FACIAL FEATURE DETECTION IN NEAR-INFRARED IMAGES

Dai-Yun Li    Wen-Hung Liao

Department of Computer Science, National Chengchi University, Taipei, TAIWAN

ABSTRACT

We propose to employ near-infrared (NIR) images for face recognition in reduced illumination or total darkness. A homomorphic processing technique has been developed to effectively reduce the artifact of NIR images [1]. In this paper, we proceed to construct a facial feature detection system that would function independent of the surrounding lighting condition. Firstly, we propose a classification method based on local histogram analysis to separate NIR images captured at a short range from those in other circumstances. Afterwards, we present an algorithm to mark predominant facial features in a nearly frontal-face NIR images acquired at a short range. Experimental results demonstrate that facial feature points can be located accurately in homomorphic-filtered NIR images.

1. INTRODUCTION

Robust facial feature detection requires the ability to locate and identify the predominant facial features despite poor illumination conditions. By employing camcorders with near-infrared imaging capability, it becomes cost-effective to capture images under dim light or in total darkness. However, images thus acquired exhibit undesirable over-exposure artifact near its center. In [1], we have formulated an image formation model for NIR images and constructed a homomorphic filtering technique to address the foregoing issue. In this paper, we proceed to investigate the feasibility of using NIR images as the basis for object recognition. To begin with, we show that homomorphic processing technique is beneficial to the subsequent feature extraction task. We have also applied a similar procedure to non-uniformly illuminated color images, and have gained promising results.

For the purpose of locating meaningful features, a closer look at the characteristics of NIR images is in order. It is observed that the effective range of NIR cameras is about 10 meters when there is absolutely no ambient light. Beyond this scope, the strength of the near infrared light reflected by the target object is too weak to generate recognizable images. Therefore, an incoming image should be analyzed and classified so as to direct it to a proper destination for further processing. In accordance with the imaging distance and illumination factor, there are three categories of NIR images, namely, (1) images captured under normal lighting, (2) images captured in total darkness at a short distance, and (3) images captured in total darkness at a large distance. NIR images acquired at a large distance contain very few salient features and cannot be employed for feature detection tasks. Images captured under normal lighting condition can be processed with existing facial feature detection algorithms [2]. The focus of this paper will be on the analysis of images belonging to category 2.

We propose a classification method based on local histogram analysis, as well as the relationship between intensity level and the locations, to separate NIR images captured in a short range from those in other circumstances. With the properly classified and preprocessed NIR images at hand, we then develop a feature detection algorithm to locate and identify facial features in NIR images containing single frontal view of a human face. A logical next step is to use these detected features as the basis for face recognition.

The rest of this paper is organized as follows: Section 2 shows that preprocessing NIR images by homomorphic filtering techniques is indeed beneficial to the subsequent feature detection task. Color image correction results are also included in this section. Section 3 discusses the characteristics of the input images to motivate and assist the formulation of an image classification algorithm. In Section 4, we present a facial feature detection algorithm for homomorphic-filtered NIR images. Experimental results of facial feature detection are given in Section 5. Section 6 ends with a brief conclusion and an outlook on how the proposed method can be applied to face recognition.

2. HOMOMORPHIC FILTERING OF NIR IMAGES

2.1. Preprocessing NIR Images

A “0 LUX” operateable camera works by emitting invisible near-infrared (NIR) lights toward the direction of target objects and collecting the reflected light to form the image. The strength of the reflected energy, or equivalently, the intensity of the image, is a function of
the distance between the camera and the target object. The response decays as one moves away from the center of the image, as shown in Fig. 1(a). The apparent non-uniformity introduced during the image formation process can be effectively reduced using the homomorphic processing techniques reported in [1]. Fig. 1(b) shows the processed outcome.

![Image](image1.jpg)

Figure 1: (a) Image of a human subject recorded in total darkness and (b) its corresponding homomorphic-filtered image.

We conduct the following experiment to demonstrate that the preprocessing stage is indeed necessary. We first apply histogram equalization to both images so that the levels of the histogram-equalized image will span a fuller range of the gray scale. Afterwards, the Sobel operator [3] is used to highlight the edges in the NIR image with and without homomorphic filtering. The results are shown in Fig. 2. It is evident that the magnitude of the detected edges from the processed image is much more significant than the original NIR image. In other words, preprocessing NIR images by homomorphic filtering techniques proves to be beneficial to the succeeding feature extraction task.

![Image](image2.jpg)

Figure 2: Detected edges from images shown in Fig. 1.

2.2. Color Image Correction

Color images captured under a certain lighting condition may exhibit the artifact that is similar to that of the NIR images. In common with the NIR images, it is desirable to reduce or totally eliminate such artifacts using a similar approach. To begin with, the transformation from a vector-valued color images to a scalar-valued gray scale images is essential. We convert color images from the RGB color system to the Hue-Saturation-Value (HSV) space. Only the component (V) that represents the intensity of the image is utilized and processed by the homomorphic filtering techniques. The results are then converted back to the original RGB color system by combining the original H, S components and the processed component V.

The corrections of the color images with seemingly over-exposed region around the center are demonstrated in following illustrations. Fig. 3 is the color image simulated with non-uniformity in illumination and its corresponding color components, i.e. hue, saturation, and intensity, are also shown. The filtered result is shown in Fig. 4.

![Image](image3.jpg)

Figure 3: A full-color RGB image with highly non-uniform brightness around the center and its corresponding HSV components.

![Image](image4.jpg)

Figure 4: (a) preprocessed V component image; (b) resulting RGB image.

3. NIR IMAGE CLASSIFICATION

3.1. Characteristics of NIR Images

Even though NIR cameras can function under a 0 LUX condition, they do have a limited range of sight. General speaking, the image of an object in a totally darkness environment captured by the camcorder with NightShot® function could be recognized if the distance between the camera and the object is less than 10 meters. Beyond this scope, the strength of the near infrared light reflected by the target object is so weak that it is incapable of forming recognizable images, such as the image shown in Fig. 5. Therefore, before preprocessing an input image with homomorphic filtering techniques for further facial feature detection, a classification task must be performed. Three different circumstances under which the images were captured, including (1) normal lighting condition, (2) short range under low illumination, and (3) long range under low illumination are identified. Only those images captured in the condition described in 2 merit further investigation and analysis.
3.2. Image Classification

To accurately classify an input image into one of the three categories, we will make use of both the local histogram analysis of the image and the relationship between intensity level and the locations. Firstly, we divide an input image into two parts, the inner portion and the outer portion, according to the distance from the image center. In the next step, we exploit the local histogram curves of the inner and outer parts to compute the \( D \) value, which represents the degree of confidence that the input image belongs to the NIR image in a short range. If the calculated \( D \) of the image is greater than 1, we will classify the image into an NIR image captured at short range; otherwise, we have to compute the average intensity of the overall image to differentiate the NIR image at long range from normal light image. Fig. 6 illustrates the concept just described. The pixels contained within the black frame constitute the inner part; while those reside in the outer of the white frame constitute the outer part. The respective histogram curves of these two areas are depicted in Fig. 6(b). The details of image classification algorithm, including the exact definition of \( D \), are explained below.

**NIR Image Classification Algorithm**

In this algorithm, \( inner \) stands for the inner 1/36 portion of the image, and \( outer \) stands for the image minus from the inner 1/4 portion. \( global\_mean \) denotes the average intensity of the overall image, \( inner\_mean \) denotes the average intensity of the inner portion of the image, and \( outer\_mean \) is defined similarly. The \texttt{histogram} function returns a vector with 256 elements, which represents the number of pixels in the image with an intensity value of \( i \).

1) Compute histogram

\[
\begin{align*}
  h\_outer &= \text{normalize}(\text{histogram}(\text{outer})) \\
  h\_inner &= \text{normalize}(\text{histogram}(\text{inner})) 
\end{align*}
\]

2) Compute \( D \) value

\[
\begin{align*}
  \text{maxp} &= \max(h\_outer) \\
  \text{for each index } i \\
  &\begin{cases} \\
    \text{if } ((h\_inner[i] == 0) \text{ and } (h\_outer[i] > 0)) \\
    \quad \text{then } K\_L = K\_L + h\_outer[i] \\
    \text{if } ((h\_inner[i] > 0) \text{ and } (h\_outer[i] == 0)) \\
    \quad \text{then } K\_L = K\_L + h\_inner[i] \\
    \text{if } (\text{abs}(h\_inner[i] - h\_outer[i]) > (\text{maxp} \times 0.7)) \\
    \quad \text{then } K\_L = K\_L + \text{abs}(h\_inner[i] - h\_outer[i]) \\
  \end{cases} \\
  D &= (\text{mean}\_inner - \text{mean}\_outer) \times K\_L / \text{mean}\_outer
\end{align*}
\]

3) Classification

\[
\begin{align*}
  &\text{if } (D > 1) \text{ then category 2} \\
  &\text{else} \{ \\
  &\text{if } (\text{global}\_mean > \text{threshold}) \text{ then category 1} \\
  &\text{else category 3} \\
  \}
\end{align*}
\]

4. A FACIAL FEATURE DETECTION ALGORITHM

After discriminating the NIR image acquired in a short range from the other circumstances, we can apply homomorphic filtering techniques to enhance the image quality. With the enhanced images at hand, our goal now turns to the construction of an algorithm capable of automatically detecting facial features in a frontal-view face in homomorphic-filtered images regardless of background cluster.

The key idea in our approach is to make use of the density of horizontal edges of an input image. In addition, \textit{a priori} knowledge about the human face would be helpful in locating facial features of concern. Below is an explanation of the facial feature detection algorithm. As discussed previously, NIR images captured at a short distance often display a highly bright region around the center. It implies that the face is located at the “super-bright” area in a NIR image. Therefore, we first search for the threshold value \( T \) to convert the homomorphic-filtered image into a binary image, which reserves the portion of the face for detecting facial features and discards the irrelevant portion of the image. Mentioned earlier that the intensity histogram distribution of a NIR image resembles a bimodal distribution, thus it would be reasonable to assume that the threshold value \( T \) falls in the local minimum of the border between the two main modes. Fig. 7 shows one example where the cursory boundary of the face is extracted and binarized by the threshold method with \( T = 145 \).
After converting the homomorphic-filtered images into binary ones, some morphological filtering techniques such as opening and closing are put in use to get rid of noise and close up small holes in binary images. Here, we wish to locate the most stable and informative features in the face image. Typically these feature points are located near the center of the face. Locating the boundary of a face first by using a projection method would certainly constrain the search area. Let \( I(x,y) \) be the intensity value of an image at position \((x,y)\), the horizontal projections of the image are defined as:

\[
H_I(x) = \sum_{y=1}^{n} I(x,y),
\]

i.e., the intensities accumulated by each column of an input image. In the same way, the vertical projections of an image are defined as:

\[
V_I(y) = \sum_{x=1}^{n} I(x,y),
\]

i.e., the intensities accumulated by each row of an input image. Horizontal profile of a binary image is obtained first. The peak of the profile corresponds roughly to the column that the nose tip is located. Next, we apply some heuristics, namely, the first \( x \) from left side with projection value exceeding one of fourth of the maximum projection value of \( H_I \) is selected as left boundary of the face. The right boundary of the face is defined in the same manner. Similarly, the vertical profile is obtained and the abrupt change in plain curve is assumed to correspond to the location of the hair line. We also make use of the property that the ratio of length to width of human face in general is within a certain range, i.e. around the Golden Ratio (1.618) [4], to get the approximate location of the jaw line. Fig. 8 shows the detected boundary of face candidate and the corresponding processed binary image. The horizontal profile and vertical profile of Fig. 8 (b) are illustrated in Fig. 9.

However, after the foregoing preprocessing, we suffer from the problem: loss of information in the preprocessing stage. Some feature boundaries can be weakened and become hard to detect in the binary image, while the others are still retained. This additional filtering stage proves to be very useful in eliminating noise and confusion by other features. It makes the detection of the eyes more robust.

The first phase of the feature detection algorithm aims to compute the horizontal edge among the filtered binary image. The pixel with the highest density is regarded as the strongest features: in our case, the eyes. The second phase refines the horizontal edge density by scanning the homomorphic-filtered image area surrounded by the eyes to locate the pupils more precisely. The third phase exploits the positions of the detected features and the prior knowledge of relative relation of feature locations to find less salient features, such as the mouth, the jaw line, and the eye brows.

Fig. 10 shows the detected horizontal edges of an image in Fig. 8(b). Evidently, the area surrounded by the eyes possesses the most significant density of horizontal edges. Therefore, we take advantage of this property to search for the eyes, which are located at the opposite side of the nose in the upper-half face. With the approximate positions of the eyes, we could continue to find the exact location of the pupils with ease.

After that, we continue to search for the mouth, which has a similar \( x \)-coordinate as that of the nose and also has the significant density of horizontal edges in the ‘lower face’. With these criteria, the algorithm might fail to detection of the mouth, but acquiring the nose tip or the jaw line. For this reason, we have formulated the
following heuristics in order to prevent from erroneously detecting the lower jaw or the nose tip. Firstly, having locations of eyes, we are able obtain preliminary location of the mouth. Secondly, the length of horizontal edges of detected mouth candidate should not exceed a certain proportion of the width between two eyes. Finally, we can keep on detecting the jaw line below the mouth and two eyebrows above the pupils within the face boundary easily. In Fig. 11, we illustrate the flow chart of the facial feature detection algorithm.

**Facial Feature Detection Algorithm**

![Flow chart of facial feature detection algorithm.](image)

**6. CONCLUSIONS**

We first demonstrate the noticeable improvements of the homomorphic processed NIR images. We also show that the preprocessing facilitates the subsequent feature extraction step. We then present a new facial feature detection algorithm, which have successfully located several fiducial points, including the pupils, the eye brows, the mouth and the jaw line, on a human face. It is speculated that with appropriate pre-processing and recognition algorithms, a face recognition system using NIR images can perform with great accuracy under various lighting conditions.

**7. ACKNOWLEDGEMENTS**

This work was supported by NSC, under Grant NSC-90-2213-E-004-006.

**8. REFERENCES**