Quality driven iris recognition improvement

Sandra Cremer, Nadege Lemperiere, Bernadette Dorizzi, Sonia Garcia-Salicetti

Thales Communications & Security
20-22 rue Grange Dame Rose, 78141 Velizy Cedex
Institut Mines-Telecom, Telecom SudParis
9 rue Charles Fourier, 91011 Evry Cedex
sandra.cremer@thalesgroup.com, nadege.lemperiere@thalesgroup.com
bernadette.dorizzi@telecom-sudparis.eu, sonia.garcia@telecom-sudparis.eu

Abstract: The purpose of the work presented in this paper is to adapt the feature extraction and matching steps of iris recognition to the quality of the input images. To this end we define a GMM-based global quality metric associated to a pair of normalized iris images. It quantifies the amount of artifact in these images as well as the amount of texture in artifact-free regions. First we use this metric to adjust, for each pair of irises, the proportion of the normalized image selected on a local quality criteria for feature extraction. This approach is tested with two matching techniques: one performs a bit to bit comparison of binary feature vectors and the other one computes local cross-correlations between real valued vectors. We show that our approach is effective with both techniques. Then we use our metric to choose the matching technique that is best adapted to each image pair in order to make a good compromise between accuracy and speed.

1 Introduction

Irises possess a very rich pattern that is believed to be different between persons, therefore iris recognition has become one of the most reliable and accurate biometric identification systems available. A detailed survey on iris recognition was published by Bowyer et al. [BHF08].

The first successful algorithm for iris recognition was proposed by Daugman [Dau93]. After a preprocessing of the iris that includes a segmentation and a normalization step, this algorithm uses convolution with two dimensional Gabor filters to extract the texture from the polar iris image. Each filter’s phase response is then quantized in a pair of bits, so the information from the iris image is represented by a binary code. Following this, different images of irises can be compared through their binary codes using bitwise operations.

Daugman’s algorithm is used in most of today’s commercial iris recognition systems for it is very efficient computationally and its evaluations on very large databases have proven its high recognition accuracy [Dau04]. However it has been shown that its performance considerably decreases on images with degraded quality, for instance images affected by defocus [KDS02], occlusion [KDS02] or pupil dilation [HBF09].
A lot of research has been done to quantify iris image properties that are influential on recognition performance. The National Institute of Standards and Technology has conducted the Iris Quality Calibration and Evaluation (IQCE) [TGS] to this end. A number of global quality metrics were defined to qualify iris images. They can be used to reject low quality images in the recognition process and thus improve performance. Some work has also been done on defining local quality metrics, namely by Chen et al. [CDJ05], Li and Savvides [hLS09] or Cremer et al. [CDGSLpt]. They are used to diminish the impact of low quality regions on the computation of the Hamming distance when using techniques for feature extraction and matching that are similar to Daugman’s.

In [CDGSLpt], we showed that our local quality metric enables to localize regions of the polar iris images that contain artifacts, as well as to quantify the amount of texture in artifact-free regions. In this paper we exploit this local quality metric to define a global quality metric. Unlike most global quality metrics defined in literature, it is computed from the polar iris images. It qualifies these images in terms of amount of artifacts and of iris texture which enables in particular to detect segmentation inaccuracies. Our aim is to use it to improve the feature extraction and matching steps of the recognition process when dealing with different levels of quality image degradations.

More specifically, the purpose of this work is to combine the exploitation of the local quality metric presented in [CDGSLpt] and of our global quality metric in order to adapt the feature extraction and matching steps to each pair of iris images to compare. In [CDGSLpt] local image quality had been incorporated in Daugman’s algorithm to select the best regions in the polar iris images in terms of occlusions and amount of texture. It had been shown that the proportion of the initial image selected is a parameter that is very influential on performance. Here we go further and show that the optimal proportion of the initial image to select is directly linked to the global quality of the image pair to compare. We therefore chose an approach in which we use a local quality metric to localize the regions to select for feature extraction and matching and a global quality metric to decide how many regions should be selected. We test this approach with two different techniques for feature extraction and matching: Daugman’s [Dau93] and Krichen’s [KGSD09]. We chose Krichen’s technique because it is presented as being more robust to degraded input iris images. Krichen et al. keep Daugman’s idea of using Gabor phases to describe the iris image, however they use a real-valued feature vector to represent the iris and perform local correlation between these vectors to compute a similarity score.

Moreover we exploit our global quality metric to decide which feature extraction and matching technique out of Daugman’s or Krichen’s is the best adapted to each iris pair in order to make a good compromise between recognition accuracy and speed. The main novelty of this paper is to use a global quality metric to adapt the recognition process to each pair of images instead of using it to eliminate low quality images, as is done in most systems.

This paper is organized as follows. Section 2 describes briefly the local quality measure defined in [CDGSLpt] as well as the region selection strategy for feature extraction and matching. Section 3 shows how the proportion of the image that is selected can be adapted to the global quality of images, both when it is used with Daugman’s algorithm and Krichen’s. In Section 4, we show that one or the other algorithm performs best de-
pending on the value of this global quality measure. Finally, in Section 5, we propose an
iris recognition process in which we combine the local and global quality measures. We
work on two databases ND-IRIS-0405 [PSO+10] and CASIA-IrisV3-Lamp [CI] for they
are large databases that contain images with a large range of qualities. The ND-IRIS-0405
database contains 64,980 images from 356 unique subjects and the CASIA-IRISV3-Lamp
database contains 16,212 irises from 411 subjects.

2 Using local quality to select the right regions for feature extraction

2.1 Definition of the exploited local quality measure

In the same way as in [CDGSLpt], we used a Gaussian Mixture Model (GMM) to charac-
terize high quality sub-images. The sub-images that we have chosen to train the GMM are
free from occlusion, well-focused and highly textured. Four local observations, measured
in a neighborhood of each pixel form the input vector of our GMM: the local mean, local
variance and local contrast in addition to the values of the pixel grey-level.

The quality measure associated to a sub-image \( w \) is obtained by applying the formula 1,
where \( a \cdot b \) is the size of the sub-image \( w \) and \( x_i \) is the input vector of our GMM associated
to a pixel indexed \( i \). \( P(x_i|\lambda) \) is the likelihood given by the model \( \lambda \) to the input vector \( x_i \).
\( \mathbb{L}_{\text{train}} \) is the mean log-likelihood on the training set, computed over the training set of
sub-images and the spatial region of each sub-image.

\[
Q(w) = \exp\left(-\frac{1}{a \cdot b} \sum_{i=1}^{a \cdot b} \left| \log(P(x_i|\lambda)) - \mathbb{L}_{\text{train}} \right| \right)
\]  

(1)

\( Q(w) \) will be comprised between 0 and 1. It is shown in [CDGSLpt] that it allocates the
highest values to sub-images that are artifact-free and highly textured and the lowest values
to sub-images containing artifacts. Intermediate values are given to artifact-free and lowly
textured sub-images.

2.2 Selection of the best regions for feature extraction

It has often been demonstrated as in [KDS+02] that iris recognition performs best when
the features have been extracted in high quality regions of the polar iris image, namely non-
occluded and highly textured zones. As in [CDGSLpt], we use the local quality measure
described here above to find those regions. More specifically we proceed in 4 steps for
each pair of gallery and probe iris images to compare:

Step 1: In each polar iris image sized 64x512 pixels, we compute the local quality of \( M \)
sub-images sized 11x51 pixels. \( M \) is the same for all the images in the database.

Step 2: For each pair of gallery and probe polar images, we fuse the gallery local qualities
with the probe ones. This is done by selecting, for each one of the $M$ pairs of sub-images, the minimum between the gallery and the probe quality value.

**Step 3**: We sort the pairs of (gallery, probe) sub-images according to their local quality.

**Step 4**: We keep the $N$ best sub-images for feature extraction and matching.

This process is done for different translations of the probe polar image in order to allow in-plane rotations of the raw iris in the image.

In [CDGSLpt], this strategy has been set-up with Daugman’s feature extraction and matching algorithm on three public databases: ND-IRIS-0405 [PSO+10], CASIA-IrisV3-Twins [CI] and CASIA-IrisV3-Interval [CI]. It was shown that, this strategy performs better than other strategies that select the regions used for recognition by computing a segmentation mask, for it allows to keep control on the quantity of information exploited through the choice of the value $N$, for a set value of $M$. Moreover, different values of the ratio $N/M$ have been tested. It has been demonstrated that when using the same value of $N/M$ for all (gallery, probe) pairs of images, there is an optimal value of $N/M$ equal to $1/3$ whichever database was used and whichever was the value of $M$. This ratio represents the proportion of the polar iris image that is exploited for recognition. Its optimal value compromises between minimizing the amount of occluded regions taken into account for feature extraction and maximizing the amount of information available for matching. Indeed, when this ratio is too high, regions of bad quality are taken into account for recognition. On the other hand, when it is too low, the amount of information exploited for matching is not high enough and the impostor distribution is biased.

3 Using global quality to decide what proportion of the polar iris should be kept

In Section 2, the proportion of the polar iris image exploited for feature extraction and matching was the same for all (gallery, probe) pairs of images to compare. In this section, we wish to demonstrate that the optimal value of this proportion, represented by the ratio $N/M$, can change for each (gallery, probe) pair depending on the pair’s quality. This is one of the novelties of the present work.

It appears obvious that the less occluded an image is, the more regions of the image can be used for matching without taking into account occlusions. Therefore the higher the image quality is in terms of occlusion, the higher the optimal value of the ratio $N/M$ should be. Moreover, non-occluded regions that contain very few texture won’t necessarily bring discriminating information for matching. So the amount of texture can also have an influence on the optimal value for $N/M$.

To qualify a pair of (probe, gallery) images globally in terms of occlusion and amount of texture we have exploited the local quality described in Section 2. Since we want to define the quality of each pair of (probe, gallery) images, we fuse the $M$ local qualities of the probe image with the $M$ local qualities of the gallery image. This is done by selecting the minimum value between the gallery and the probe local quality value. Indeed,
when matching the features corresponding to a gallery and a probe sub-image, one of both sub-images containing an artifact is enough for the chances of false rejection to increase drastically, so the selection of the minimum quality seems logical.

As a result, we define the global quality measure associated to a pair of images as the mean of the $M$ fused local measures corresponding to the $M$ pairs of (gallery, probe) sub-images as in equation 2

$$GlobQ = \frac{1}{M} \sum_{i=1}^{M} \min(Q_{gal}(w_i), Q_{probe}(w_i))$$

(2)

$GlobQ$ is directly linked to the local measure used for feature selection. We therefore expect the ratio $N/M$ to depend on $GlobQ$.

### 3.1 Experiments using Daugman’s algorithm

As in [CDGSLpt], we have used the OSIRIS-V2 version of Daugman’s algorithm. However we have replaced the segmentation module by the one described in [LDGS+13], in order to better model the iris’ inner and outer boundaries. The segmented images are then normalized using Daugman’s rubber-sheet model [Dau93] to perform a polar transformation of the iris. After this, we apply the strategy described in Section 2.2 to select the $N$ best sub-images out of the total number $M$, for each (probe, gallery) pair of polar images. $M$ is set to a constant value throughout this experiment. The $N$ selected sub-images will then be convolved with two dimensional Gabor filters and each filter’s phase response will be quantized in a pair of bits. So the information of each iris image is summarized in a binary code. The probe and gallery binary codes will be compared by the computation of their Hamming distance.

We have applied this recognition process on images from the ND-IRIS-0405 and the CASIA-IrisV3-Lamp databases. More specifically, in order to demonstrate the link between the global quality defined in formula 2 and the optimal value of $N/M$, we have generated matching lists for these databases and then divided them into five categories according to the value of the global quality associated to each (gallery, probe) pair. Then we have applied our recognition process on each category for different values of $N$ and compared the achieved performance. The values of the global quality measure that define the boundaries of each of the five categories were determined by analyzing the range of the global quality on the entire list and dividing the resulting quality interval into five equal-sized intervals. Table 1 presents the percentage of the full matching list in each category for the ND-IRIS-0405 and CASIA-IrisV3-Lamp databases and Figure 1 gives examples of pairs of (gallery, probe) polar iris images belonging to categories 1, 3 and 5.

We have started by applying our recognition process to the first category of matches ($0 < GlobQ < 0.13$) for different values of $N$. The values tested for $N/M$ were 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9. Figure 2 presents the ROC curves, i.e the False Rejection Rate (FRR) as a function of the False Acceptance rate (FAR), obtained for 4 values of $N/M$
Table 1: Percentage of the full matching list for each category of GlobQ.

<table>
<thead>
<tr>
<th>GlobQ</th>
<th>[0; 0.13]</th>
<th>[0.13; 0.26]</th>
<th>[0.26; 0.39]</th>
<th>[0.39; 0.52]</th>
<th>[0.52; 0.65]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ND-IRIS</td>
<td>11%</td>
<td>16%</td>
<td>34%</td>
<td>33%</td>
<td>6%</td>
</tr>
<tr>
<td>CASIA</td>
<td>13%</td>
<td>21%</td>
<td>39%</td>
<td>24%</td>
<td>3%</td>
</tr>
</tbody>
</table>

Figure 1: Examples of (gallery, probe) polar iris image pairs for different global qualities.

on this category corresponding to the lowest-quality images of the CASIA-IrisV3-Lamp database.

As we can see, the best performance is achieved for \( N/M = 0.4 \). We have carried out the same experiment on the four other categories of matches. Table 2 presents the values of \( N/M \) that lead to the best performance for each category for the CASIA-IrisV3-Lamp and the ND-IRIS-0405 databases.

As we can see, the optimal value of \( N/M \) is the same for the two first categories and increases with GlobQ for the three other ones. This can be explained by the fact that a higher value of GlobQ implies that the quality of the images is better in terms of occlusion and amount of texture, so taking into account more regions for feature extraction and matching adds useful information for recognition. It is interesting to note, that the optimal value of \( N/M \) is the same for categories 1 and 2, which means that when GlobQ goes beneath a certain threshold, \( N/M \) stays stable and does not decrease anymore. One reason for this is that this value of \( N/M \) is the value underneath which too few information is taken into account for the matching to be significant. Consequently, under a certain value of GlobQ, taking into account regions of low quality has a less negative impact on recognition performance than taking into account too few regions.

Table 2 shows that the optimal value of \( N/M \) for \((0.39 < \text{GlobQ} < 0.52)\) is 0.6. On the

Table 2: Optimal values for \( N/M \) for different databases and different values of GlobQ when using Daugman's algorithm.
Figure 2: ROC curves for 4 values of $N/M$ and for matches from the CASIA-IrisV3-Lamp databases for which $0 < \text{GlobQ} < 0.13$.

On the other hand, Figure 2 shows that using $N/M = 0.6$ for $0 < \text{GlobQ} < 0.13$, instead of $N/M = 0.4$ leads to a 30% increase of the FRR for an FAR=$10^{-4}$. This demonstrates that setting $N/M$ appropriately with the knowledge of $\text{GlobQ}$ can considerably improve performance compared to the use of an inadequate value of $N/M$.

Moreover, for a given range of $\text{GlobQ}$, the optimal value for $N/M$ is the same whether the exploited database is ND-IRIS-0405 or CASIA-IrisV3-Lamp. So this value only depends on the image quality.

### 3.2 Experiments using Krichen’s algorithm

We have conducted the same type of experiment with the recognition algorithm proposed by Krichen et al. [KGSD09]. More specifically we have segmented and normalized the iris images in the same way as is section 3.1. and applied our strategy to select the $N$ best sub-images in polar iris images. We have then performed local normalized cross-correlations on the Gabor phases of these sub-images to compare the iris images. So the recognition process we have used here for a (gallery, probe) pair can be cut into the following steps after segmentation and normalization of the gallery and probe images:

**Step 1:** computation of the local qualities of $M$ sub-images from the galley and probe images

**Step 2:** fusion of the local qualities of the gallery and probe sub-images

**Step 3:** selection of the $N$ best sub-images for convolution with 2D Gabor filters

**Step 4:** local normalized cross-correlations for the $N$ pairs of (gallery, probe) sub-images

**Step 5:** similarity score computation

The 3 first steps are identical to the ones done in section 2.2 with $M$ set to a constant.
We have demonstrated that the global quality we have defined in section 3 can be used to improve performance by adjusting the proportion $N/M$ of the images used for feature extraction and matching, whether the technique used is Daugman’s or Krichen’s. In this section we present another useful way of exploiting this global quality measure.

We compare the performance of Daugman’s and Krichen’s algorithms on the five categories of image lists described in section 3.1 when they are used with the local-quality-based region selection strategy for feature extraction as in sections 3.1 and 3.2. We use the results from Table 2, to set the value of $N/M$ for each category as the value that maximizes performance. The comparative performance of the Daugman and Krichen based processes is presented in Table 3 and Table 4 for the ND-IRIS-0405 database and CASIA-IrisV3-Lamp database respectively.

It is interesting to notice that the Daugman based algorithm performs less well than the one based on Krichen’s technique on the categories of lower global quality. However the

<table>
<thead>
<tr>
<th>$\text{GlobQ}$</th>
<th>$[0; 0.13]$</th>
<th>$[0.13; 0.26]$</th>
<th>$[0.26; 0.52]$</th>
<th>$[0.39; 0.52]$</th>
<th>$[0.52; 0.65]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N/M$</td>
<td>0.4</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
</tr>
<tr>
<td>Daugman</td>
<td>33%</td>
<td>11%</td>
<td>3%</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>Krichen</td>
<td>19%</td>
<td>6%</td>
<td>2%</td>
<td>1%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Table 3: FRR for FAR=$10^{-3}$ on the ND-IRIS database for different values of GlobQ and $N/M$ when using the recognition process on Daugman’s and Krichen’s techniques.
Table 4: FRR for FAR=$10^{-3}$ on the CASIA-IrisV3-Lamp database for different values of $GlobQ$ and $N/M$ when using the recognition process on Daugman’s and Krichen’s techniques.

<table>
<thead>
<tr>
<th>$GlobQ$</th>
<th>[0; 0.13]</th>
<th>[0.13; 0.26]</th>
<th>[0.26; 0.52]</th>
<th>[0.39; 0.52]</th>
<th>[0.52; 0.65]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N/M$</td>
<td>0.4</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
</tr>
<tr>
<td>Daugman</td>
<td>17%</td>
<td>11%</td>
<td>3%</td>
<td>1%</td>
<td>0%</td>
</tr>
<tr>
<td>Krichen</td>
<td>13%</td>
<td>7%</td>
<td>2%</td>
<td>1%</td>
<td>0%</td>
</tr>
</tbody>
</table>

performance is equivalent on the category of best global quality.

This can be explained by the fact that Krichen’s algorithm does local correlations between real feature vectors, whereas Daugman’s algorithm does a bit to bit comparison allowing only global radial translations. Therefore Krichen’s algorithm is more robust to local distortions as well as a reduced amount of texture available than is Daugman’s. Both of these elements are linked to our global quality measure. Indeed local distortions mostly occur in the presence of occlusions which are detected by our quality measure. Moreover the amount of texture is also directly measured by our quality measure. We can therefore understand why Krichen’s algorithm performs better than Daugman’s on pair of images that have a low global quality according to our measure.

Moreover, it is important to note that the algorithm based on Krichen’s technique is much slower. The best compromise between speed and precision would therefore be to apply Krichen’s technique only on the image pairs for which it will perform best. In practice this can be done thank to our global quality measure, since the benefit of Krichen’s algorithm towards Daugman’s depends on $GlobQ$.

5 Final strategy

Based on the results presented in Section 3 and 4, we propose a complete iris recognition strategy that combines Krichen’s and Daugman’s algorithms as well as the local and global use of a GMM-based quality measure.

More specifically, for each pair of iris images to compare, we compute the associated local and global quality measures. We then set the proportion of the polar images to keep for feature extraction depending on the value of the global quality value, based on the results presented in Table 2. The regions to keep are chosen as the ones with the highest local quality values. Following this we use the global quality value once more to decide which feature extraction and matching technique to apply. If $GlobQ < GlobQ_{th}$ we perform the correlation-based method described in section 3.2. Otherwise we perform the bit-to-bit comparison method described in section 3.1.

Based on the results of Section 4, we have set the value of the threshold $GlobQ_{th}$ to 0.4. Indeed, since Krichen’s technique is more than one hundred times slower than Daugman’s, we only wish to use it when it performs significantly better than Daugman’s in terms of accuracy. Table 3 shows us that this is the case when the global quality is under 0.4.
Figure 3: ROC curves for iris image comparisons of the ND-IRS-0405 database for which $GlobQ < 0.4$ (upper graph) or $GlobQ \geq 0.4$ (lower graph).

Figure 3 presents the ROC curves obtained when applying the iris recognition we propose to image pairs from the ND-IRIS-0405 database. Note that the images used for these comparisons are different ones than those used in sections 3 and 4. To plot these ROC curves we have separated the comparisons in two categories: those for which $GlobQ < GlobQ_{th}$ and those for which $GlobQ \geq GlobQ_{th}$. We have compared our strategy to two other algorithms:

**Krichen04:** using the local quality-based region selection strategy with the same value of the $N/M$ parameter for all images: $N/M = 0.4$ and Krichen’s feature extraction and matching technique

**Daugman06:** using the local quality-based region selection strategy with the same value of the $N/M$ parameter for all images: $N/M = 0.6$ and Daugman’s feature extraction and matching technique

For each of these two algorithms, the constant value of $N/M$ for all comparisons has been chosen as the one that maximizes performance for the range of global qualities considered.

We can see that the best performance is achieved with the method we propose, whatever
the value of $GlobQ$. When $GlobQ < GlobQ_{th}$, the performance is very close to the one achieved with Krichen04, but it outperforms Krichen04 when $GlobQ \geq GlobQ_{th}$. On the contrary, the performance of our method is very close to the one of Daugman06 when $GlobQ \geq GlobQ_{th}$ but it exceeds Daugman06 performance when $GlobQ < GlobQ_{th}$.

6 Conclusion

In this article we associated a global GMM-based quality measure to each pair of (gallery, probe) polar iris images we wish to match in order to quantify the amount of occlusions present in the images as well as the amount of texture in the non-occluded areas. We demonstrated that this global quality measure can be used for two purposes. First it can be exploited to adjust, for each pair of irises, the proportion of the polar image selected for feature extraction on a local quality criteria. The value of this parameter has a major influence on recognition performance whether the selection strategy is applied to Daugman’s feature extraction and matching technique or Krichen’s. Setting it too high can imply taking into account regions with artifacts for matching, but setting it too low can lead to missing out on useful information for matching. Its optimal value decreases when the global quality of the image pair is reduced until a threshold value under which matching becomes insignificant being based on too few information. Moreover, the relationship between the global quality of an image pair and the optimal value of the proportion of the polar image to use for matching is the same whether the matching technique used is Daugman’s or Krichen’s. Therefore, the knowledge of the global quality of an image pair allows us to set this parameter so as to maximize performance whichever technique is used for matching.

Secondly, the global quality measure we define can be used to choose between using Daugman’s or Krichen’s feature extraction and matching technique for each pair. Indeed, we showed that Krichen’s technique performs better than Daugman’s on low quality images according to our measure, but is much slower.

In consequence we proposed a new strategy for iris recognition that combines Krichen and Daugman’s algorithms as well as the local and global use of a GMM-based quality measure. We use the local quality measure to select the regions in the probe and gallery polar iris images that are of highest quality in terms of non-occlusion and amount of texture. These regions will then be used for feature extraction and matching using either Krichen’s or Daugman’s techniques. The global quality measure we have defined is associated to a pair of polar images and is used for two purposes: to decide how many regions are selected for feature extraction and matching, and to choose which technique to use for this extraction and matching. More precisely, these two decisions are taken automatically for each pair of images given the value of the associated global quality measure.

The experiments we carried out showed that our strategy leads to a higher recognition accuracy than a strategy for which we would exploit the same proportion of the iris polar image for all images and the same technique for feature extraction and matching.
References


