NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer’s or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.
Thermal-Polarimetric and Visible Data Collection for Face Recognition

by Matthew D Thielke and Patrick J Rauss

Sensors and Electron Devices Directorate, ARL

Approved for public release; distribution unlimited.
**4. TITLE AND SUBTITLE**
Thermal-Polarimetric and Visible Data Collection for Face Recognition

**14. ABSTRACT**
One of the less-explored topics in automated face recognition research is exploiting the thermal IR bands along with the visible band. One such application is matching a thermal face image with visible spectrum face images for interoperability with existing biometric face databases and watch lists. One of the challenges to this cross-modal matching is the substantial differences in thermal and visible face signatures due to phenomenology. In this data collection, we explore the use of a relatively new sensor for face recognition applications, a polarimetric IR imager. When processed, the data collected by this imager provide geometric face features that result in better cross-modal recognition performance than conventional IR data. This data collection also explores the effect of several independent factors on face recognition performance: 1) subject-to-sensor distance, 2) face aspect/pose angle, and 3) facial expressions.

**15. SUBJECT TERMS**
face recognition, infrared, IR, polarimetric, cross-modal, thermal IR

<table>
<thead>
<tr>
<th>19a. NAME OF RESPONSIBLE PERSON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matthew D Thielke</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>19b. TELEPHONE NUMBER (Include area code)</th>
</tr>
</thead>
<tbody>
<tr>
<td>301-394-0810</td>
</tr>
</tbody>
</table>
Contents

List of Figures iv

List of Tables iv

1. Introduction 1

2. Sensor Description 1
   2.1 Imaging Polarimeter 1
   2.2 Visible Imagers 4
   2.3 Uncooled Thermal IR Imager 5
   2.4 Sensor Platform 5
   2.5 Sensor Calibration 7

3. Data Acquisition 8
   3.1 Imaging Polarimeter Data Acquisition 8
   3.2 Uncooled Thermal IR Imager Data Acquisition 8
   3.3 Visible Imager Data Acquisition 8

4. Institution Review Board Process 9

5. Experimental Setup 10

6. Experimental Procedure 12

7. Experimental Data 13

8. Conclusion 16

9. References 18

List of Symbols, Abbreviations, and Acronyms 19

Distribution List 20
List of Figures

Fig. 1  VELA LW sensor ..................................................................................2
Fig. 2  Retardance of half-wave (λ/2) retarder ..............................................2
Fig. 3  Polarimetric VELA sensor (red), Basler CCD cameras (white), and FLIR 660 camera (yellow) .................................................................6
Fig. 4  Gibraltar pan-tilt head ........................................................................6
Fig. 5  Mikron M345 blackbody source positioned in front of VELA objective sense .......................................................................................................7
Fig. 6  Subject’s chair with flood lamp and background board ....................10
Fig. 7  Thermal and visible registration makers .............................................11
Fig. 8  Geometry of the “pose” image sequence ...........................................12
Fig. 9  VELA polarimetric S₀ image ...............................................................13
Fig. 10 VELA polarimetric S₁ image .............................................................14
Fig. 11 VELA polarimetric S₂ image .............................................................14
Fig. 12 VELA polarimetric DOLP image .......................................................14
Fig. 13 FLIR SC600 image, 25° FOV ..............................................................15
Fig. 14 Images from the Basler Scout cameras: 5°, 17°, 34°, and 53° FOVs...16

List of Tables

Table 1  Polarimetric data products produced by the VELA LW sensor ..........4
Table 2  Camera configurations .........................................................................4
1. Introduction

Face recognition has been an active area of research for over 2 decades due its wide range of applications. The focus of face recognition has primarily been on visible imagery in the 0.35- to 0.74-μm wavelength range. However, the military commonly employs thermal imaging systems, which operate in the midwave IR band (MWIR; 3–5 μm) and longwave IR band (LWIR; 8–14 μm), for nighttime operations. Therefore, it would be beneficial to the Army to be able to exploit the thermal IR bands along with the visible band for face recognition applications, such as matching a given thermal face image acquired at nighttime with visible light face images in an existing database. To achieve this capability, a database of simultaneously acquired thermal and visible face imagery is needed to support the development of multimodal face recognition algorithms and methods.¹

Previous work suggests that IR face images may be less sensitive to facial expression changes than visible band face images² Other work has focused on improving face recognition performance with respect to different head poses. A relatively new development in face recognition is exploiting polarimetric imaging sensors, which provide information related to the surface geometry of the face in addition to the intensity information. This additional information has been demonstrated to improve cross-modal face recognition over conventional thermal imagery³

2. Sensor Description

The data collection sensors consisted of the instruments discussed in this section and Section 3. They were mounted to a single optical plate and were oriented so that the subject’s face was in the field of view (FOV) for all the sensors for most of the data collection distances and sensors’ FOVs.

2.1 Imaging Polarimeter

The primary sensor for this face recognition data collection was the 640 LWIR multispectral imaging polarimeter developed by Polaris Sensor Technologies, Inc., called the VELA LW sensor (Fig. 1). This instrument has a 6-position filter wheel, which enables it to operate in several thermal IR spectral bands. For this data collection, Filter 1 was employed, which provided a spectral band of 7.5–11.1 μm.
This instrument employs a rotating half-wave (λ/2) retarder. Half-wave retarders are sometimes called polarization rotators. “The half-wave plate can be used to rotate the polarization state of a plane polarized light.” If the incoming light is linearly polarized and a half-wave retarder is rotated by the angle (phi) relative to the incoming light, the output will be linearly polarized at an angle of 2*(phi) relative the incoming light. The ideal half-wave retarder would have a retardance of 180° and would be stable with respect to wavelength, temperature, and angle of incidence. The retarder used in this instrument is a custom cadmium sulfide/cadmium selenide achromatic retarder fabricated by Gooch and Housego, Inc. The measured retardance (in degrees) of this optical device is close to 180° and does not vary greatly over the operating wavelengths, as shown in Fig. 2.
The rotating retarder is located in front of the cooled optics containing the fixed wire-grid polarizer and bandpass filter wheel. The rotating retarder working in conjunction with the wire-grid polarizer results in changes in the intensity levels on the IR focal-plane array (FPA). These intensity changes are a function of the linear polarization states of the IR light radiated from the surfaces of objects in the FOV, such as a human face. Rotating retarder type imaging polarimeters can suffer from pixel- and subpixel-level image misregistration between successive images. It is worst for the sharp edges in the scene. The evidence of this misregistration is artificial edge enhancement. These potential data quality issues did not significantly impact our study.

The VELA LW sensor is a division-of-time spinning retarder system. The measurements by the FPA are taken sequentially in time. Between measurements, the retarder rotates and the angle of the linear polarized IR light that is passed to the FPA changes. “In its principle mode of operation, the system acquires a set of 16 images per rotation of the λ/2 retarder. Images are captured at 0°, 22.5°, 45° … to 337.5°.”

The basic sensor parameters for the VELA LW are as follows:

- Format: 640 × 480
- Waveband: 7.5–11.1 μm
- Pixel size: 15 × 15 μm
- NEΔT: 25 mK at f/2
- Frame rate: 120 Hz
- Dynamic range: 14-bit
- Narrow FOV: 3.3° × 2.5°
- Wide FOV: 10.6° × 7.9°

The polarimetric data products produced by this sensor are shown in Table 1. The Stokes polarization parameters, S₀, S₁, S₂, and S₃, describe the polarization state of the electromagnetic field. S₃ is a measure of the right-hand circular (RHC) polarization minus the left-hand circular (LHC) polarization. For this sensor, S₃ is 0, because the wire grid polarizer passes only linearly polarized light. RHC=LHC=0.
Table 1  Polarimetric data products produced by the VELA LW sensor

<table>
<thead>
<tr>
<th>Image</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_0$ image</td>
<td>Total radiance image, W/cm²-sr</td>
</tr>
<tr>
<td>$S_1$ image</td>
<td>Horizontally polarized radiance minus vertically polarized radiance</td>
</tr>
<tr>
<td>$S_2$ image</td>
<td>45° polarized radiance minus 135° polarized radiance</td>
</tr>
<tr>
<td>$S_1/S_0$ image</td>
<td>$S_1$ normalized radiance image $S_0$</td>
</tr>
<tr>
<td>$S_2/S_0$ image</td>
<td>$S_2$ image normalized by radiance image $S_0$</td>
</tr>
<tr>
<td>DOLP image</td>
<td>Degree of linear polarization image $\sqrt{S_1^2 + S_2^2}/S_0$</td>
</tr>
</tbody>
</table>

### 2.2 Visible Imagers

For the visible face images, we used 4 imagers from the Basler Scout series of progressive-scan charge-coupled device (CCD) cameras. All 4 cameras used the Sony ICX424 AL/AQ sensor, which provides either grayscale (monochrome) or red-green-blue (RGB) Bayer mosaic images. Two models of Basler cameras were used for face image collection, and we employed 2 cameras of each model:

1) **scA640-70gm** (monochrome)
   - Sensor size: 659 × 494 pixel images
   - Optical size: 1/3 inch
   - Pixel size: 7.4 μm × 7.4 μm
   - Maximum frame rate: 70 fps

2) **scA640-70gc** (RGB color)
   - Sensor size: 658 × 492 pixel images
   - Optical size: 1/3 inch
   - Pixel size: 7.4 μm × 7.4 μm
   - Maximum frame rate: 70 fps

Using 4 cameras with virtually identical imaging and data properties allowed us to collect images of the same face at 4 different resolutions simultaneously by using objective lenses with 4 different focal lengths. The configuration of the cameras is shown in Table 2.

<table>
<thead>
<tr>
<th>Camera</th>
<th>Type</th>
<th>Objective FL (mm)</th>
<th>FOV (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>scA640-70gc</td>
<td>RGB color</td>
<td>50</td>
<td>5</td>
</tr>
<tr>
<td>scA640-70gm</td>
<td>Monochrome</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>scA640-70gc</td>
<td>RGB color</td>
<td>8</td>
<td>34</td>
</tr>
<tr>
<td>scA640-70gm</td>
<td>Monochrome</td>
<td>4.5</td>
<td>53</td>
</tr>
</tbody>
</table>
The 50-mm lens provided the greatest number of pixels between the eyes and the 4.5-mm lens the least number of pixels between the eyes for the same face.

2.3 Uncooled Thermal IR Imager

An uncooled thermal imager, the FLIR SC660, provided broadband unpolarized thermal face images for this study. The sensor size in pixels, 640 × 480, was similar to the polarimetric VELA sensor and the Basler Scout CCD cameras for comparison.

The basic specifications for the FLIR SC660 used in this study were as follows:

- FOV: 24° × 18°
- Spatial resolution: 0.65 mRad
- Thermal sensitivity: 30 mK at 30 °C
- IR resolution: 640 × 480 pixels
- Spectral range: 7.5–13 μm
- Analog image output: NTSC analog video
- Digital image output: Firewire radiometric, 14-bit digital video to PC

The analog video was not used for this study. The radiometric, 14-bit digital data provided temperature measurement information for comparison with the polarimetric VELA sensor.

2.4 Sensor Platform

The sensors were mounted to a single optical plate as shown in Fig. 3. The VELA polarimetric IR sensor is indicated by the red rectangle, the 4 Basler CCD visible cameras are indicated by the white rectangle, and the FLIR uncooled IR sensor is indicated by the yellow rectangle. The optical plate was mounted to a heavy duty QuickSet Gibraltar geared pan-tilt head; the Gibraltar tripod as shown in Fig.4. This setup provided a very stable platform for the image collection.
Fig. 3  Polarimetric VELA sensor (red), Basler CCD cameras (white), and FLIR 660 camera (yellow)

Fig. 4  Gibraltar pan-tilt head
2.5 Sensor Calibration

The FLIR SC660 IR camera is factory calibrated for accurate temperature measurements and thermal imaging. The FLIR SC660 used in this study was sent to the factory for an inspection and calibration to maintain its accuracy in data collection. No user calibration procedures were necessary for the FLIR SC660 during the face image collection.

The VELA LW sensor requires a polarimetric calibration and a radiometric non-uniformity correction (NUC). For the polarimetric calibration, a calibration polarizer consisting of a wire grid polarizer on a barium fluoride substrate and 2 blackbody sources are used. During the extent of the face image collection, we used premeasured polarimetric calibration files. No polarimetric calibrations measurements were taken during the test. For the radiometric NUC, we positioned a blackbody source (Fig. 5) at a stable temperature in front of the objective lens. The NUC measurements consist in collecting uniform image frames at several reference temperatures. For the face image collection, we employed a Mikron M345 as our blackbody calibration source and performed a 2-point NUC at 20 °C and then at 40°C.

Fig. 5 Mikron M345 blackbody source positioned in front of VELA objective sense
3. Data Acquisition

3.1 Imaging Polarimeter Data Acquisition

The VELA LW imaging polarimeter had a gigabit Ethernet (GigE) port for data output. To record the data, we used the Grave Control software developed by Polaris Sensor Technologies. We used the software to adjust the focus on each subject and set the following parameters:

- Number of frames to collect: 500
- Number of frames per second: 60
- Rotation rate of the λ/2 retarder: 60 Hz
- Camera lens: wide FOV

3.2 Uncooled Thermal IR Imager Data Acquisition

The FLIR SC660 had an IEEE 1394a (Firewire 400) port for output of image data. To record the data, we used the Research IR software from FLIR, Inc. We generally used the autofocus on the FLIR SC660 camera to set the focus, and then verified the focus using the Research IR software. We used the software to acquire sequence of images for 10 s at 30 fps.

3.3 Visible Imager Data Acquisition

The 4 Basler Scout CCD cameras had GigE ports for data output. To record the data, we used a custom configured Quazar video recording and processing platform from Boulder Imaging, Inc. The Quazar had 4 GigE ports capable of recording 4 data streams simultaneously. We recorded each video stream to a separate internal disk for maximum system throughput. We used the Quazar software and set up files to set the following parameters:

- Elapsed time: 10 s
- Frames per second: 70
- File option: TIFF
- Image type: Bayer RGB (color), mono
- Image size: 640 × 492 (color), 640 × 480 (mono)
- Image bit depth: 8 bits/pixel
4. Institution Review Board Process

Because the face image collection involved human subjects, we needed to obtain approval from the US Army Research Laboratory (ARL) Institutional Review Board (IRB) before collecting face images. To obtain the approval, the investigators needed to do the following:

1) Take an online course from the Collaborative Institutional Training Initiative (CITI) and pass the online test. In addition, the investigators needed to take periodic refresher courses and pass additional online tests.

2) Submit the research protocol. This protocol provides an abstract, the location of the research, the data collection dates, background information, and the research objective. In addition, there is a description of the equipment and facilities. The protocol needs to provide the number subjects planned, a justification for the sample size, whether there is compensation for subjects, and the methods of subject recruitment. The protocol must include the data collection procedures, the experimental design, and the planned data analysis. An important part of the protocol is the identification of risks and discomforts to the subjects, and the plan to mitigate the risks and discomforts. The protocol must address confidentiality issues and the methods we will use to maintain the privacy of the subjects.

3) Submit a copy of the consent form that the subjects of the study would be required to sign.

4) Submit technical reviews of the research protocol by at least 3 “individuals qualified by training and experience to evaluate the validity of the project”.

The IRB provided feedback to the investigators. The feedback resulted in additional work on the protocol and consent form before the IRB granted approval.

Our face data collection did not fit into the usual parameters of human subject research that was familiar to the IRB and the Human Research and Engineering Directorate of ARL. The sample files and templates provided by the IRB generally presumed that humans subjects would be asked to perform tasks while subjected to various stimuli and it was the subject’s performance that being tested or evaluated. We did ask the subjects to perform a few tasks, but these were to cause the subjects to vary their facial pose or expression. We were not evaluating the subject’s performance. For this reason, there was “learning curve” for the face recognition researchers and the IRB. Our face data collection also not fit into the standard practices covered in the CITI training. The CITI training primarily targeted social and behavioral research investigators.
5. Experimental Setup

For the initial phase of our study, we collected the face images indoors. The sensors were mounted as shown in Fig. 4. For the duration of the data collection, the sensor tripod was stationary and the subjects moved to provide different distances to the sensors. The sensor platform could pan in azimuth and tilt in elevation several degrees to adjust for the different heights of various subjects.

The subjects were seated a nonrotating chair (shown Fig. 3). Directly behind the chair was rectangular piece of “sound-soak” wall material to serve as uniform background in the visible and IR. Located 2 m in front of the chair was a flood lamp (Fig. 6) with a compact fluorescent bulb. For our study, we used Ecosmart compact fluorescent “daylight” 23-W bulb, with a brightness of 1,550 lumens and a color temperature of 5,000 K.

![Fig. 6 Subject’s chair with flood lamp and background board](image)
We used the flood lamp in order to increase illumination on the face for the visible sensors and minimize differences in facial illumination that occurs when the chair is moved to different distances from the sensors and the overhead or surrounding lighting changes.

Attached to the background board were positional markers for the visible and IR sensors (Fig. 7). These are used for registering the thermal and visible images. These markers were constructed of common cellulose sponge material. When the sponges were filled with water, they would appear as persistent “dark” spots in the thermal images and were also highly visible to the other cameras.

The setup of the flood lamp, chair, and background board were moved together in order to change the distance from the sensors to the seated subject. The distances used for our study were as follow:

- R5: 12.5 m
- R4: 10.0 m
- R3: 7.5 m
- R2: 5.0 m
- R1: 2.5 m
6. Experimental Procedure

For each subject, we followed this experimental procedure:

1) The subjects arrived at the image collection site at a scheduled time.

2) They were given an overview of the study and allowed to read the consent form and the face image release form. Subjects were not asked to disclose age, gender, or other demographic information.

3) If the subjects signed the consent form and the face image release form, they were shown to chair at the first marked distance.

4) If they were wearing glasses, they were asked to remove them, because glasses are opaque in the IR band. Subjects were not asked to remove makeup or sunscreen.

5) The subjects were asked to face forward and sit still while images were being collected by the sensors. This was referred to as the “baseline” image sequence.

6) We asked the subjects to turn their heads from far left to far right over a span of 8 s while the sensors were collecting data. This was referred to as the “pose” image sequence. This sequence contained multiple face pose angles from approximately −90° to +90°, as shown in Fig. 8.

7) The subjects were then asked to face the cameras and count numbers out loud starting from “1, 2, 3,...” for 10 s while the sensors were collecting data. This sequence was referred to as the “expression” image sequence. Requiring the subject to count aloud resulted in different facial expressions on the subject’s face while speaking the different numbers.

Fig. 8  Geometry of the “pose” image sequence

Approved for public release; distribution unlimited.
8) We would then move the chair and associated equipment to the next marked distance. We would ask the subjects to repeat the “baseline”, “pose”, and “expression” sequence at the new distance. We would continue this process until we had collected image data at 5 marked distances: 2.5, 5, 7.5, 10, and 12.5 m.

7. Experimental Data

We collected face data on 50 subjects during the period from August to December 2014. These subjects were employees or contractors of the US Army Garrison Adelphi Laboratory Center, Adelphi, Maryland. There were 38 males and 12 females. We did not record the age, race, or ethnicity of the subjects, but all were at least 18 years of age or greater. Several researchers in the field of cross-spectral face recognition have used 50 subjects for their studies. Fifty subjects were used for a visible and MWIR study. The West Virginia University Multi-spectral database consists of visible and shortwave IR face images of 50 subjects. In 2012, researchers at Night Vision and Electronic Sensors Directorate, in their study “Thermal to Visible Facial Recognition Data Collection,” used 50 subjects.

Figures 9 through 12 show the image data produced by the VELA polarimetric IR hardware and the data collection software.

![Fig. 9 VELA polarimetric S0 image](image)
Figure 9, the VELA $S_0$ “total reflectance” image, closely resembles the images produced by the broadband unpolarized FLIR SC660 as would be expected. The VELA $S_0$ image generally showed greater thermal detail than the FLIR SC660.
given that the sensor had a better thermal sensitivity and a smaller FOV. Figures 10 through 12, the $S_1$, $S_2$, and DOLP images, show more shading and stronger edge information related the surface geometry of the face than traditional thermal images.

Figure 13 shows an image produced by the uncooled thermal IR imager, the FLIR SC660. The thermal markers shown in Fig. 7 are plainly visible as 4 dark spots on the board behind the subject.

![Image](image.jpg)

**Fig. 13** FLIR SC600 image, 25° FOV

Figure 14 show the images produced by the Basler Scout “visible light” cameras. These imagers had 5°, 17°, 34°, and 53° FOVs, respectively. Opto-mechanical issues with the 8- and 4.5-mm objective lenses prevented the imagers with 34° and 53° FOVs from forming completely in-focus images.
8. Conclusion

The face images collected by the polarimetric thermal IR imager and visible light cameras have already found useful applications.

Alex J Yuffa, Kristan P Gurton, and Gorden Videen, of ARL’s Computational and Information Science Directorate have used the Stokes images $S_0$, $S_1$, $S_2$, and DOLP image to generate 3-D face images. They used the Fresnel relations to determine the surface normal at each pixel and integrated over the surface normals to obtain a 3-D face image.$^{10}$

Shuowen Hu of ARL and Nathaniel Short of Booz Allen Hamilton used the polarimetric and visible faces images obtained from the 50 subjects to improve the thermal-to-visible, cross-modal face recognition accuracy by more than 20% versus using conventional thermal imagery alone.$^{11}$

A subset of the collection database is now releasable to researchers outside of ARL for face recognition research. There are relatively few thermal polarimetric and visible face image datasets currently available to researchers. This Multi-Spectrum Face Dataset is intended to promote research on polarimetric and multi-spectrum face recognition. We hope that this database will enable the development of
improved thermal-to-visible face recognition. We are currently announcing the data set in various face recognition research forums.
9. References


List of Symbols, Abbreviations, and Acronyms

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-D</td>
<td>3-dimensional</td>
</tr>
<tr>
<td>ARL</td>
<td>US Army Research Laboratory</td>
</tr>
<tr>
<td>CCD</td>
<td>charge-coupled device</td>
</tr>
<tr>
<td>CITI</td>
<td>Collaborative Institutional Training Initiative</td>
</tr>
<tr>
<td>DOLP</td>
<td>degree of linear polarization</td>
</tr>
<tr>
<td>FOV</td>
<td>field of view</td>
</tr>
<tr>
<td>FPA</td>
<td>focal-plane array</td>
</tr>
<tr>
<td>GigE</td>
<td>gigabit Ethernet</td>
</tr>
<tr>
<td>IR</td>
<td>infrared</td>
</tr>
<tr>
<td>IRB</td>
<td>Institutional Review Board</td>
</tr>
<tr>
<td>LHC</td>
<td>left-hand circular</td>
</tr>
<tr>
<td>LWIR</td>
<td>longwave infrared</td>
</tr>
<tr>
<td>MWIR</td>
<td>midwave infrared</td>
</tr>
<tr>
<td>NUC</td>
<td>non-uniformity correction</td>
</tr>
<tr>
<td>PC</td>
<td>personal computer</td>
</tr>
<tr>
<td>RGB</td>
<td>red-green-blue</td>
</tr>
<tr>
<td>RHC</td>
<td>right-hand circular</td>
</tr>
</tbody>
</table>