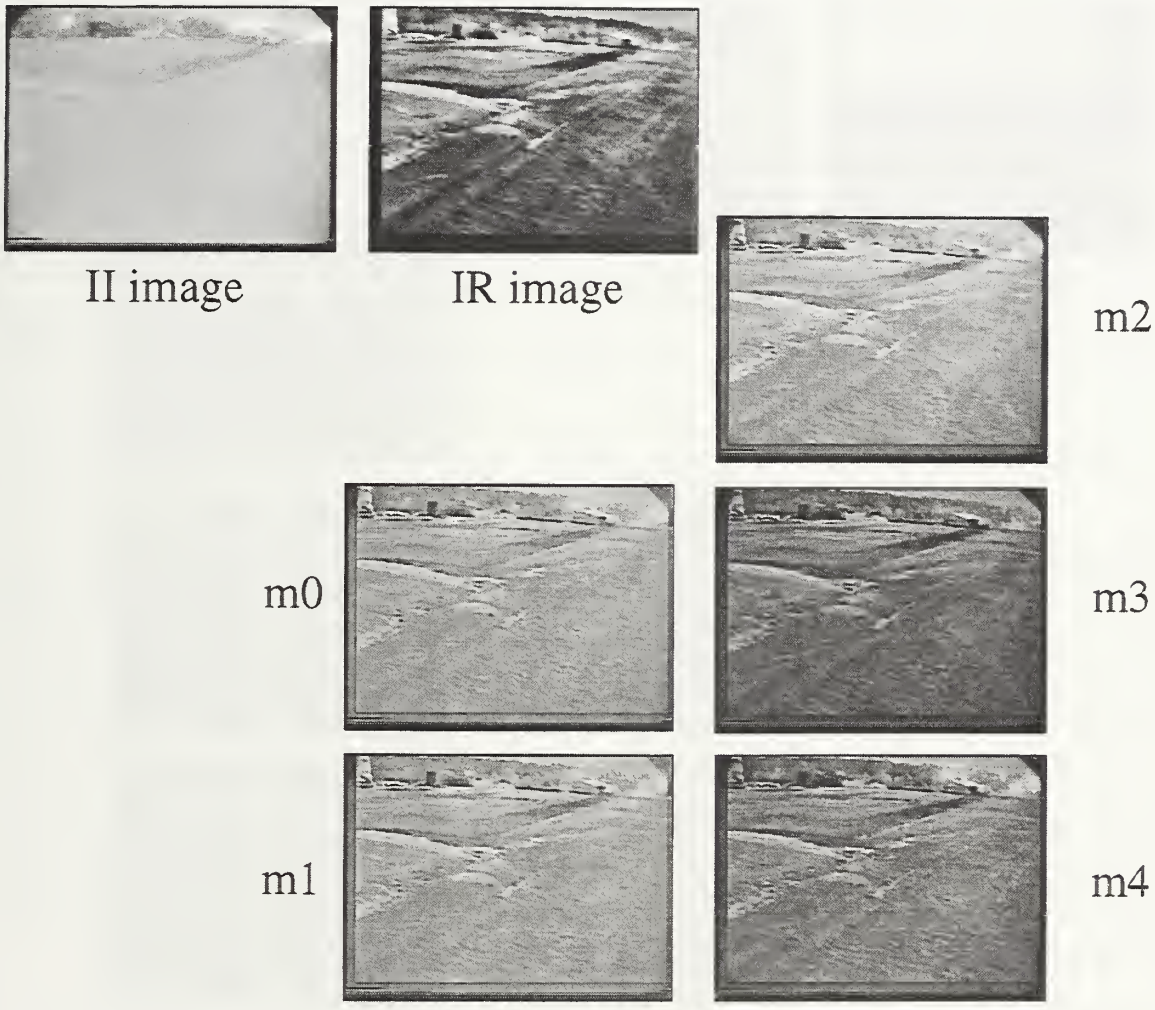

Figure 33. Set of Processed Images for Scene 4

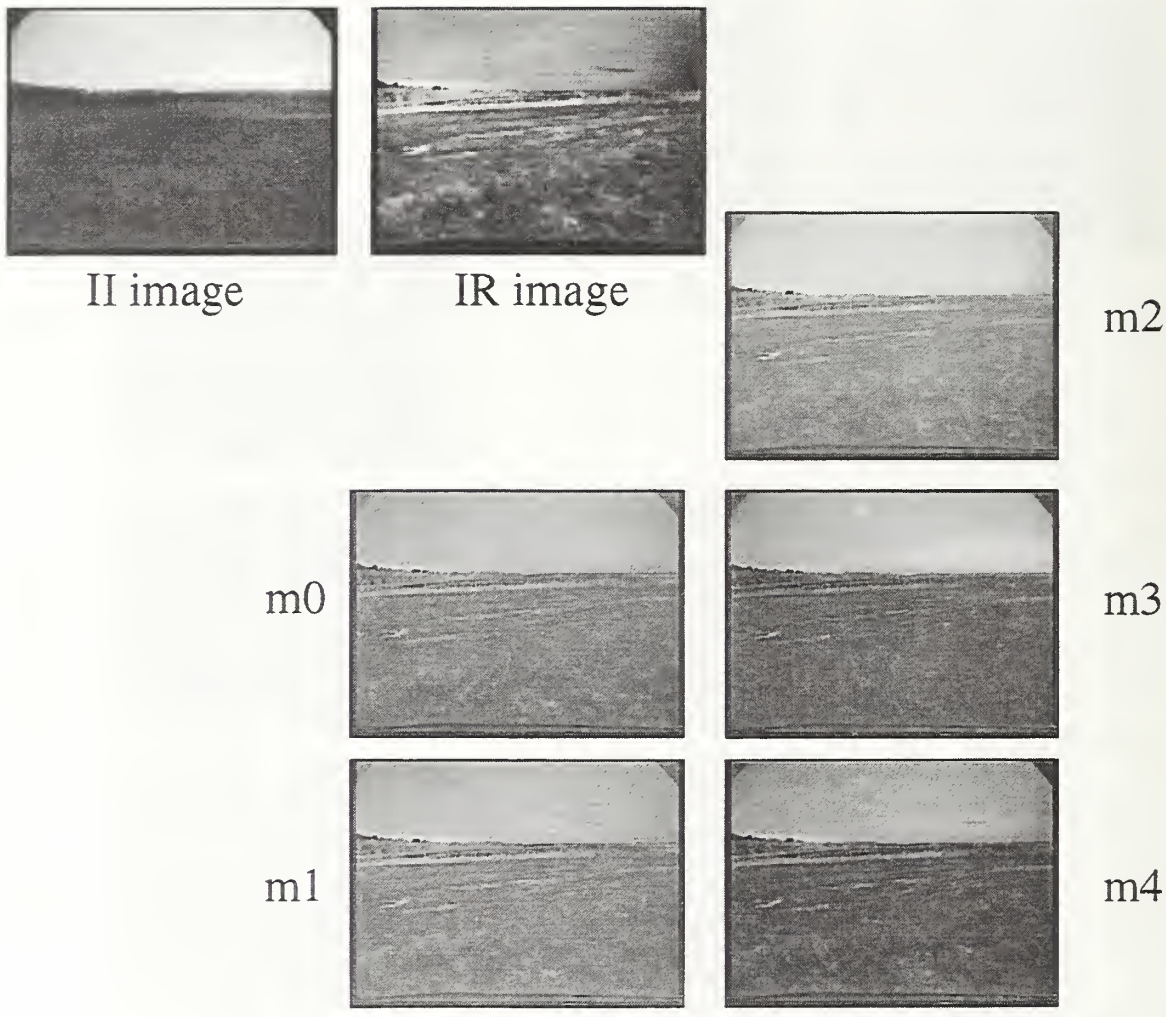

Figure 34. Set of Processed Images for Scene 5

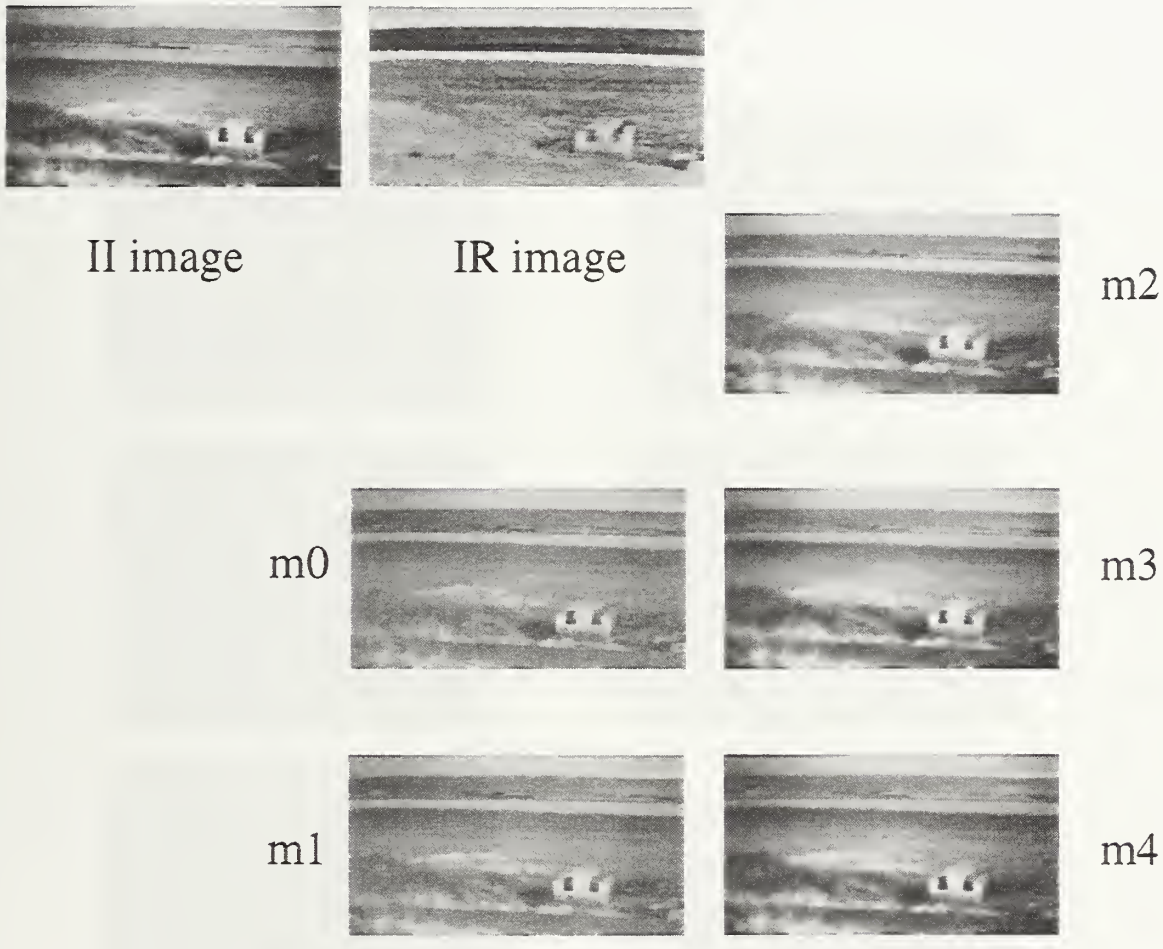

Figure 35. Set of Processed Images for Scene 6

APPENDIX B. GENERAL ENHANCEMENT/FUSION RESULTS

The following images result from the general enhancement/fusion method discussed in Chapter II. The image labels (1–6) refer to the image scene.

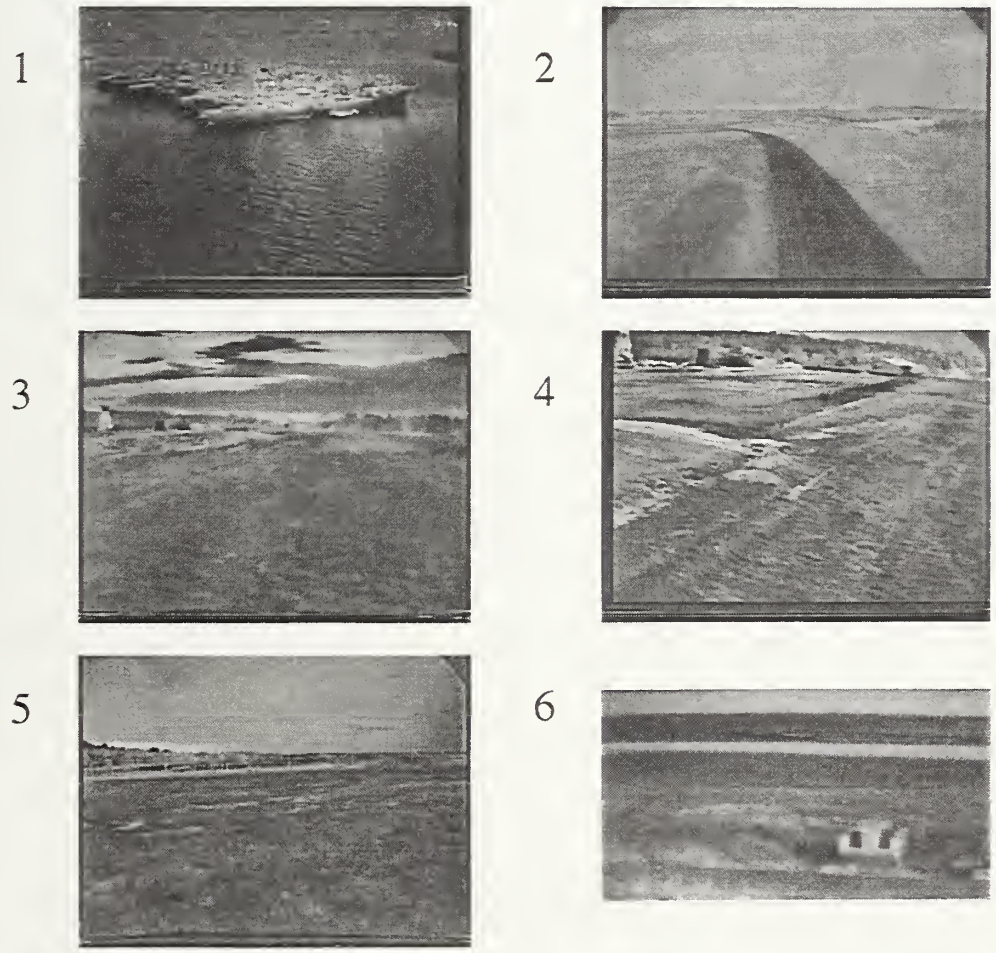

Figure 36. Set of General Processed Images

APPENDIX C. GAIN AND LUMINANCE TRANSFORMATION CURVES

Transformation curves used for the results shown in Appendix A.

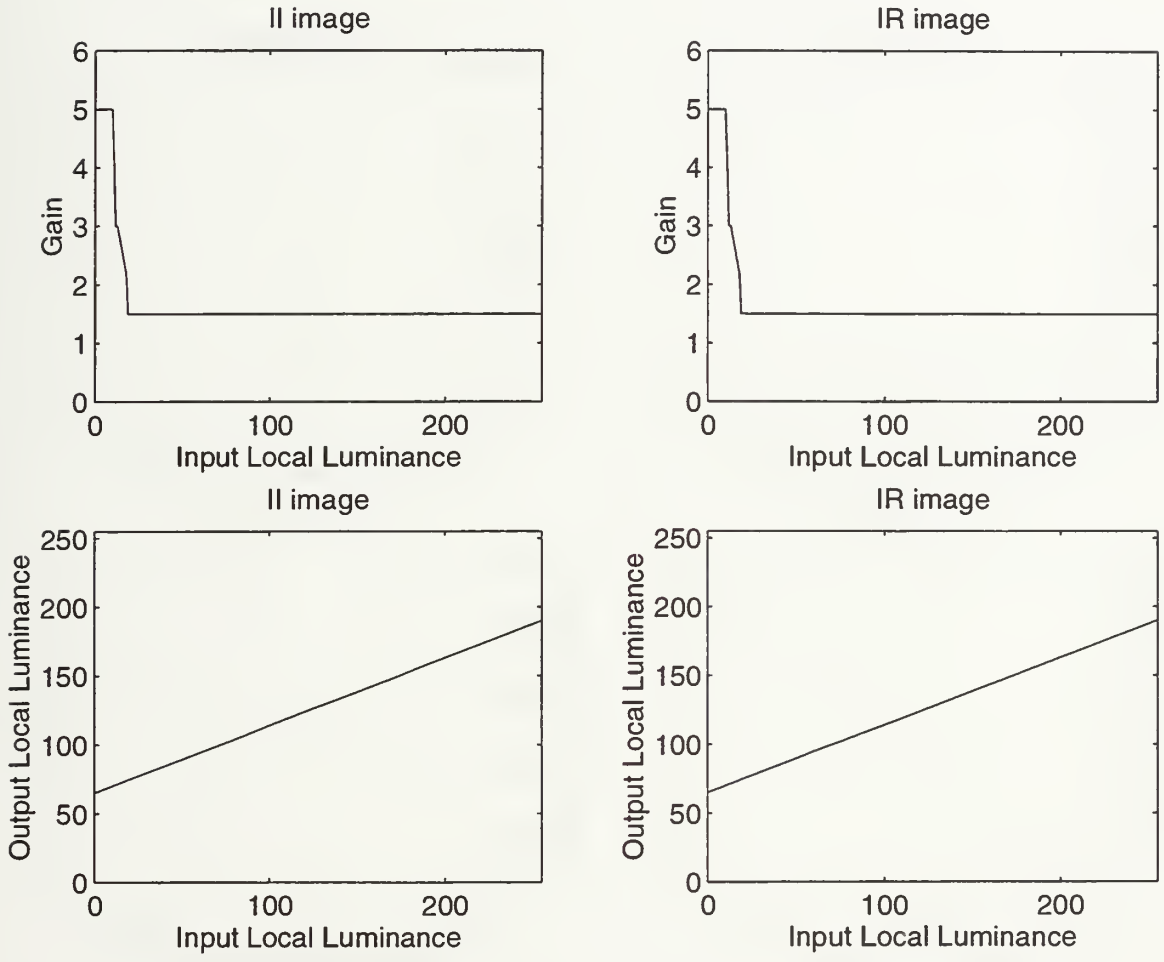

Figure 37. Gain and Local Luminance Transformation Curves for Scene 1

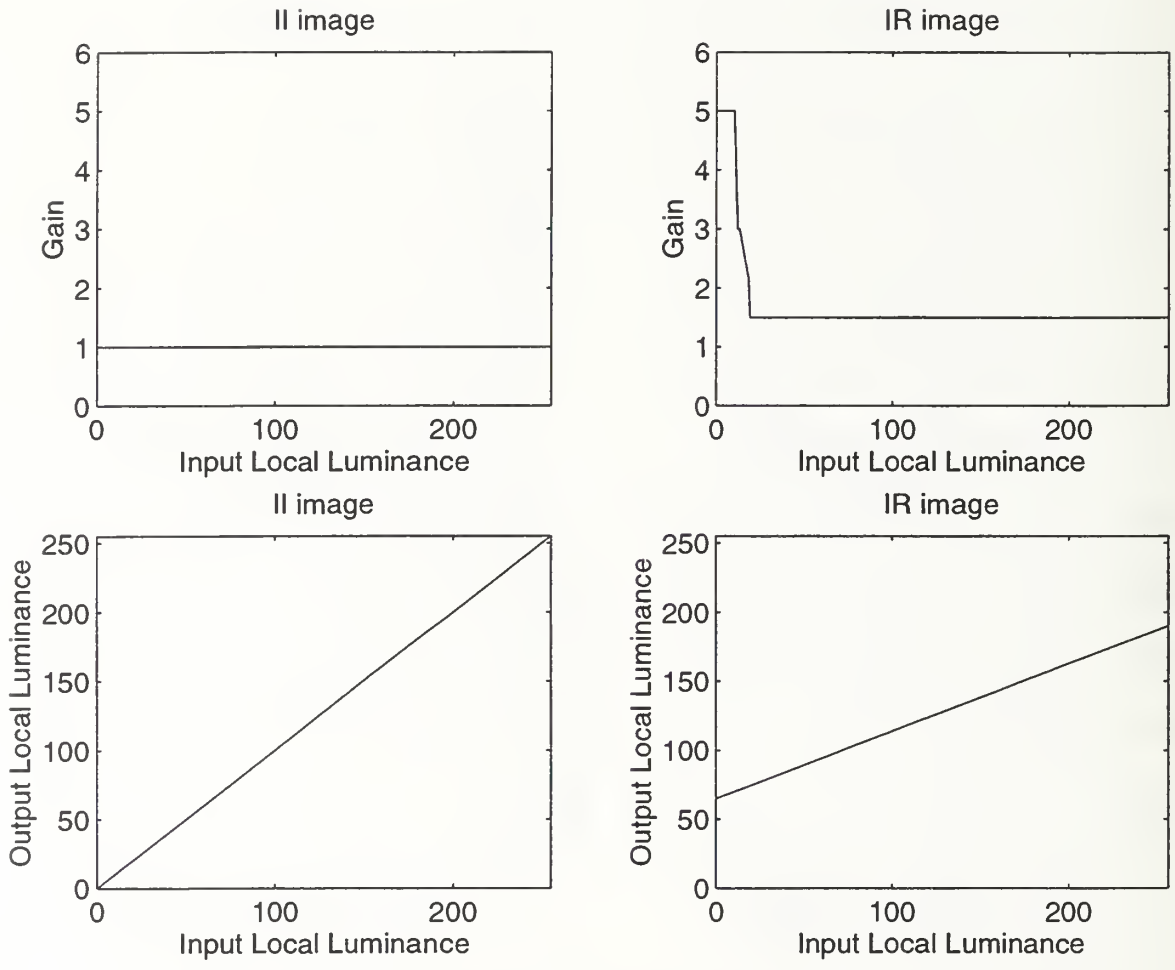

Figure 38. Gain and Local Luminance Transformation Curves for Scene 2

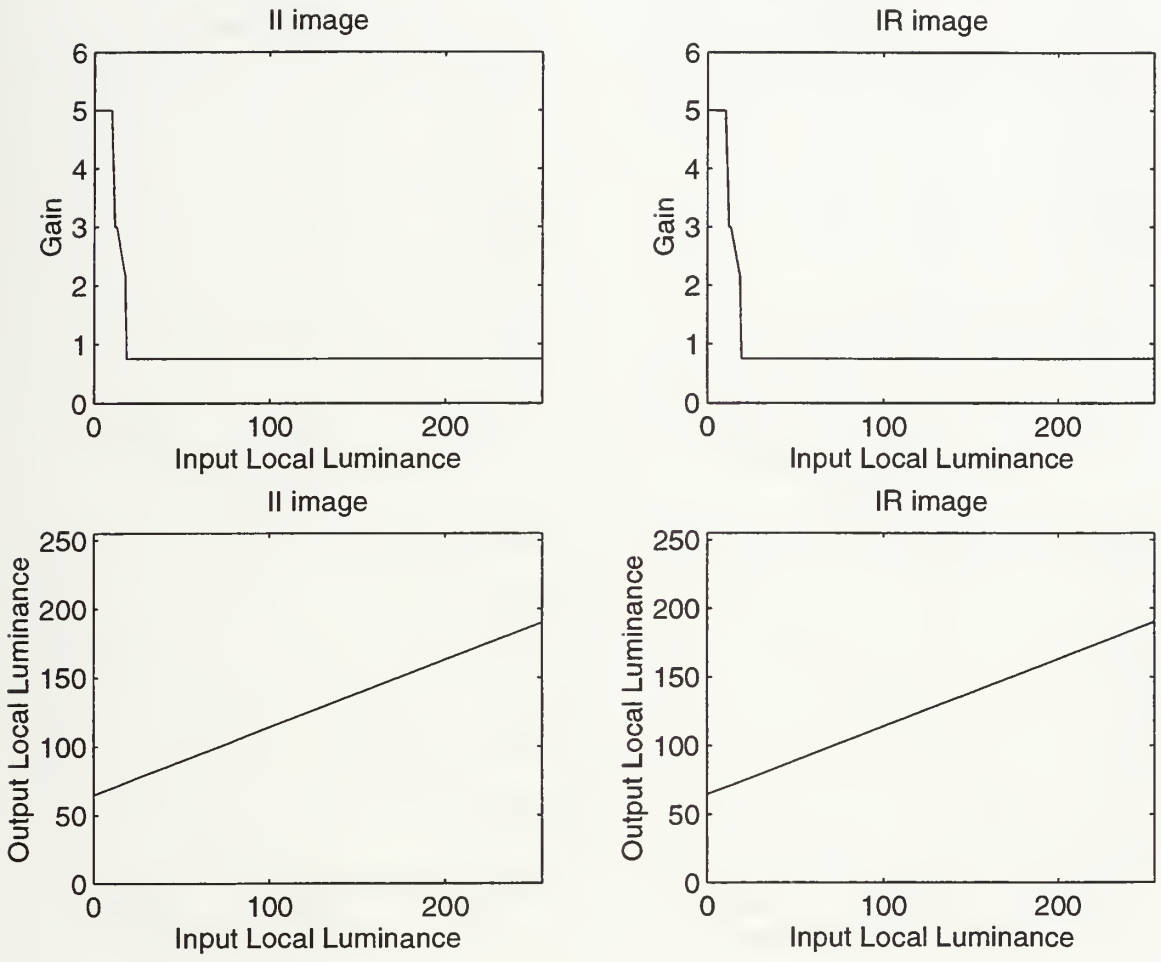

Figure 39. Gain and Local Luminance Transformation Curves for Scene 3

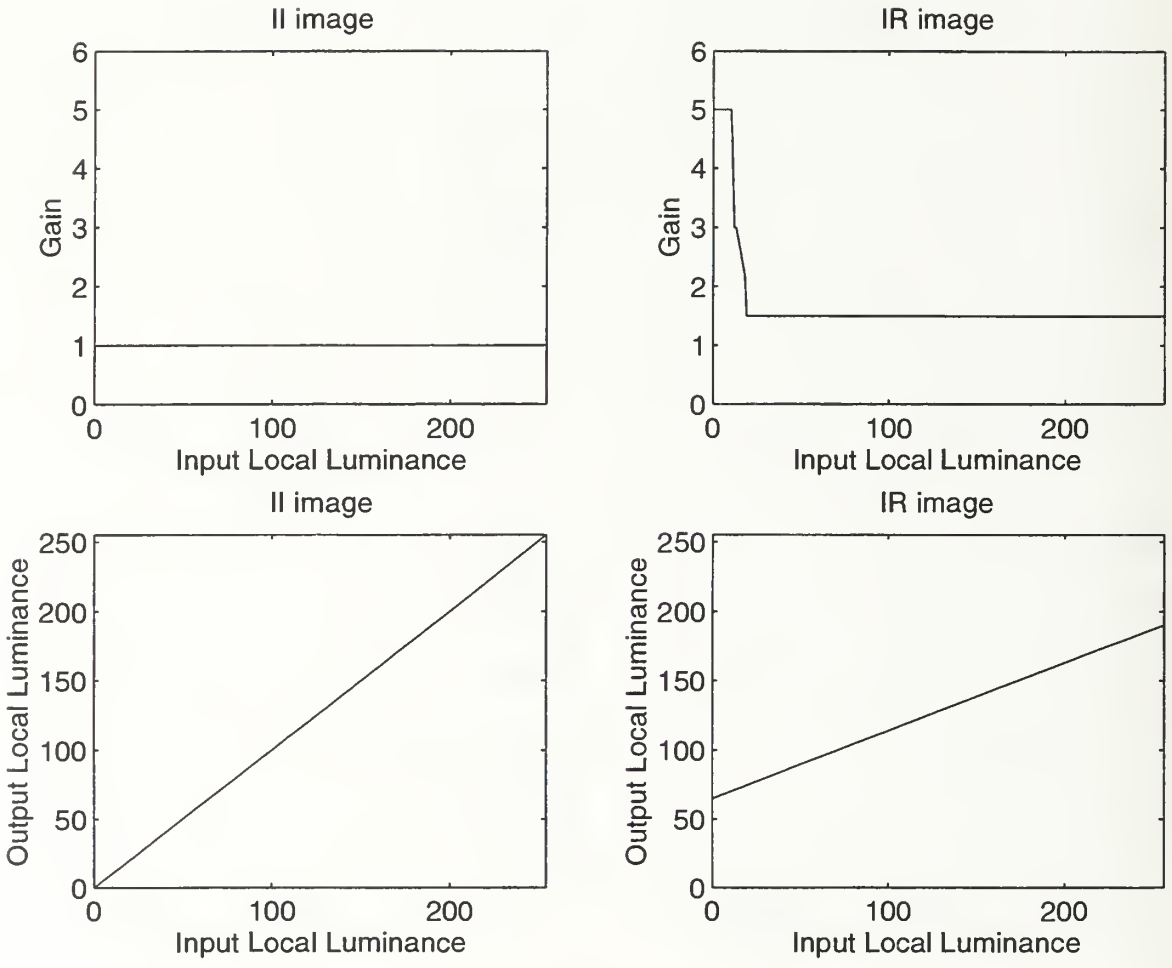

Figure 40. Gain and Local Luminance Transformation Curves for Scene 4

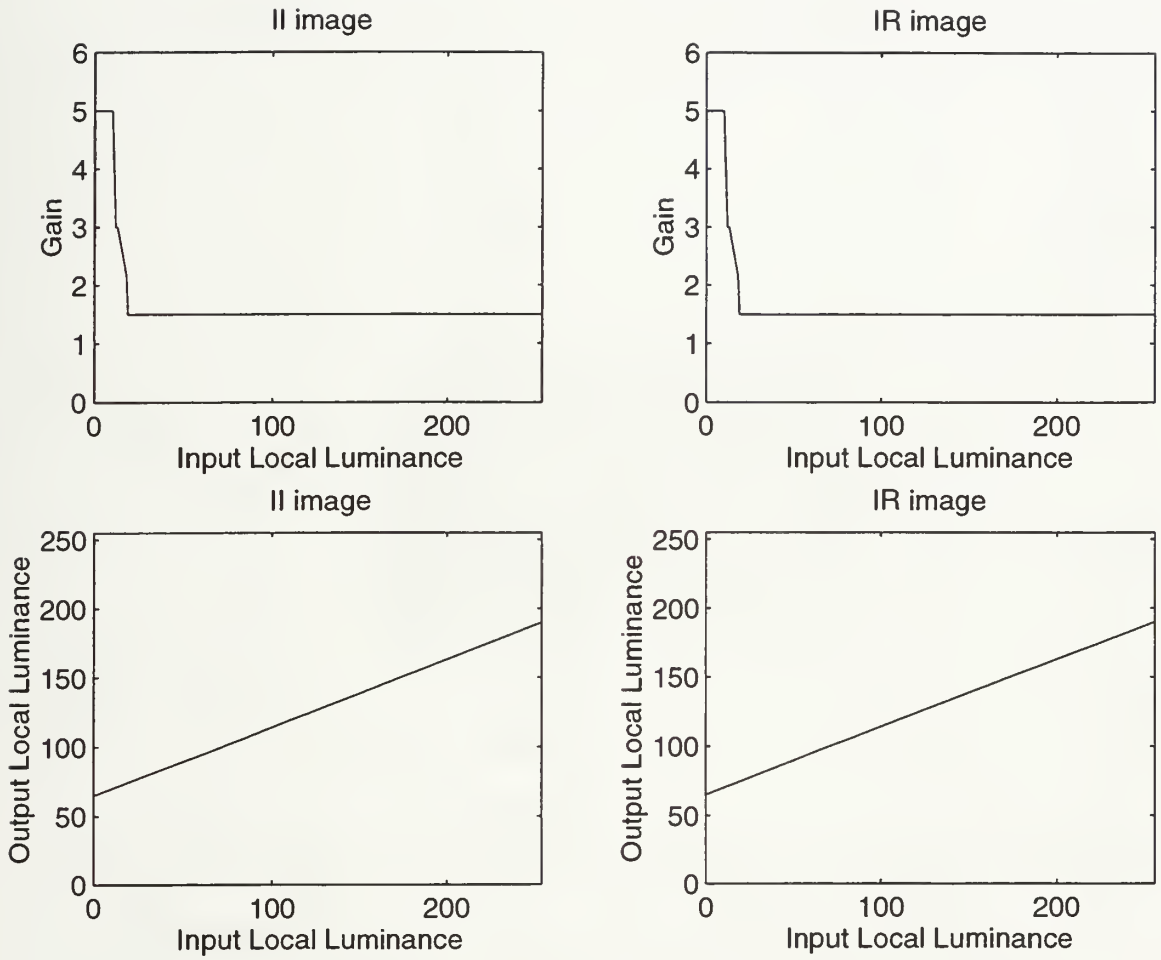

Figure 41. Gain and Local Luminance Transformation Curves for Scene 5

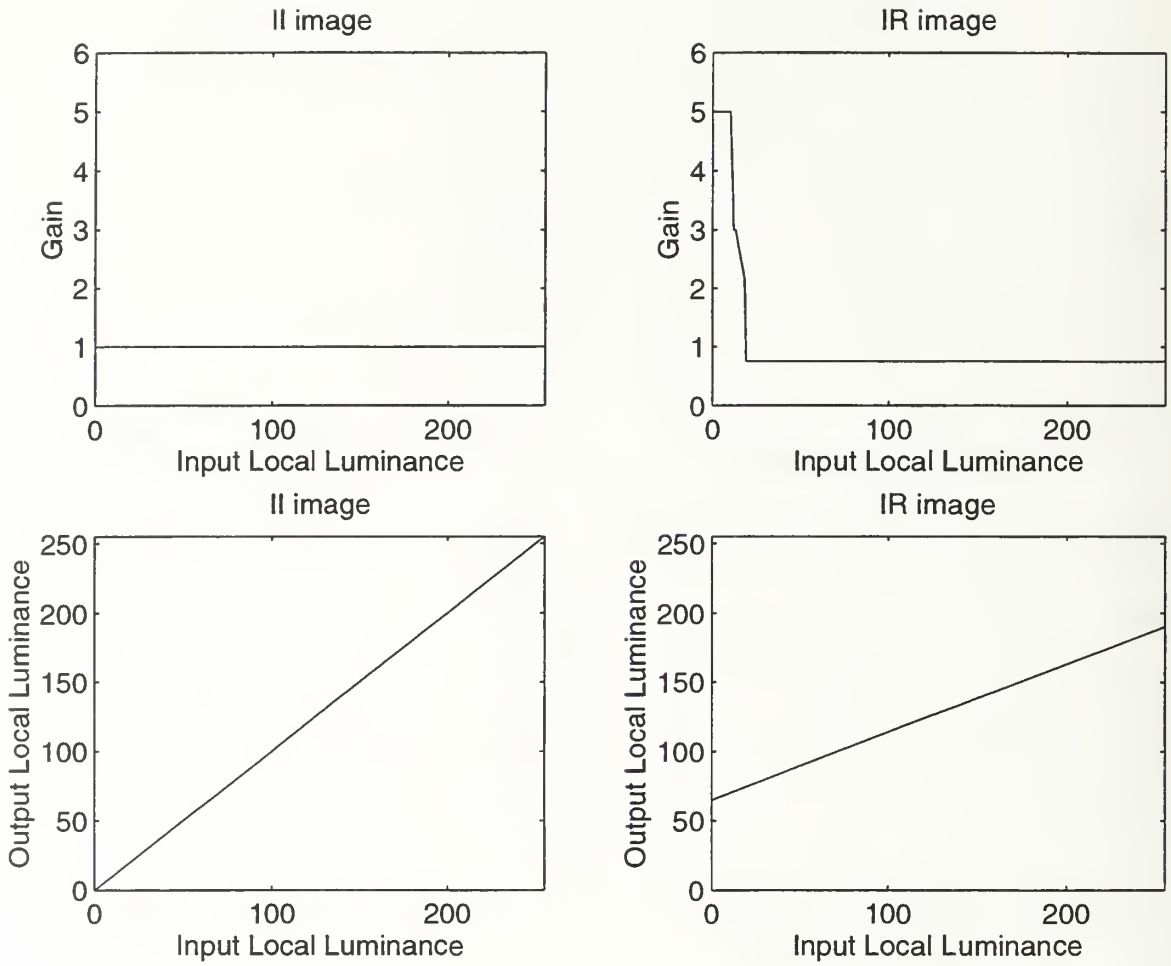

Figure 42. Gain and Local Luminance Transformation Curves for Scene 6

APPENDIX D. CLUSTERING RESULTS USED TO IDENTIFY NONLINEAR MAPPING COEFFICIENTS

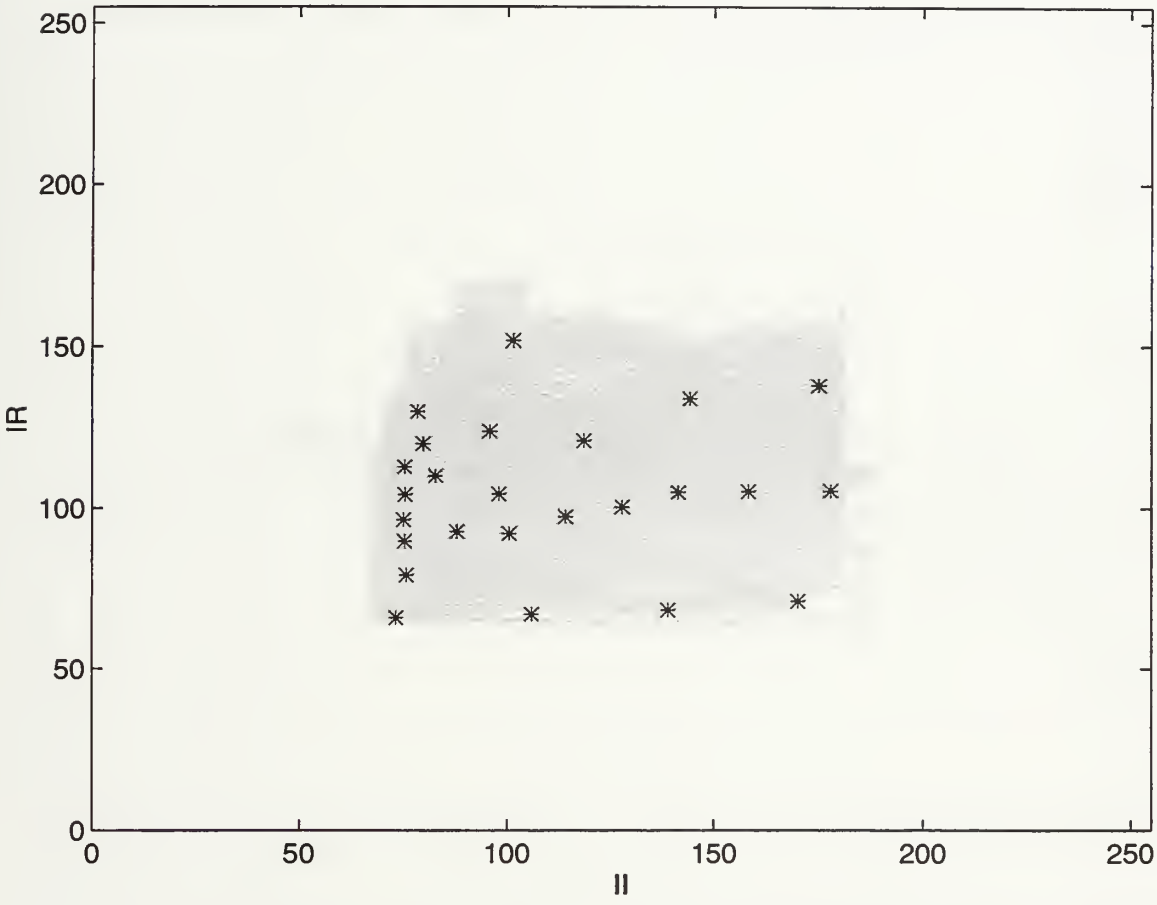

Figure 43. Modified Low-Pass I-R Plane with Associated Centers for Scene 1

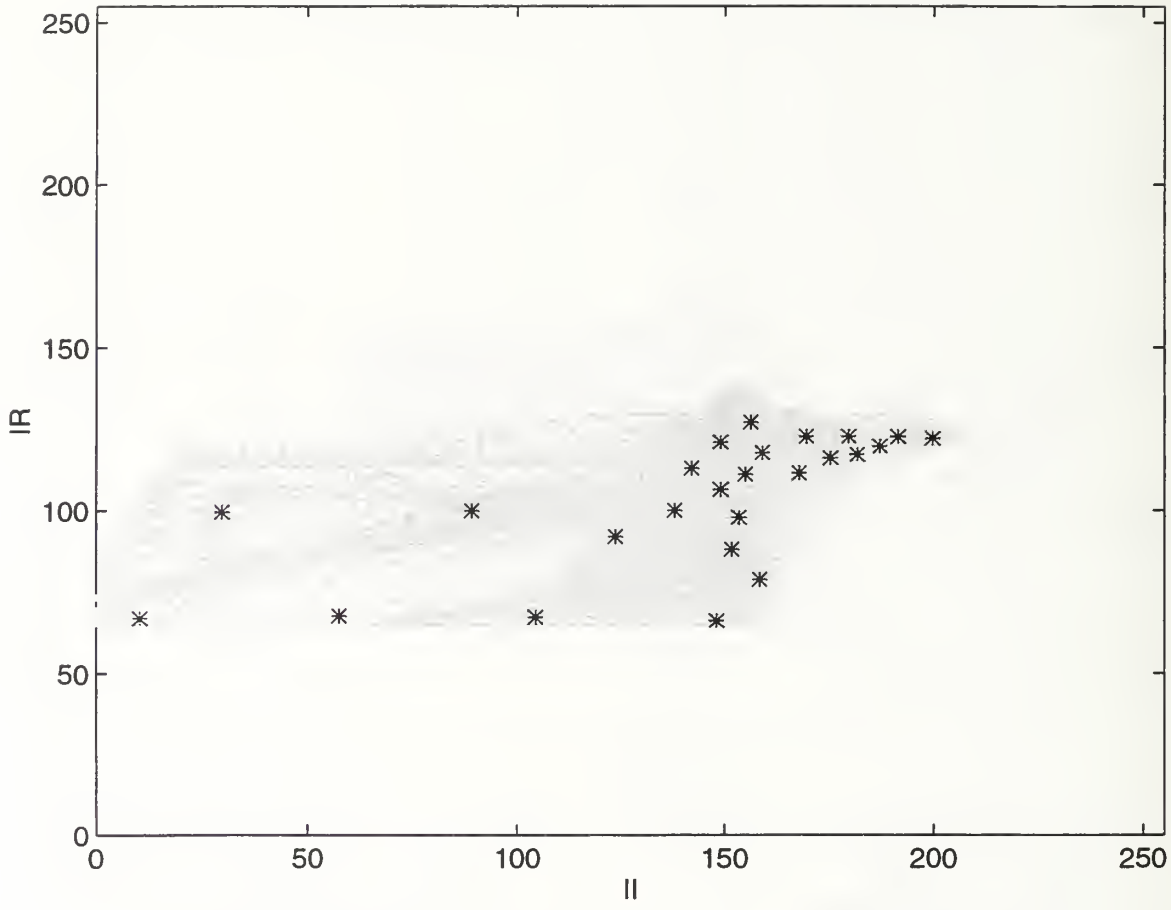

Figure 44. Modified Low-Pass I-R Plane with Associated Centers for Scene 2

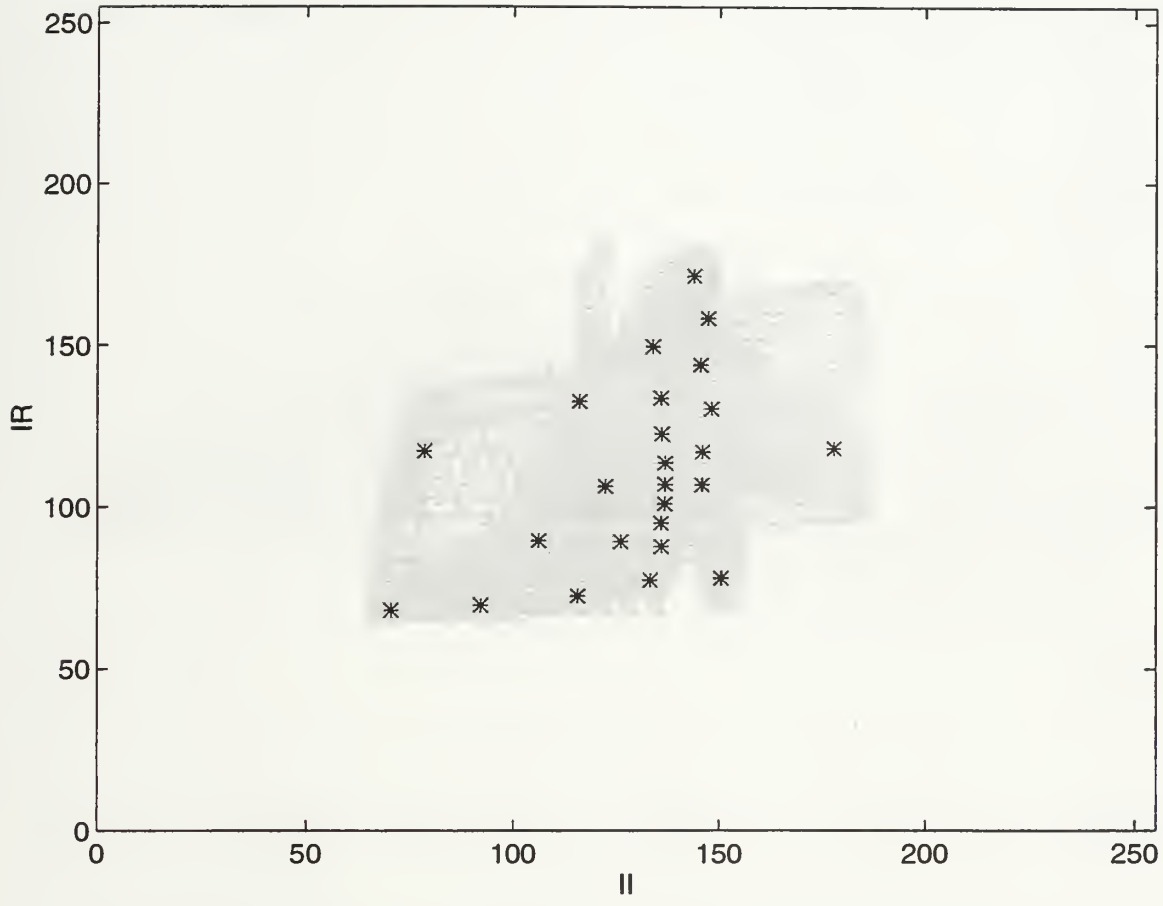

Figure 45. Modified Low-Pass I-R Plane with Associated Centers for Scene 3

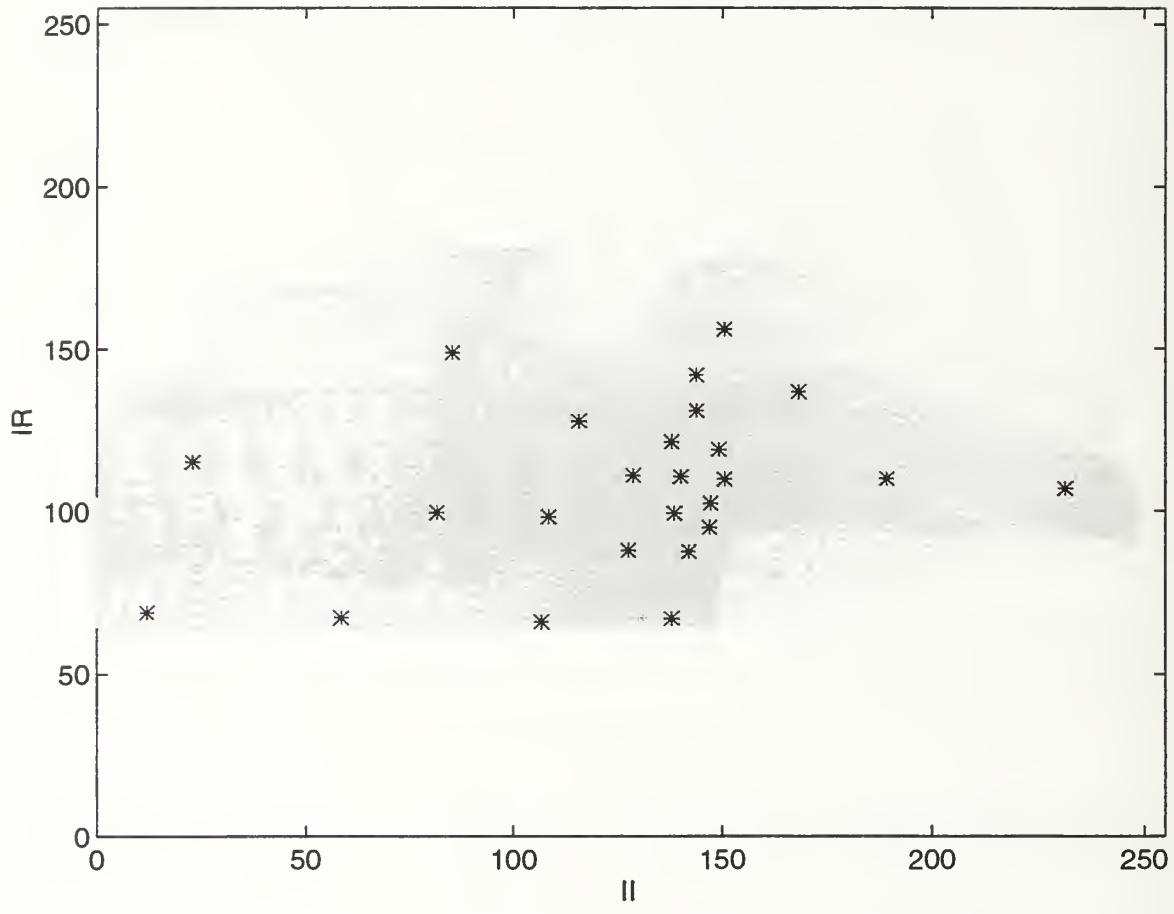

Figure 46. Modified Low-Pass I-R Plane with Associated Centers for Scene 4

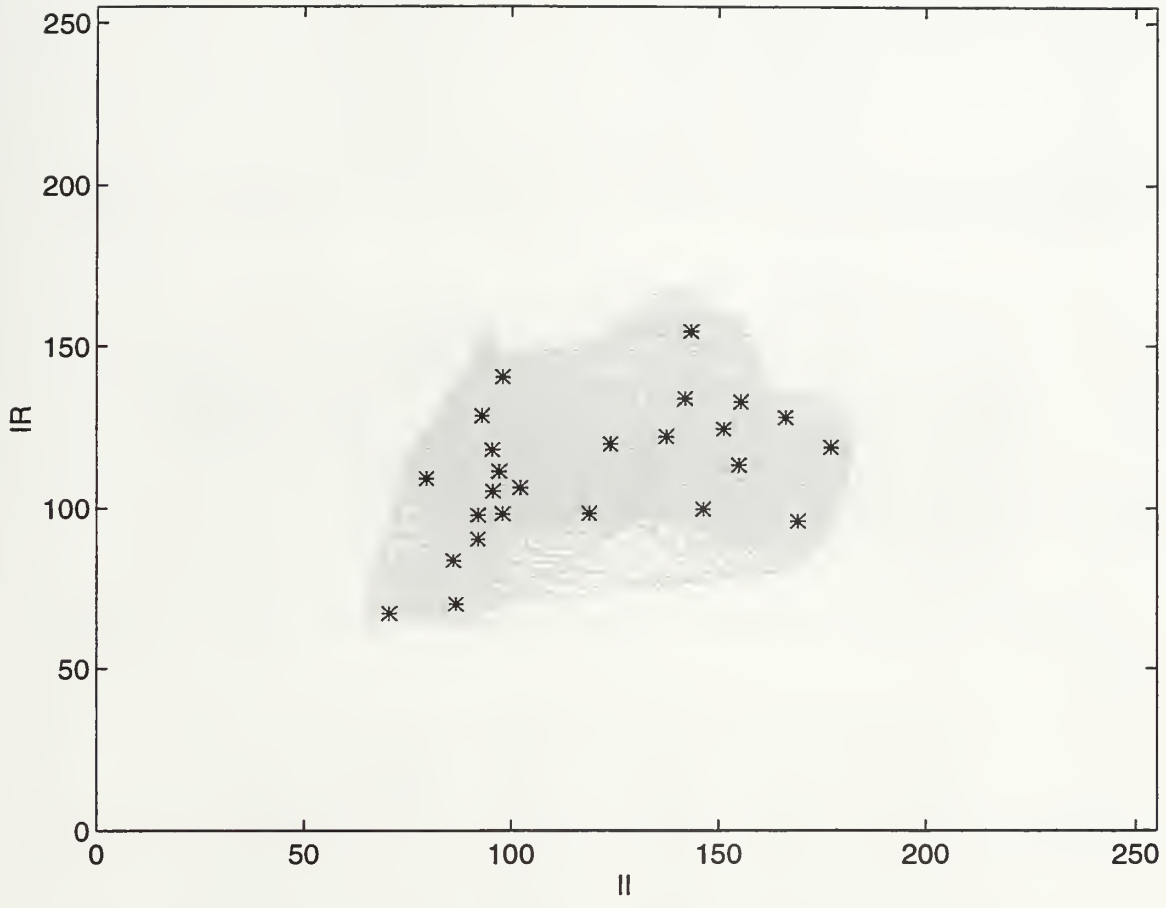

Figure 47. Modified Low-Pass I-R Plane with Associated Centers for Scene 5

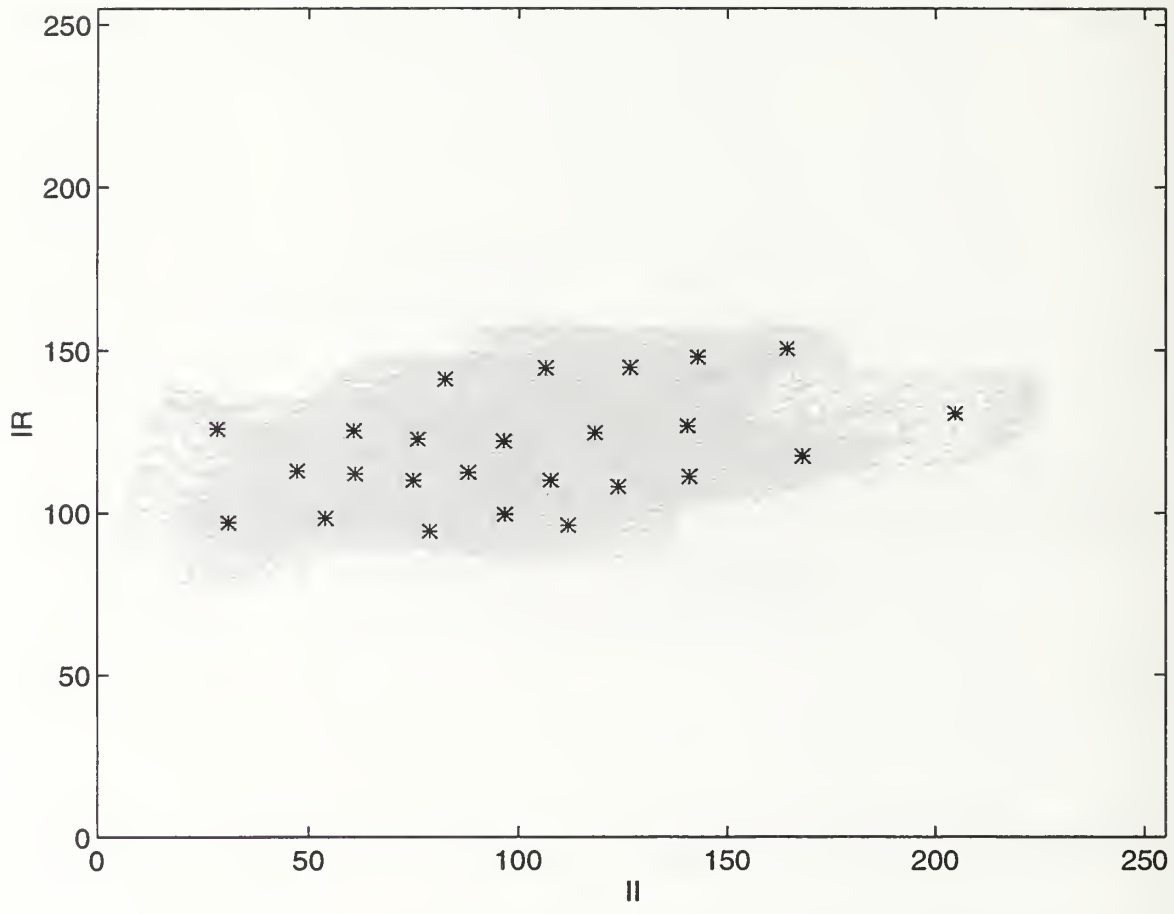

Figure 48. Modified Low-Pass I-R Plane with Associated Centers for Scene 6

APPENDIX E. MAPPING COEFFICIENTS

	a	(1-a)
Scene 1	0.2	0.8
Scene 2	0.7	0.3
Scene 3	0.4	0.6
Scene 4	0.4	0.6
Scene 5	0.4	0.6
Scene 6	0.6	0.4

Table X. Optimization Coefficients for the Weighted Linear Mapping Method

	a1	a2	a3	a4	a5	a6
Scene 1	0.7637	0.0000	0.0000	0.0000	1.9256	-0.0024
Scene 2	0.0000	0.0484	0.0000	0.0175	0.2029	-0.0004
Scene 3	0.0000	0.2231	-0.0004	0.0000	0.0703	-0.0001
Scene 4	0.0000	0.0775	-0.0001	0.0000	0.1895	-0.0004
Scene 5	0.0000	0.2597	-0.0005	1.5927	0.0479	-0.0001
Scene 6	0.0000	0.2026	-0.0004	2.8830	0.0491	-0.0001

Table XI. Optimization Coefficients for the Nonlinear–1 Mapping Method

	a1	a2	a3	a4	a5
Scene 1	0.0000	4.0715	-0.0074	0.4586	0.0000
Scene 2	0.0000	1.2094	0.5827	0.0067	0.0000
Scene 3	0.0000	1.6022	-0.0031	0.0001	0.0049
Scene 4	0.0000	3.2911	-0.0065	0.0000	0.0024
Scene 5	0.0000	13.9587	0.0089	0.0616	0.0000
Scene 6	0.0000	0.0079	126.6983	0.0000	0.0000

Table XII. Optimization Coefficients for the Nonlinear–2 Mapping Method

	a1	a2	a3	a4	a5
Scene 1	0.6700	0.0000	0.0000	0.8587	0.0025
Scene 2	0.0000	0.0051	66.8386	1.0036	-0.0020
Scene 3	5.9179	0.0000	0.0000	0.2197	-0.0002
Scene 4	0.0000	0.0080	0.0072	0.9761	-0.0019
Scene 5	0.1309	0.0034	0.0000	2.0199	-0.0040
Scene 6	0.1509	0.0000	0.0000	6.3589	0.0010

Table XIII. Optimization Coefficients for the Nonlinear–3 Mapping Method

LIST OF REFERENCES

[1] United States Marine Corps Marine Aviation Weapons and Tactics Squadron One, Commanding Officer, Marine Aviation Weapons and Tactics Squadron One, Box 99200, Yuma, AZ 85369-9200, *Tactical jet night vision device manual*, 1st edition, 1994.

[2] Dee Ryan and Richard Tinkler, "Night pilotage assessment of image fusion," In *Helmet- and Head-Mounted Displays and Symbology Design Requirements II*, Volume 2465, pp. 50–67. SPIE–The International Society for Optical Engineering, 1995.

[3] Allen M. Waxman, Alan N. Gove, and Michael C. Seibert, et. al., "Progress on color night vision: Visible/IR fusion, perception & search, and low-light CCD imaging," In *Enhanced and Synthetic Vision 1996*, Volume 2736, pp. 96–107. SPIE–The International Society for Optical Engineering, 1996.

[4] Matthew T. Sampson, "An assessment of the impact of fused monochrome and fused color night vision displays on reaction time and accuracy in target detection," Master's Thesis, Naval Postgraduate School, Monterey, CA, 1996.

[5] Naval Research Laboratory, Optical Sciences Division (Code 5636), Washington D.C. "Nite hawk project: Color fusion of IR and ICCD sensors," Briefing, June 1996.

[6] Hui Li, B.S. Manjunath, and Sanjit K. Mitra, "Multi-sensor image fusion using the wavelet transform," *IEEE*, 0-8186-6950-0:51–55, 1994.

[7] Jae S. Lim, *Two-Dimensional Signal and Image Processing*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1990.

[8] Tamar Peli and Jae S. Lim, "Adaptive filtering for image enhancement," *International Conf. on Acoustics, Speech and Signal Processing*, 1610:1117–1120, 1981.

[9] Major Claude R. Sasso, "Soviet night operations in World War II," *Leavenworth Papers*, 6:ix–x, 1982.

[10] William K. Krebs, "A solid state visible/FLIR color night vision system to improve situational awareness and tactical efficiency," Navy Advanced Technology Demonstration Proposal, 1997 start, 1994.

[11] Allen M. Waxman, David A. Fay, and Alan Gove, et. al., "Color night vision: Fusion of intensified visible and thermal IR imagery," In *Synthetic Vision for*

Vehichle Guidance and Control, Volume 2463, pp. 58–68. SPIE–The International Society for Optical Engineering, 1995.

[12] Department of the Navy, Commander Operational Test and Evaluation Force, Commander Operational Test and Evaluation Force Squadron One, 7970 Diven Street, Norfolk, VA 23505-1498. *Tactics guide TZ-1031-03-93 for A-6E night attack*, 1993.

[13] Dee Ryan, Richard Tinkler, and Howard Creswick, "Assessment of image fusion in a night pilotage system," NATO AC/243 Panel 3/4 Symposium on Mulitsensors and Sensor Fusion, 1993.

[14] Allen M. Waxman, Alan N. Gove, and Michael C. Seibert, et. al., "Color night vision: Opponent processing in the fusion of visible and IR imagery," *Neural Networks Letter*, 10(1):1–6, 1997.

[15] John W. Sammon, Jr., "A nonlinear mapping for data structures analysis," *IEEE Transactions on Computers*, 5:401–409, 1969.

[16] Charles W. Therrien, *Decision, Estimation, and Classification: An Introduction to Pattern Recognition and Related Topics*, John Wiley & Sons, New York, 1989.

[17] The Mathworks Incorporated, *MATLAB© Optimization Toolbox.*

[18] W.K. Krebs, P. Lewis, S. Buttrey, and E. McKenzie, "Pairwise comparison of low-light, infrared, and multi-sensor nighttime imagery," (in preparation), 1997.

[19] W.K. Krebs, M.T. Sampson, J. Ogawa, D. Scribner, and E.A. Essock, "Visual search: comparison between a matched filter and behavioral performance," (in preparation), 1997.

[20] S. Siegel and J. Costellon, *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill, Inc., New York, NY, 1988.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center 2
   8725 John J. Kingman Road., Ste 0944
   Ft. Belvoir, VA 22060-6218

2. Dudley Knox Library 2
   Naval Postgraduate School
   411 Dyer Rd.
   Monterey, CA 93943-5101

3. Chairman, Code EC 1
   Department of Electrical and Computer Engineering
   Naval Postgraduate School
   Monterey, CA 93943-5121

4. Prof. Charles W. Therrien, Code EC/Ti 3
   Department of Electrical and Computer Engineering
   Naval Postgraduate School
   Monterey, CA 93943-5121

5. Prof. William K. Krebs, Code OR/Kw 3
   Department of Operations Research
   Naval Postgraduate School
   Monterey, CA 93943-5000

6. LT James W. Scrofani, USN 4
   Department of the Navy
   Program Management Office
   PMOSSP Detachment
   P.O. Box 157
   Magna, UT 84044-0157