Gradient Edge Map Features for Frontal Face Recognition under Extreme Illumination Changes

Ognjen Arandjelović
ognjen.arandjelovic@gmail.com
Swansea University, UK

The aim of this work is to match images of frontal faces across extreme illumination changes. This is a problem of importance in a broad range of practical scenarios. For example, the user is commonly asked to face the camera in security applications which perform authentication before granting access to a resource. Retrieval systems also often focus on nearly frontal faces because face detection is most reliable for this pose.

Discriminative information is not uniformly distributed across different parts of a face. Rather, most of it is contained in the regions which exhibit substantial variation in either geometry or albedo and which can be readily detected using direct computations on pixel intensities. One of the simplest methods of accomplishing this is by applying a 2D high pass filter. However, when applied on images acquired under extreme illuminations, the simple high pass filter fails in achieving a satisfactory result. One of the reasons can be readily observed by examining Figures 1(a) and 1(b). Notice that the discontinuities in the poorly lit, shadowed regions are less pronounced than those in well illuminated regions.

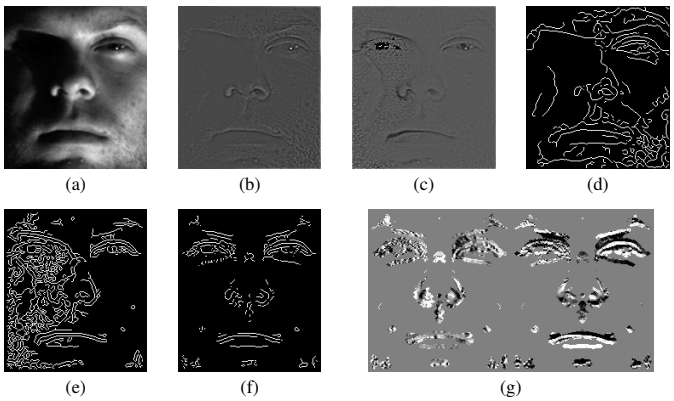

Figure 1: (a) Raw and (b) band pass filtered appearance, (c) self-quotient image, edges computed from (d) the original and (e) the self-quotient image, (f) symmetrically consistent and reliable edges, and (g) the proposed representation.

The dependence of the magnitude of the intensity discontinuities preserved by high pass filtering can be addressed by using one of the Retinex-like methods. These are in part inspired by the human visual system and the observation that humans perceive brightness in a relative rather than absolute manner. In other words, a discontinuity of a small magnitude in a dark region should have a greater effect than a discontinuity of the same magnitude in a bright region. The high pass filter can thus be modified simply by dividing pixel-wise the filtered result with the low pass filtered image which has the effect of averaging image intensity. This is a variant of the self-quotient image:

$$I_{SQI}(x,y) = I_{HP}(x,y) \;/\; I_{LP}(x,y) \tag{1}$$

The result of applying this filter is shown in Figure 1(c) which indeed appears to be an improvement over the output of the high pass filter. However, when this representation is used for matching on our data set, as discussed in detail in the paper, the error rate is increased to nearly that achieved by using raw appearance. A more detailed inspection of the resulting image reveals insight into the causes. Specifically, the noise in the poorly lit regions of the face has been amplified, as has the originally imperceptible boundary of the shadow caused by (weak) ambient illumination. The filter also causes the appearance of artefacts around interfaces between very bright and very dark image regions.

Steps 1 and 2: Provisional Edges Our method avoids the described difficulties associated with the use of absolute intensity by concentrating on the binary edge image. This is the first step of the proposed cascade. Note that we do not detect edges directly in the original image. Instead, we apply the Canny edge detector on the self-quotient image, to ensure that very weak edges in poorly illuminated regions are correctly detected. The difference between the two approaches is illustrated in Figures 1(d) and 1(e). Note that our approach results in higher automatically estimated Canny thresholds. While this has the effect of producing fewer false edges in the well lit regions of the face and more true edges in the poorly lit regions, the number of spurious edges in the poorly lit regions is also increased. This problem is addressed in the next step of our cascade.

Step 3: Spurious Edge Removal The edge map computed from the self-quotient image may contain many false edges e.g. from the amplification of noise in poorly lit regions or from the boundaries of cast shadows. Highly saturated image regions may also cause the hallucination of edges. Regardless of what the underlying cause is, false edges can decrease the matching accuracy. For example, it is straightforward to see that the left hand side of the edge map in Figure 1(e), full of densely packed false edges, will match nearly any true face edge map rather well.

We remove false edges by exploiting the vertical symmetry of frontal faces by requiring agreement between the left hand and right hand sides of the edge map. If E is the binary edge image and $\overset{\frown}{\cdot}$ the vertical mirroring operator the edge image E_T with spurious edges removed is computed as:

$$E_T(x,y) = \begin{cases} 1 & : E(x,y) = 1 \text{ and } \overset{\frown}{E}_{DT}(x,y) \leq 2 \\ 0 & : \text{otherwise} \end{cases} \tag{2}$$

where E_{DT} is the distance transformed edge image. In other words, we remove all edge segments which are not within ≈ 2 pixels from the corresponding mirrored edges, see Figure 1(f).

Steps 4 and 5: Edge Reliability Refinement After spurious edges are removed in the previous step of the cascade, the resulting binary image E_T is not necessarily vertically symmetric. We interpret this lack of symmetry as arising from true but unreliable edges. To ensure that the final representation contains only those true edges which are repeatably detectable, we again exploit the vertical symmetry of frontal faces. We first dilate the edge image E_T using a 4×4 pixel solid circle structuring element S and then combine the dilated edge information from the left hand and right hand sides of the face:

$$E_R(x,y) = \min\left[E_T(x,y) \oplus S, \overset{\frown}{E}_T(x,y) \oplus S\right] \tag{3}$$

Steps 6 and 7: Merging Edge and Gradient Information In the last step of the proposed cascade, we incorporate into our representation further discriminative information. The specific limitation of the edge map that we wish to overcome is its limited ability to robustly capture shape. This is a consequence of the observation that each edge map pixel by itself only contains information about whether an edge passes through it or not. Edge pixels carry no additional information about the directionality of the corresponding edge. We demonstrate that a highly discriminative representation can be obtained by combining the dilated reliable edges map and the corresponding gradient phase. This is achieved by computing a 3D image comprising two "stacked" 2D images which contain horizontal and vertical gradients at the dilated edges:

$$E_{GM}(x,y,1) = \begin{cases} \frac{\partial I}{\partial x}/|\nabla I| & : E_R(x,y) > 0 \\ 0 & : E_R(x,y) = 0 \end{cases} \tag{4}$$

$$E_{GM}(x,y,2) = \begin{cases} \frac{\partial I}{\partial y}/|\nabla I| & : E_R(x,y) > 0 \\ 0 & : E_R(x,y) = 0 \end{cases} \tag{5}$$

The normalization by magnitude is performed to account for the unreliability of absolute or even relative image intensity across different illuminations. The directionality of the gradient, on the other hand, is preserved well in the vicinity of strong discontinuities (but not necessarily elsewhere). The proposed representation is illustrated in Figure 1(g), displayed as the two stacked images side by side.

The effectiveness of the proposed representation was demonstrated on the notoriously challenging YaleB data set, which covers a wide range of illumination conditions, many of which are extreme (rear lateral, overhead). Unlike most of the previous work we used only a single image per person for training and a single probe image as test, and did not eliminate any of the images from the evaluation. Our gradient edge map achieved outstanding results, incorrectly recognizing in only 0.8% of the cases and exhibiting nearly perfect receiver-operator characteristic behaviour. This performance vastly exceeds that reported previously in the literature on this data set and using the same evaluation methodology.