Face Detection in the Near-IR Spectrum

J. B. Dowdall[a], I. Pavlidis[a], G. Bebis[b]
[a]Dept. of Computer Science., University of Houston, Houston, TX 77204
[b]Dept. of Computer Science, University of Nevada, Reno, NV 89557

ABSTRACT

Face detection is an important prerequisite step for successful face recognition. Face detection methods reported in the literature are far from perfect and deteriorate ungracefully where lighting conditions cannot be controlled. We propose a method that could potentially outperform state-of-the-art face detection methods in environments with dynamic lighting conditions. The approach capitalizes upon our near-IR skin and face detection methods reported elsewhere [11][12][13]. It ascertains the existence of a face within a skin region by finding the eyes and eyebrows. The eye-eyebrow pairs are determined by extracting appropriate features from multiple near-IR bands. In this paper we introduce a novel feature extraction method we call *dynamic integral projection*. The method is relatively simple but highly effective because the processing is constrained within the skin region and aided by the near-IR phenomenology.

Keywords: face detection, near infrared imaging, integral projection, and feature extraction.

1. INTRODUCTION

Face detection and recognition have been active research areas for more than thirty years. Face detection is an important preprocessing stage of a face recognition system. Although, it may appear rudimentary to a layman, face detection is a challenging machine vision operation, particularly in outdoor or semi-outdoor environments where illumination varies greatly. This is one of the primary reasons that face recognition is currently constrained to access control applications in indoor settings.

In the current paper we present a novel face detection system based on near-IR phenomenology, and multi-band feature extraction. Facial signatures are less variable in near-IR aiding significantly the detection work. Illumination in the scene can be maintained at an optimal level through a feedback control loop that adjusts a near-IR illuminator. Since, near-IR light is invisible to the human eye the system can remain unobtrusive and covert. The aforementioned advantages in combination with the unique reflectance characteristics of the human skin in the near-IR spectrum allow for simple algorithmic-based face detection methods to perform extremely well.

In recent years a sizable body of research in the area of face detection has been amassed. An excellent survey of the relevant literature can be found in [1]. The majority of face detection research aims to find structural features that exist even when the pose and viewpoint vary. The existence of such features is associated with the existence of faces in the image. Feature extraction methods utilize various properties of the face and skin to isolate and extract desired data. Popular methods include skin color segmentation [2][3], principal component analysis [4][5], eigenspace modeling [6], histogram analysis [7], texture analysis [8], and frequency domain features [9].

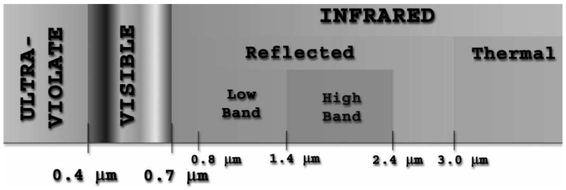

Figure 1: The EM spectrum.

All of the above approaches are associated with visible spectrum imagery. Therefore, they are susceptible to light changes [10] and the variability of human facial appearance in the visible band. A distinct line of research pursued by our group proposed the fusion of two near-IR bands for the detection of face and other exposed skin areas of the body [11][12]. The method capitalizes upon some unique properties of the human skin in the near-IR spectrum. Our tri-band system maintains an optimal illumination in the scene through the liberal use of artificial non-distracting near-IR lights. As a result, the system performs superb skin detection both in indoor and outdoor settings. In [13] we reported further algorithmic work that located the face within the detected skin region. In the present paper we introduce a modification to our original face detection algorithm. We call the new method *dynamic integral projection*.

The rest of the paper is organized as follows: In Section 2 we give a top-level description of the hardware and software architecture of our face detection system. In Section 3 we elaborate on our face detection method, which builds upon our skin detection method. Finally, in Section 4 we conclude the paper and present our plans for future work.

2.System Overview

The latest version of our face detection system uses three cameras as the input medium. Two of the cameras have Indium Gallium Arsenide Focal Plane Arrays (FPA), which are sensitive to a portion of the near-IR spectrum in the range 0.9-1.7 μm. This range clearly falls within the reflected portion of the infrared spectrum and has no association with thermal emissions (see Figure 1). The third camera is a color visible band camera. A system of beam splitters (see Figure 2) allows all three cameras to view the scene from the same vantage point, yet in different sub-bands. The splitters divide the light reflected from the scene into the visible band beam (0.3-0.6 μm), the lower band beam (0.8-1.4 μm), and the upper band beam (1.4-2.4 μm). The three beams are funneled to the FPAs of the corresponding cameras. Each camera is connected to a frame grabber, which digitizes the incoming video.

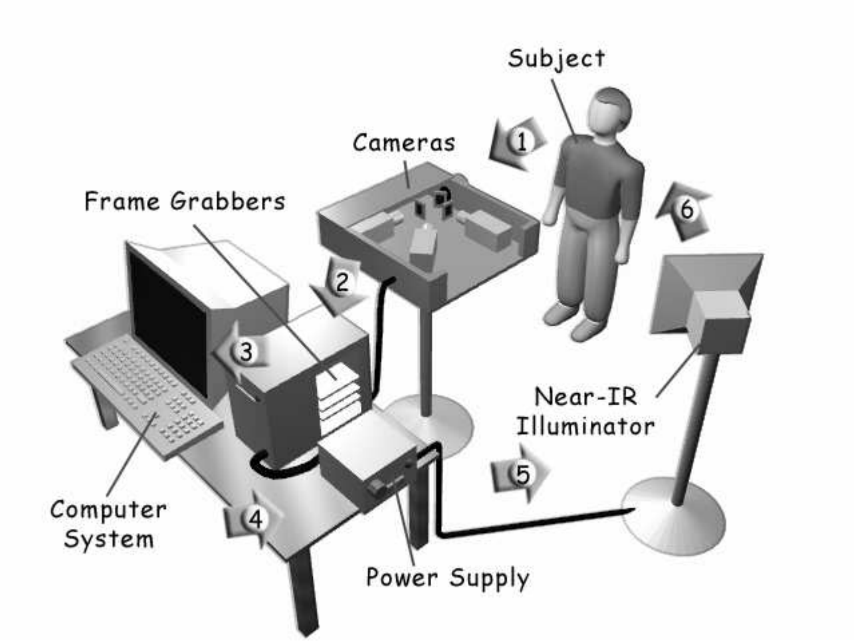

Figure 2: Harware diagram of the tri-band system.

Although we have designed and implemented a tri-band system we use only the two near-IR bands in our approach. At the moment, we use the visible band only for comparative testing purposes with other face detection and recognition software. One of the main benefits of using the near-IR spectrum is that subjects in the scene are unaware that they are being illuminated by the system. This is especially beneficial for covert operation in surveillance applications.

2.1 Software Architecture

The system's software consists of six modules (see Figure 3):

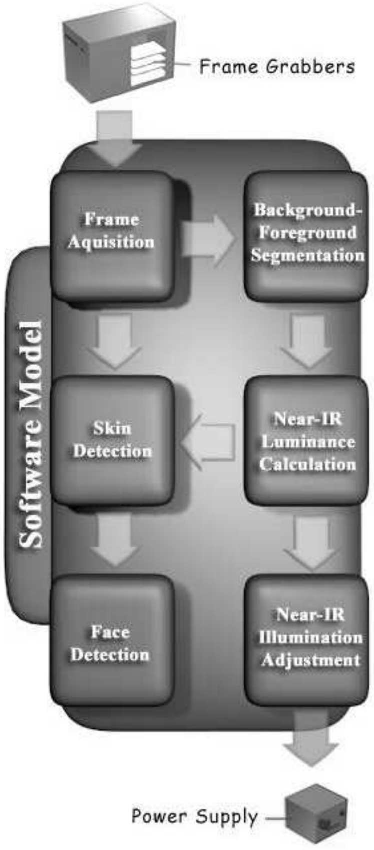

Figure 3: Software diagram of the tri-band system.

- **Frame Acquisition:** Initially the system gets the input frames for all three bands from the respective frame grabbers. The near-IR frames are sent to: a) the Background-Foreground Segmentation and b) the Skin Detection modules. The visible-band frame is made available to the Identix face detector and recognizer [5].
- **Foreground-Background Segmentation:** The foreground-background segmentation is performed based on frame differencing. The binarized along with the original frames are sent to the Near-IR Luminance Calculation module.
- **Near-IR Luminance Calculation:** This module calculates the luminance levels present in the lower and upper near-IR bands. The calculation takes into account the background portions of the frames only.

- **Near-IR Illumination Adjustment:** Based on the computed luminance levels the system adjusts the output on the power supply. The objective is to maintain a constant near-IR luminance level by appropriately adjusting the power of the illuminator in response to environmental changes.
- **Skin Detection:** Upon receiving the two near-IR frames the skin detector performs a series of operations to isolate the skin [11][12]. The output of the skin detection module is a binary image where all skin appears black against a white background. The skin image along with the original near-IR frames is then passed to the Face Detection module.
- **Face Detection:** The face detector uses correlated multi-band dynamic integral projections to detect the existence and location of eyes within the skin region. Eventually, if at least one eye is detected the skin region is declared a facial region.

3.Face Detection Methodology

The system tries to find the facial features within the skin region using the *dynamic horizontal integral projections* of the skin region in the lower and the upper band near-IR images. Using integral projections for facial feature detection is not a new idea [14][15]. The novelty of our approach lies in taking integral projection one step further and adapting it to the dynamic extraction of features. One of the key elements of our approach lies in correlating the information extracted from the lower and upper near-IR bands to improve the robustness of feature extraction (see Figure 4). In particular, eyebrows show up very nicely in the upper near-IR band because human hair is highly reflective in this band and contrasts with the highly non-reflective skin. Eyes show up better in the lower near-IR band because they are non-reflective in this band and contrast with the highly reflective skin.

3.1 Integral Projections

Horizontal (and vertical) integral projections (or profiles) have been used in association with visible band imaging for facial feature extraction [14][15]. Assuming that the search region is a HxW rectangle, the horizontal integral projection can be computed as follows:

$$P(i) = \sum_{j=1}^{W} I(i, j), ..0 \le i \le H \ , \qquad (1)$$

where $I(i,j)$ is the intensity function of our search window. Locating the facial features is then equivalent to finding certain local minima and maxima in $P(i)$. This method works only when the face is facing fairly forward and is unobstructed.

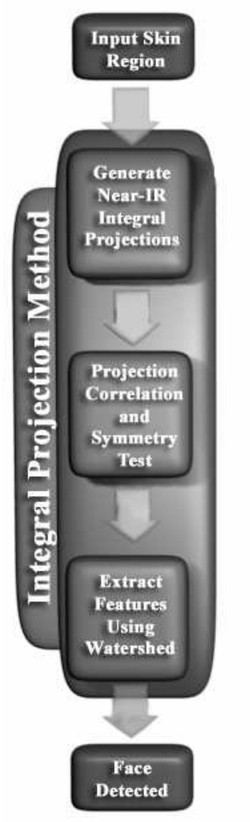

Figure 4: Outline of the integral projection steps.

Figure 5: An example of the integral projection in the visible band. **(a)** The visible band image**. (b)** The gray scale version of the visible band image with its integral projection overlaid in red. The dark stripe in the background creates a significant valley that would make eye detection very hard.

There are two main difficulties with using integral projections in the visible spectrum. First, it requires that the skin region has been extracted, a non-trivial task in the visible spectrum. Without this assumption, it would be quite difficult to locate accurately the correct minima or maxima due to noise introduced by non-trivial backgrounds (see Figure 5). Second, even moderate illumination changes can affect the shape of the integral projection significantly.

Within the context of our method the background noise is not an issue, since we apply the integral projection on the segmented skin regions only. The feedback control mechanism that maintains constant scene luminance further facilitates the effectiveness of integral projection. The effectiveness of the integral projections is also aided by the facial phenomenology in near-IR. In the lower near-IR band, the eyes appear dark while the skin is light. This creates a consistent relative minimum where the eyes are located in the horizontal integral projection of the skin region (see Figure 6(a)). In the upper near-IR band, the eyebrows appear light while the skin is dark. This creates a consistent relative maximum in the horizontal profile where the eyebrows are located (see Figure 6(b)). By correlating the minima in the lower band with the maxima in the upper band, we can find the eye-eyebrow pair more robustly than carrying out the detection in the visible spectrum (see Figure 6(c)).

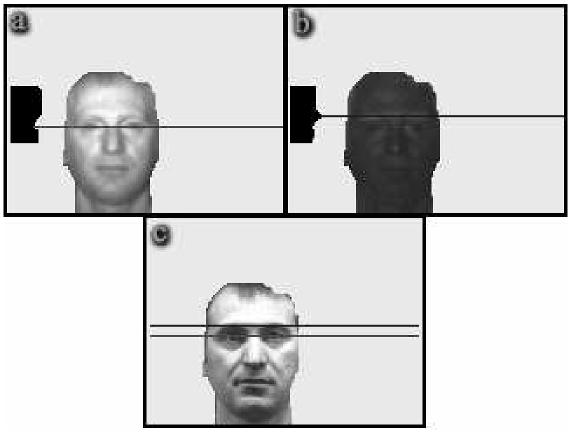

Figure 6: **(a)** The integral projection of the lower band skin region with the min location in red. **(b)** The integral projection of the upper band skin region with the max location in blue. **(c)** Both the min and the max locations overlaid on the visible band image for visualization purposes.

3.2 Dynamic Integral Projection

One of the key obstacles that we encountered in our previous work [13] was the poor performance of the system when it encountered rotated faces. The undesirable results for these cases were in large part due to the method's assumption that both eyes are located on the same horizontal axis. This limitation was the primary impetus for reworking the standard integral method into a dynamic approach that could be used to find each eye independently.

Figure 7: (a) The integral projection of a cross section in the lower band. The horizontal red line denotes the min location. **(b)** All of the lower band cross sections containing a substantive feature. **(c)** The grouping of adjacent substantive features into aggregate substantive features.

In the new method integral projection is performed not once but multiple times in a dynamic process on selected areas of the skin region (see Figure 7). The selected areas used are vertical cross sections of the skin region in question each composed of one or more columns of skin pixels. The pixel width of these cross sections (W_c) is determined to be the equivalent of a third of an eye's width ($W_e=3xW_c$). Given that on average the width of the human face (W_f) is four eye widths, we divide the skin region into twelve equal vertical cross sections ($W_f = 4xW_e = 12xW_c$) to be used for integral projection.

Another key modification to the previous methodology is the addition of a dynamic definition of substantive feature. A *substantive feature* is defined as a portion of an integral projection that conforms to desired phenomenological manifestations (i.e., either a potential eye or eyebrow). Substantive features are determined based on the angles formed between the extremum and its corresponding bounds in the projections of cross sections (see

Figure 8). The bounds of the formations are defined by a 1-dimensional watershed transformation [16]. Specifically, if $\min(\theta_1, \theta_2) > \theta_{th}^{k}$ (2) , where θ_{th}^{k} is the threshold angular value, then a substantive feature is declared for the cross section. A cross section can have at most one substantive feature. Once all the cross sections are processed from left to right (*epoch*), adjoining substantive features are joined together to form aggregate substantive features. If at least two aggregate substantive features are being detected the feature extraction process terminates. Otherwise, a new threshold value θ_{th}^{k+1} is being established, such that $\theta_{th}^{k+1} < \theta_{th}^{k}$ and the Formula (2) is being reapplied in a new epoch. The iteration of epochs continues by successively relaxing the angular threshold until the termination criterion is met.

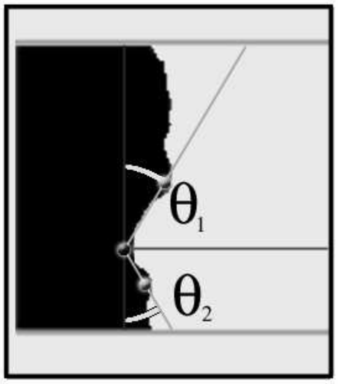

Figure8: The black region is an integral projection of a cross section with the min location by the blue dot and the red line. The two orange dots are placed at the bounds of the concave formation. The corresponding angles θ_1 and θ_2, illustrated by the yellow arcs, are used to determine whether the formation is a substantive feature.

Once two or more aggregate features are formed on each band the feature extraction phase is followed by the feature verification phase. In feature verification, the aggregate features in the lower band (potential eyes) are matched with the aggregate features in the upper band (potential eyebrows). As a result of this matching only the eye-eyebrow pairs that conform to anthropometrics are retained. The rest are discarded as spurious features.

4.Conclusions and Future Work

We have expanded the face detection work reported earlier by our group [13] by developing a face detection method that is robust to rotation, occlusion, and a wide diversity of subjects. The system capitalizes on the observed phenomenology of the near-IR and uses correlated dynamic multi-band integral projections to detect the eyes and the eyebrows. This method shows great promise for overcoming many of the current obstacles facing contemporary face detection methods (see Figures 9, 10, and 11).

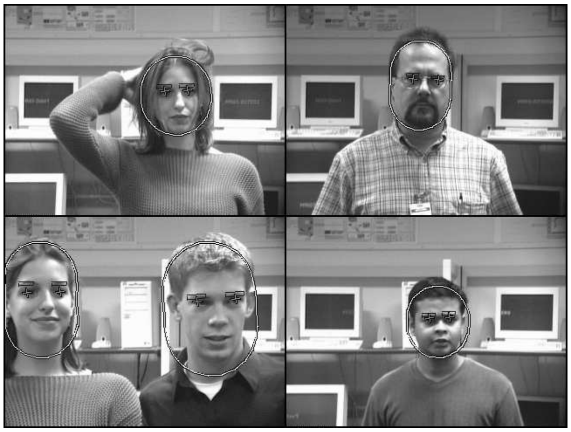

Figure 9: Examples of the system's performance using frontal faces. The superimposed crosses indicate the locations of the eyes. The detection results have been overlaid onto the visible band image for visualization purposes only.

Figure10: Examples of the system's performance using rotated faces. The superimposed crosses indicate the locations of the eyes. The detection results have been overlaid onto the visible band image for visualization purposes only.

In our future work we plan to test the system out on a statistically significant data set. These results would then be compared to an existing commercial system such as Identix operating on the same data. In our previous work we conducted such a comparison and our system outperformed Identix [13]. Preliminary results indicate that the new method discussed in this paper exhibits significantly superior performance over our previously documented approach not only for rotated subjects, but also for frontal facing subjects. Other interesting questions related to the project include determining whether the reflectance properties of the skin in the near-IR band fluctuate due to moisture, exertion, or other external factors such as sunburn. To improve the performance of our system, we plan to model the probability distribution of the features using more powerful models (e.g., mixtures of Gaussians).

Figure 11: Example of our system detecting the driver of a car in an outdoor environment (a) Low near-IR image with the eye positions overlaid in green, (b) High near-IR image with the eye positions overlaid in green.

Acknowledgements

The dynamic integral projection algorithm was developed at the University of Houston. The skin detection method was developed at Honeywell Laboratories. We would like to thank Mr. Jeff Radke and Murray Cooper from the management team of Honeywell Laboratories for their financial support. We would also like to thank Mr. Pete Reutiman, Alan Greisbach, and Justin Droessler for their valuable technical contributions. Part of this project was also supported by an NSF grant (NSF/CRCD 0088086) through the University of Nevada at Reno. The views expressed in this article reflect the opinions of the authors only and should not be linked in any way to the funding institutions.

References

1. M.-H. Yang, D.J. Kriegman, and N. Ahuja, Detecting Faces in Images: A Survey, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 1, pp. 34-58, 2002.

2. S. Kawato and J. Ohya, "Two-step Approach for Real-time Eye Tracking with a New Filtering Technique," in *Proceedings 2000 IEEE International Conference on Systems, Man, and Cybernetics*, 2000, Vol. 2, pp. 1366 –1371.

3. S.H. Kim, and H.G. Kim, "Face Detection Using Multi-modal Information," in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, 2000, pp. 14–19.

4. C. Morimoto, and M. Flickner, "Real-Time Multiple Face Detection Using Active Illumination," in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, 2000, pp. 8-13.

5. http://www.faceit.com

6. Y. Li, S. Gong, S. Liddel, and H. Liddel, "Multi-view Face Detection Using Support Vector Machines and Eigenspace Modeling," in *Proceedings Fourth International Conference on Knowledge-Based Intelligent Engineering Systems & Allied Technologies*, 2000, Vol. 1, pp. 241-244.

7. X. Lv, J. Zhou, and C. Zhang , "A Novel Algorithm for Rotated Human Face Detection," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, 2000, Vol. 1, pp. 760–765.

8. W. Huang, and R. Mariani, "Face Detection and Precise Eyes Location," in *Proceedings 15th International Conference on Pattern Recognition*, 2000, Vol. 4, pp. 722–727.

9. B.H. Jeon, S.U. Lee, and K.M. Lee, "Rotation Invariant Face Detection Using a Model-Based Clustering Algorithm," in *Proceedings 2000 IEEE International Conference on Multimedia and Expo, 2000*, Vol. 2, pp. 1149-1152.

10. J. Wilder, P. Phillips, C.Jiang, and S. Wiener, "Comparison of Visible and Infra-Red Imagery for Face Recognition," *Proceedings Second IEEE International Conference on Automatic Face and Gesture Recognition*, 1996, pp. 182–187.

11. I. Pavlidis, and P. Symosek, "The Imaging Issue in an Automatic Face/Disguise Detection System," in *Proceedings IEEE Workshop on Computer Vision beyond the Visible Spectrum: Methods and Applications*, 2000, pp. 15–24.

12. Pavlidis, V. Morellas, and N. Papanikolopoulos, "A Vehicle Occupant Counting System Based on Near-Infrared Phenomenology and Fuzzy Neural Classification," *IEEE Transactions on Intelligent Transportation Systems*, Vol. 1, No. 2, pp. 72-85, June 2000.

13. J. Dowdall, I. Pavlidis, G. Bebis, "A Face Detection Method Based on Multi-Band Feature Extraction in the Near-IR Spectrum," Proceedings 2001 IEEE Workshop in Computer Vision Beyond the Visible Spectrum: Methods and Applications, Kauai, HI December 14, 2001.

14. G. Bebis, S. Uthiram, and M. Georgiopoulos, "Face Detection and Verification Using Genetic Search", *International Journal of Artificial Intelligence Tools,* vol. 9, no. 2, pp. 225-246, 2000.

15. T. Kanade, "Picture processing by computer complex and recognition of human faces", *Technical Report,* Kyoto University, Dept of Information Sciences, 1973.

16. K. Sobottka and I. Pitas, "A novel method for automatic segmentation, facial feature extraction, and tracking", *Signal Processing: Image Communication,* vol. 12, pp. 263-281, 1998.