Abstract—The International Civil Aviation Organization (ICAO) endorsed the use of face recognition as main biometric characteristic for identity validation. According to ICAO directives, ISO proposed the ISO/ IEC 19794-5 standard, which specifies some quality requirements for facial images. Verifying the compliance of a face image to these quality specifications requires an prior detection of eyeglasses. This paper presents a novel method for detecting glasses in frontal face images. The eyeglass lenses are connected by a piece which usually appears in the horizontal direction. The proposed method explores this idea performing a skin clustering operation followed by a horizontal edges detection. The skin detection process uses the Cr band of the YCbCr color system, which is widely used for skin detection. The performance was evaluated on the AR and FRGC databases, obtaining success rates of 98.58% and 89.33% respectively.

Keywords—Glasses Detection; YCbCr; ICAO.

I. INTRODUCTION

A high-performance glasses detection method in facial images is important to computer vision research areas, as face recognition, document photos automatic validation, etc [1]. A robust face recognition system has to be able to perform under non-ideal conditions, as individuals wearing glasses or poor lighting. Such factors affect the success of these systems [2]. According to Jia et al. [3], due to the convenience of facial image acquisition, face recognition has been widely used, such as in access control, staff supervision, immigration, etc.. However, high recognition rates are difficult to achieve because of the many possible variations in illumination, facial expression, angle, and accessories. Numerous studies have shown that eyeglasses have significant negative impacts on the face recognition.

Automatic glasses detection is also relevant in different scenarios. As eyeglasses differ too much (e.g., size, shape, color), analyzing information related to the glasses may aid in age and gender definition, since women’s glasses have distinct features from men’s glasses, as well as children’s glasses distinguish from adult’s glasses. It is also useful since presence of eyeglasses in face image interferes in the precise location of the eyes, which is an important task in biometric face recognition. Accurate eyes location and the distance between them is important because it allows inferring face dimensions and location of other face components as mouth and nose. So, assuming that the individual wear glasses makes the eye detection better, reducing errors in the biometric process [4].

Another scenario where glasses detection is quite important is in automatic validation of document photos. Face is a very relevant biometric characteristic which is used in many promising forensic and commercial applications such as access control, video-surveillance, etc. The International Civil Aviation Organization (ICAO) endorsed, in 2002, the use of face recognition as the globally interoperable biometric characteristic for machine assisted identity confirmation with machine readable travel documents [5].

In sequence, according to ICAO directives, the International Standard Organization (ISO) proposed the ISO/ IEC 19794-5 standard [6], which specifies some quality requirements for facial images. Some of those requirements establish rules related to the use of eyeglasses, such as: the person can’t use dark tinted eyegasses; the frame should not cover the eyes; no flash reflection in the lenses, etc. Verifying the compliance of a face image to these quality specifications requires an prior detection of the presence of eyeglasses in the image.

Although glasses detection is a challenging subject, this paper proposes a computationally simple approach and establish a comparative with other methods among the related literature, with no need of previous training.

This paper is organized as follows. In Section II is made a literature review related to glasses detection. Section III contains a brief introduction about the YCbCr color system. In Section IV the databases used to evaluate the algorithms are described. Section V presents in details the methods proposed by this work. Section VI shows the obtained results. Finally, Section VII has the conclusions from the results analysis and possible future works.

II. RELATED WORKS

Several methods were proposed to perform glasses detection. Jiang et al. [7] [8] approaches the problem using edge detection. The method starts by locating the eyes and estimating the upper cheeks region of the face. Edges information is collected from the interest region by using the Sobel filter [9]. Over the edge information are calculated Three measures to determine the presence of glasses. These metrics are describe in the following equations [1] [2] and [3].
locates and removes glasses from a facial image. The extraction or frame shape.

Therefore, it is possible to highlight a separation between restrictions, a set of features are identified in the glasses frame. So, based on the three-dimensional plane and given some restrictions, a set of features are identified in the glasses frame.

The detection assumes that the lenses are in the same plane of the three-dimensional space. The 3D Hough Transform is used to obtain the plane in which the features are concentrated.

These three measures are used as input vector for a class separation operation in which vector from no glasses images differs from faces wearing eyeglasses. According to the author, results showed that such approach achieves good separation between the two considered classes (with or without glasses).

Jiang [8] expands the idea applying those metrics in a new region of interest. Based on the Principle of Fisher Information [10], it was found that the region in the face which highlights the presence of glasses is the area above the nose, where the nosepiece of the eyeglasses is supported.

Jing et al. [2] presents a method for detection and feature extraction from glasses in facial images. In this work the detection is based on Jiang’s [7] while the extraction is made by deformable contours, matching edges intensity and orientation. At the end, after many iterations, the lenses contour is outlined. Jing [11] proposes a probabilistic approach. Initially, the rotation is fixed aligning the eyes with the horizontal axis and the size is normalized to standard values. Next, a binary edge detection filter is applied to the image and the presence of glasses is probabilistically estimated taking into account edge pixels only. According to the author, the tests show that this approach has higher performance than Jing et al. [2].

Haiyuan Wu et al. [12] describes an innovative method comprising glasses detection in 3D images, got from stereoscopy. The detection assumes that the lenses are in the same plane of the three-dimensional space. The 3D Hough Transform is used to obtain the plane in which the features are concentrated. So, based on the three-dimensional plane and given some restrictions, a set of features are identified in the glasses frame. Therefore, it is possible to highlight a separation between images with or without glasses. This approach does not require further information about face pose estimation, eyes location or frame shape.

Chenyu Wu et al. [13] present a system that recognizes, locates and removes glasses from a facial image. The extracted features are orientation patterns. These features are discriminated by a classifier based on error reconstruction. According to the author, this method achieved a success rate of 94.2%.

Bo Wu [14] shows a detection system based on machine learning. A variation of a boosting algorithm called Real AdaBoost is applied to improve the effectiveness of a simple wavelet classifier. The tests were made over 3000 images coming from the FERET database [15] and from the Web, a half with glasses and the other half without. After a five-round cross validation and detailed description of the experiments, the author states that the general success rate of his method is 94.0%.

Jia et al. [3] propose an eyeglasses detection and removal algorithm based on phase congruency and progressive inpainting. Taking into account the importance of human eyes in face recognition, the method first constructs a model to shield the eyes in an facial image based on a priori knowledge of facial features, then computed phase congruency to detect textures, and marked the eyeglasses frame. Finally, inpainting process is used to fill in the marked region by using information from nearby pixels. Image inpainting provides a means for reconstruction of small damaged portions of an image, so that the image looks more natural. In this case, the damaged portion is the glasses frame, and inpainting is used not to make the glasses-removed-image look subjectively better, but mainly to increase face recognition rate. Experimental results showed that this method improves recognition rate significantly and exhibited strong robustness against high variations in facial images, such as skin colors, poses, background, illuminations, etc.

Finally, the recent work of Fernandez et al. [4] brings an innovative approach to the subject, using Local Binary Patterns (LBP) as features representation and robust alignment. First, the image is preprocessed and the eyes region is delimited. Next, the LBP descriptors are built to describe the selected region. The classification of the extracted features is made by Support Vector Machine (SVM). Aiming to evaluate the method in a realistic environment, this algorithm was tested over the Labeled Faces in the Wild database (LFW) [16], whose purpose is the face analysis in unconstrained environments, containing 13,233 famous people images got from the internet highly varied in lighting, position, clothing, etc. The method achieved 98.65% success rate on this set of images. There is no other paper in related literature which uses the LFW as testing scope, so, to endorse that the method outperforms the state-of-the-art it was tested over the FERET database using five-round cross validation, yielding 99.89% success rate.

III. YCbCr Color System

According to Chitra et al. [17], color space is a mathematical model to represent color information as three or four different color components. Different color models are used for different applications such as computer graphics, image processing, TV broadcasting, and computer vision. There are many color spaces which are useful for the skin detection.
They are: RGB based color space (RGB, normalized RGB), Hue Based color space (HSI, HSV, and HSL), Luminance based color space (YCBCr, YIQ, and YUV), and perceptually uniform color space (CIEXYZ, CIELAB, and CIELUV).

The RGB color information presents a lot of redundancy. So, according to Kakumanu et al. [18] the orthogonal color spaces reduce the redundancy present in RGB color channels and represent the color with statistically independent components (as independent as possible). The Y value represents the luminance (or brightness) component, computed as a weighted sum of RGB values, the Cr and Cb values, also known as the color difference signals, represent the chrominance component of the image, they are computed by subtracting the luminance component from B and R values. As the luminance and chrominance components are explicitly separated, these spaces are a favorable choice for skin detection, and because of that, YCbCr space is one of the most popular choices for skin analysis. The values of the YCbCr components are calculated according to the following equations:

\[
\begin{align*}
Y &= 0.229 \cdot R + 0.587 \cdot G + 0.144 \cdot B \\
Cb &= 0.168 \cdot R + 0.3313 \cdot G + 0.5 \cdot B + 128 \quad (4) \\
Cr &= 0.5 \cdot R + 0.4187 \cdot G + 0.0813 \cdot B + 128
\end{align*}
\]

IV. MATERIALS

On the performance evaluation of the proposed method were used frontal face images from the AR Face Database [19] and Face Recognition Grand Challenge (FRGC) [20] Face Database. Some image samples of these bench are shown in Fig. 1. The FERET and LFW databases were not used due to lack of ground-truth.

Images from AR Database are 768x576 pixel sized and present variations regarding to smile, open mouth, sunglasses, scarf over face, unnatural skin color, flash on skin, etc. Besides, there is a deliberate and controlled variation in face illumination, switching between frontal and lateral. The background is slightly varied, mostly white. It contains over 4,000 color images corresponding to 126 people’s faces (70 men and 56 women). Images feature frontal view faces with different facial expressions, illumination conditions, and occlusions (sun glasses and scarf). The pictures were taken at the CVC (Computer Vision center, Barcelona University) under strictly controlled conditions. No restrictions on wear (clothes, glasses, etc.), make-up, hair style, etc. were imposed to participants. A subset with 1,000 images was selected, where half of the individuals wear glasses and the other half does not.

From the FRGC database 1,230 images were used in tests, again half of them with glasses, the other half without. These images are 2272x1704 and 1704x2272 pixel sized. This database provides facial image samples taken from realistic environments, presenting more varied conditions as regards lighting, face location in the image, image size and face size ratio and highly varied background. The data for FRGC consists of 50,000 recordings divided into training and validation partitions, once it focuses in face recognition. The training partition is designed for training algorithms and the validation partition is for assessing performance of an approach in a laboratory setting. The validation partition consists of data from 4,003 subject sessions.

V. METHODS

In this section, the proposed method for eyeglasses detection in facial images is presented. The method explores the upper nose region, searching for horizontal edges since the piece that join the lenses of the eye glasses usually appears horizontally in this region. The method explores color information from Cr band of YCbCr color system according to empirical tests in which Cr band got best results. This process is presented in details ahead.

Jiang [8] shows, based on the Fisher’s Information Criterion, that the facial region between eyebrows, the upper nose region, is the best for glasses detection in frontal face images. The location of this region is estimated according to the eyes position and distance between them. Therefore, the first step of the algorithm consists in locating the eyes.

The eyes location is based on the quadrilateral region provided by the Viola-Jones face detection algorithm [21]. Knowing the proportions of the human face, a rectangular region which delimits the eyes can be heuristically estimated. Since the facial region has width \( W_{face} \) and height \( H_{face} \), it was found through empiric tests that the eyes usually occupy an area fraction \( A_{eyes} \) of the facial area corresponding to the subset of pixels \( p_{i,j} \), such that:

\[
p_{i,j} \in A_{eyes} \forall \begin{cases} 
    i \in [0.15W_{face}, 0.85W_{face}] \\
    j \in [0.33H_{face}, 0.50H_{face}]
\end{cases}
\]

Having the upper left corner of the facial region as \((0,0)\) reference. After defining the rectangular eyes region, it it divided in two equal parts so that each one contains one of the eyes. The center points of the two resulting quadrilaterals
Fig. 2. Estimation of eyes location. Variables $W$ and $H$ respectively correspond to width and height of the face region.

Fig. 3. Calculation of S Region according to the distance between the eyes. This process is illustrated in Fig. 2.

Fig. 4. Some steps of the algorithm: Region S selection (a), Cr band isolation (b), After binarization (Cr$_{edge}$) (c), Edges (d), Cr band (e), Edges (f).

VI. EXPERIMENTS AND RESULTS

This section presents the performance evaluation methodology and experimental results. The proposed algorithm was...
evaluated by calculating EER for both AR and FRGC benchmarks. Detailed information is presented ahead.

Glasses detection in facial images has two possible classifications only: with glasses or without glasses. The output of the process depends on a previously established threshold. If the value of $P_G$ is higher than the threshold, the sample is classified as "with glasses", otherwise the face is classified as "without glasses". Under these circumstances, two kinds of error may occur:

1) **False Acceptance**: classify an image that shows no glasses as "with glasses";
2) **False Rejection**: classify an image that shows glasses as "without glasses".

Changing the threshold in a predefined range allows calculate the false acceptance and false rejection for each value of threshold, so, making possible to plot the curves of False Acceptance Rate (FAR) and False Rejection Rate (FRR). The meeting point of the curves is called Equal Error Rate (EER), it is the point where the previous rates are proportionally equal. The EER points out that the ratio of false acceptances detected divided by the total number is equal to the same ratio of false rejections.

Accordingly, low values of EER suggest false acceptance and false rejection rates equally low, which implies a high accuracy to the system. The curves of FAR and FRR correspond to the false acceptance and false rejection ratios for each threshold on the horizontal axis. The meeting point of the curves is the EER point. Tests performed on the image database AR got a success rate of 98.58%, whereas on the FRGC the success rate was 89.33%. Table I show the results of the proposed method in each database and works discussed in Section II, to establish a comparative. The FAR x FRR charts are presented in Fig. 6 for images from both AR and FRGC Face Databases.

VII. CONCLUSIONS AND FUTURE WORKS

This paper presented a novel method for accomplishing the detection of eyeglasses in frontal face images. The method is based on YCbCr color analysis, once it is a widely used for skin segmentation. It analyzes the region between the eyes (called here by region $S$), where the nose piece of the eyeglasses must be present. This region is defined according to the eye pupils location, which is calculated according to the human face proportions and in the face region delimited by Viola-Jones face detector. The method explores color information contained in Cr band from YCbCr color model and edges detection to accomplish the detection of the nose piece.

The Algorithm presented EER results very close to the other methods present in literature, achieving correct classification rate equals to 98.58% in AR Database benchmark and 89.33% in FRGC database. Some advantages of the proposed method concerning other state-of-the-art algorithms are the computational simplicity and the absence of a previous machine learning. Although the use of Cr band from YCbCr color model it was verified that the method suffers a bad influence of shadows in region $S$. But this problem only occurs in face images from FRGC Face Database which present faces in uncontrolled conditions. So some errors are
acceptable once working with challenging image samples from uncontrolled environments. Fig. 5 presents some samples which were wrongly classified because of bad lighting condition.

As future work, it would be important to improve the eyes localization process and S Region delimitation. It is possible to locate both eyes with bigger precision by replacing actual eye localization method, which depends on the face detection, by facial landmark detector. It enables better eyes localization and, hence, better calculation of Region S. Another possible improvement is to investigate a way of reducing lighting influence in the proposed algorithm.

REFERENCES