Illumination Invariant Face Detection with Time-Dependent Intrinsic Images under Near-IR lighting

http://www.cats.fiu.edu

1. Abstract

Face recognition is broadly used in a variety of law enforcement and security applications. Due to its non-intrusive characteristics, extensive research is performed to optimize the face detection and recognition accuracy. Although most face identification systems are designed to operate indoors, today’s face detection algorithms are still compromised in precision when changes in illumination conditions occur. In this paper, we present a novel solution for achieving illumination invariant face recognition that uses multiple cues to detect and locate faces accurately in near infrared (NIR) imaging systems.

2. Framework Overview

1. We recover the reflectance and illumination intrinsic images from the original image using intrinsic images with NIR image acquisition process.
2. We provide a methodology for object segmentation using a graph-based energy minimization technique to extract the bright pupil regions generated from the NIR light.
3. We apply a simplified radial symmetry (RS) transform method to accurately localize pupils’ positions.

3. Methods

3.1 Time-dependent Intrinsic Images Derivation:

A set of images I can be described as a composition of a static reflectance values R and varying of illumination field L as referred in [6]:

\[ h = L - R \]  (1)

We follow a similar approach to decompose a given set of NIR face images \( I(x, y) \) into a sequence set of illumination images and a static reflectance image representing the reflectance properties of the face as \( R(x, y) \) where \( L(x, y) \) represents the face incident light and the reflectance image as well as the dynamic constituent from the illumination variations change with time. The algorithm below is used to extract those components from a sequence of input images.

Figure 1 (a,b,c) illustrating the resulting image after applying the Intrinsic Images derivation method.

Algorithm 1. Time-dependent Intrinsic Images derivation

Require: Image sets are defined as time-varying 2-D images, yielding \( I(x, y) \)
Ensure: Image sets \( I(x, y) \) are processed in log-domain, \( \log(I(x, y)) \)
Require: Derivative filters \( f_t, f_{t-1} \) and \( h = \sum f_t \)
Establish: \( \mu_L \) a median filter \( h = \sum f_t \), where * represents the convolution operator
Require: reversed derivative filters \( f_{r} \)
Compute: \( g = \sum f_{r} \cdot \mu_L \)
Compute: \( h = \sum f_{r} \cdot \mu_L \)
Ensure: \( I(x, y) = \log(I(x, y)) \)
Ensure: \( L(x, y) = \exp(h) \)

3.2 Pupil Localization

A segmentation method based on intensity information is used to facilitate pupils’ localization. Since the pupil’s reflectance shape differs from one person to another, an energy minimization method is used to extract the pupil shape from the face image. We utilize the graph cut object segmentation technique to cast the energy-based objective in a structure of which the minimum cut corresponds to a globally optimal segmentation of the face. The graph cut method combines the two major approaches for image segmentation: segmentation based on objects’ gray level intensities and segmentation based on contrast in different regions in the image.

\[ E(p) = C(p) + \gamma C(p) \]  (2)

Where \( C(p) \) represents the likelihood of the object region present at pixel \( p \), \( C(p) \) represents the likelihood background region presence of the pixel \( p \), and \( \gamma \) is used to control the segmentation boundary for relative region influence.

We construct the graph representing this energy by first considering each pixel to be a graph node before it is connected with two nodes labeling the pixel as being either in the foreground or in the background level. Each pixel is connected to both either the foreground or background nodes with non-negative edge weights \( dx \) for the gray-valued pixels and \( cy \) representing the contrast between each neighboring pixels \( p \) and \( q \). Such that

\[ dx = -P(\log(x|y|) \)  (3)
\[ cy = -P(\log(x|y|) \)  (4)

Finally, each connected pair combination of neighboring pixels \( (p, q) \) with non-negative edge weight determined by the boundary discontinuity \( B(p, q) \) is determined to find the shortest possible segmentation border that gives the smallest sum along the contrast.

\[ B(p, q) = \frac{1}{\pi} \log \left( \frac{1}{1 - \rho(p, q)} \right) \]  (5)

Where \( \rho \) is the gray-value of the pixel \( p \) \( \in \Omega \), and \( || p - q || \) is the Euclidean pixel distance for normalizing edges of different length.

4. Negative Radial Symmetry

After extracting eye regions using eq. 5, a precise pupil localization algorithm is used to locate eye pupils on the eye windows. Using the pupils’ round shape property, the transform aims to detect pupil’s radial centers using both higher and lower intensities. In order to achieve a higher accuracy, we use the negative radial symmetry (NRS) as introduced in [7]. Based on a 3x3 Sobel edge detector, NRS detects the radial symmetrical centers of the pupils with lower intensity.

5. Results

The proposed method was tested on the CBSR NIR face dataset. 23 different test image sets were considered with a sequence length of 10 images (640 x 480) for each subject. The image sets feature a variety of illumination, background and face size under real-world conditions which is more challenging than images with uniform illumination. The faces were successfully detected based on eyes localization. The eye location rate achieved was 95.45%.

An illustrative example of eye detection results are shown in Figure 1.

6. Conclusion

An efficient method for detecting faces based on NIR illumination is presented. Eye pairs are localized using the structure of the eye regions. Facial reflectance textures were extracted using the intrinsic image decomposition method to eliminate the illumination variations. Furthermore, a graph-based method is used to segment the eye pupil shape prior to localizing the segmented pupils using an eye variance filter on different eye windows. The method was tested on different subjects achieving a detection rate exceeding 95% accuracy. However subject with extrinsic objects such as eye glasses were not included in this preliminary study.

7. Acknowledgment

The support provided by the National Science Foundation under grants HRD-0317692, CNS-0428125, and the BPC program under CNS-0540592 is greatly appreciated.

8. References


Supported By

NSF Grant CNS-0540592