IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 21, Issue 2, Ser. IV (Mar - Apr 2019), PP 39-48
www.iosrjournals.org

Implementation of a VLC correlator for face identification in a real-life application

Nicolas Gaborel, Marwa Elbouz, Ayman Alfalou
Yncréa Ouest, 20 rue Cuirassé Bretagne, Brest, France

Abstract: *As part of the national French Safe City project to handle access and security in schools, identification methods were implemented on an embedded system. The contract specifications required self-contained, fast, and reliable face identification via on-site cameras. To answer these constraints the Vander Lugt Correlator architecture was selected as it provides a compromise between implementation complexity, robustness and discrimination and is an efficient way to check for similarities. As is, VLC identification is limited for a number of cases including illumination variation, strong discrimination (robustness against impostors) and face rotation. To overcome these limitations, we propose and validate a pre-processing method to optimize VLC correlation plane results. This pre-processing can be used in conjunction with existing filter designs and post-processing techniques, which are widely optimized. Based on histogram equalization and phase description, this method demonstrates its efficiency on the Extended Yale Face Database B and in our experimental real-life tracking and identification embedded system.*

Keywords: *experimental implementation, correlation, pre-processing, VLC, phase-only filters, gradients, histogram equalization.*

Date of Submission: 15-04-2019 Date of acceptance: 30-04-2019

I. Introduction

Robust face recognition is the basis for smart identity and access management systems.

In most cases, to optimize correlation results for face identification using (VLC), new specific filters are created [1]. Such correlation filters can be used in low light scenarios, with compressed images, etc. [2] These filters are use-case specific and cannot adapt to new situations. Additionally, work has been conducted on the decision-making phase, developing post-processing methods to optimize decision variables such as the PCE [3], or to analyze frequency-domain spectrums [4] and therefore avoiding additional FFTs, which slow down processing. Some of these solutions often come at a price, losing useful information, while others require the use of multiple filters for a single reference, bringing the computation time and storage requirements above acceptable limits for large scale identification [2,5,6,7]. There is a need for a solution able to adapt to changes such as lighting, distortion (face perspectives and expressions) and rotation without resorting to multiple case-specific filters. We therefore present a pre-processing method aimed at enhancing the useful information in pictures, providing a general solution to weak identification results, regardless of filter choice or post-processing technique. For its simplicity, broad range of uses, and adaptability, the composite phase-only VLC filter was used in this paper [2]. What is validated for a phase-only filter will also apply for other correlation filters, making this method versatile.

The Vander Lugt Correlatoris a simple techniquebased on the use of a correlation filter [8]. Indeed, it uses only two Fourier transforms to extract relevant information from 2D scenes, it is therefore easy to implement in a computer vision algorithm and considered low-cost in processing power in view of today's technology. A correlation peak is generated after the inverse Fourier transform of the multiplication of two spectrums in the frequency domain (1).

$$CP = FT^{-1}(spectrum_{target}\,(f) \times filter(f))(1)$$

Where CP is the correlation plane, FT is the Fourier transform and $spectrum_{target}\,(f)$ is the target image in the frequency domain.

If the two spectrums are similar, a central peak with higher energy than its surroundings will appear. However, because face structures are similar, there is always a small correlation peak. This can be a problem when making sure that people who are not in the database are not identified to their closest match (impostor). To improve the discrimination of the correlation technique, its phase-only filter design is used. It produces sharper correlation peaks and increases discrimination [9].In order to keep a reasonable recognition rate, a face recognition system using correlation filters for identification should be tolerant of changes while maintaining discrimination to prevent mis-identification. Typically, this is not fully accomplished with simple VLC or phase-only VLC systems.

Even when lighting conditions are favorable, image intensities from static cameras change throughout the day. In a typical scenario where a face identification database is static (i.e. where photos and filters for individuals are not updated), it is important to find a way to extract only the relevant parts of the images. In the continuity of the line of thought that it is more efficient to use image descriptors [10,11]rather than greyscale or color intensities to draw conclusions from a scene, phase and especially image gradients were studied. Indeed, phase is the most important data in correlators such as the VLC [6,7,8]. Therefore, we need to revisit the phase reconstruction method in order to increase the recognition ratio.

First, to validate our method, tests on the Extended Yale Face Database B [12] were conducted. This database provides cropped face images of 38 individuals under various lighting conditions. This allows us to test VLC identification with unconstrained lighting, as it would be in a real-world environment. Depending on the time of day, the sun can shine brightly at an angle, casting strong shadows on faces. The Yale database adequately represents this use-case. Figure 1 shows examples from the database.

Figure 1 - 9 *examples of lighting conditions from the Extended Yale Face Database B.*

As expected, Figure 2 shows phase-only VLC filters do not perform well on the Extended Yale Face Database B. These ten ROC curves were generated using ten different lighting conditions to create the filters, therefore they also reflect the impact of filter creation on identification.Ten different reference images were used to make sure the results are not biased by a particular lighting condition. With this method of identification, if we choose to work with 0% false positive rate, the true positive rate is barely more than 40%. If implemented, the resulting system would be slow to identify individuals, often giving false negatives.

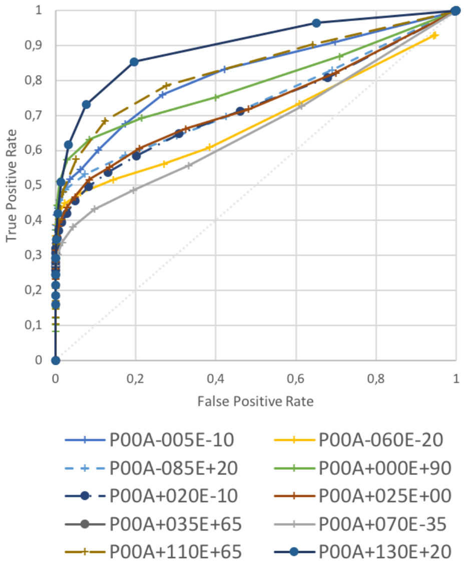

Figure 2 - 10 *ROC curvesgenerated using ten different lighting conditions,describing VLC identification performance on the Yale database without pre-processing.*

Our goal in designing a pre-processing method was to provide fast and reliable face identification with an Nvidia Jetson embedded system as part of a smart video surveillance contract(SafeCity [13]).Real-world scenarios often include unconstrained illumination and appearance changes, making them more challenging than static face databases.

The next section describes how image data can be optimized to boost correlation results.Section II-1 focuses on light invariance while section II-2 focuses on discrimination with phase optimization and II-3 summarizes the algorithmic implementation. InIVwe present results obtained with our new pre-processing method, including validation on the Yale database. Section IV, experimental implementation, covers the implementation of the pre-processing method and VLC identification in an embedded system. Finally, we conclude and discuss in the last section.

II. Correlation Data Optimization

II-1 Illumination Changes

One of the most commonly varying properties we can find between scenes is illumination. This is especially true for real-world scenarios (outdoors access management, for example) where sunlight is relied on. To overcome this recurring problem, a simple histogram equalization[14] step is added to our optimization workflow (see Figure 7). As shown onFigure 3, histogram equalization can cope with extreme illumination changes and make low-light images usable.It also deals with harsh and uneven shadows such as those presented in the Yale database.

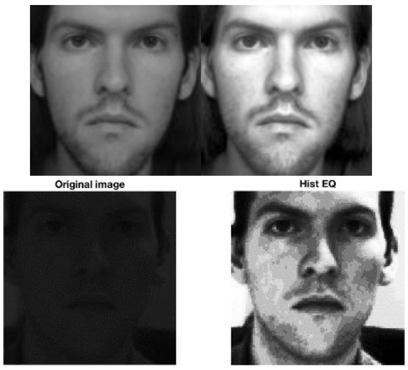

Figure 3–*Two histogram equalization examples. Top row: the original image has good illumination. Bottom row: the original image is under-exposed. Image: Nicolas Gaborel, one of the researchers.*

Table 1 shows that, although mandatory in pre-processing for poor light conditions our preprocessing optimized stepis not enough to significantly raise discrimination levels between two targets.

Method	PCE *true*	PCE *false*	Ratio
No pre-processing	0.0026	0.00026	10
Histogram EQ	0.0059	0.00055	10.7

Table 1 – *PCE comparison for person 1 vs. person 2 in Yale database.*

The ratio in the above table is calculated by dividing the true identification PCE by the false identification PCE, using (for example) the first and second person in the Yale database. Applying histogram equalization during the filter creation and comparison phase, we find only 7% discrimination increase between true and false (affecting identification using PCE criterion). The PCE, or peak-to-correlation energy, is the ratio between the total correlation plane energy and the peak's energy within 3x3 pixelsas in equation (2)[15].

$$PCE = \frac{pea\,k_{energy}}{plan\,e_{energy}} (2)$$

Identification is usually decided with the calculated PCE value against a threshold, the higher the PCE, the more resemblance between the two images.While illumination variation is dealt with, there is still potential for false identification because of the resemblance between images. Here, the PCE is used because it is a simple criterion, alternate decision metrics can be used [16] but they require more processing power.

II-2 Phase optimization

The correlation process in VLC identification makes use of phase information in images to compare two targets[17, 9]. In order to optimize these results, edge gradients are computed and their magnitude is used instead of light intensities. To do so, vertical and horizontal edges are detected with Sobel convolution kernels[18]. Square matrix kernels are convoluted with the target image, using derivative approximation to find edges. Deriving intensities along the horizontal and vertical axes results in high values around transitional areas (light to dark) and low values in areas where there is little change. The resulting approximated gradient components for the x and y directions (G_x and G_y) are then used in the following equation (3) to compute gradient magnitudes.

$$M = \sqrt{G_x^2 + G_y^2} \qquad (3)$$

Sobel convolution kernels can have various sizes (3x3, 7x7, etc.) which change the way intensity variation is detected. Smaller kernels produce thinner edges, while large kernels produce thicker ones and are less sensitive to minute changes (see Figure 4).

Figure 4–*Sobel gradient detection comparison with 3x3 pixel kernels [left] and 7x7 pixel kernels [right].*

Kernel size was chosen after several identification tests. ROC curves in Figure 5were generated by varying the PCE decision threshold. A kernel size of 7x7 pixels was chosen for our application as it gave the best true positive rate at around 0% false negative rate (non-dominant facial features are discarded in the gradient detection). Non-dominant facial features include small color/intensity shifts, wrinkles, fluff/hair, and other skin asperities as opposed to the shape of the mouth, the eyes, the nose and other bone structure-related visual characteristics.

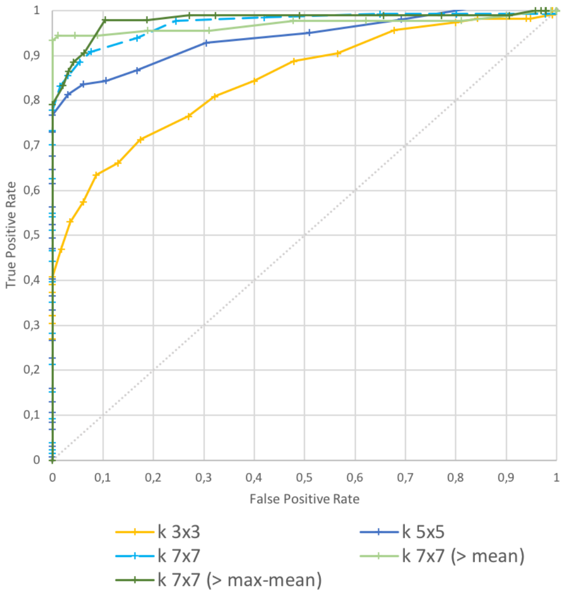

***Figure 5** - ROC curves of VLC identification results for different convolution kernel sizes.*

In the above figure, *"k 7x7 (>mean)"* represents a convolution kernel size of 7x7 pixels with gradient magnitude thresholding keepingonly the values greater than the mean magnitude in the image, as in equation (4).*"k 7x7 (>max-mean)"* means the threshold was set to the maximum minus the minimum gradient magnitude in the image.

$$Im_{thr} = Im_{gr}\,(Im_{gr} > Threshold\,)$$
$$Threshold = \begin{cases} mean(Im_{gr}) \\ max(Im_{gr}) - mean(Im_{gr}) \end{cases} \tag{4}$$

Where Im_{thr} is the thresholded image and Im_{gr} is the gradient image.

Even though smaller Sobel convolution kernel sizes give more detail, as shown in Figure 5, more discrimination is obtained with larger kernels. Focusing on strong features and facial structure allows the VLC identification to be more robust and less influenced by small-scale changes.

II-3 Implementation (illumination, phase, rotation)

To cope with illumination changes, rotation and low discrimination, the methods presented previously are combined to form our pre-processing technique. First of all, a cropped face image is extracted from a frame and used as the input of our algorithm. Histogram equalization is applied for light invariance. Gradients are then estimated with the Sobel operators and their magnitude is calculated as detailed in section II-2. The magnitude Min equation (3) for each pixel is then thresholded to remove low-magnitude gradients. Low-magnitude gradients typically represent small features and color shifts on the skin; strong gradients are found around the main descriptive features (eyes, nose, mouth, etc.). All gradient magnitudes lower than the image's mean magnitude are set to 0 (see Figure 6). Using the mean as a threshold allows for adaptability between images and has shown better results than other values (see Figure 5).

***Figure 6 –** [Left] Gradient magnitudes. [Right] Gradient magnitudes after thresholding.*

The thresholded magnitude is used as the input of the VLC process, both for the reference filter creation and as pre-processing done on target images. Figure 7reviews each step before the VLC identification process and presents the step-by-step result on a sample image.Each reference image saved in the database as frequency information (after an FFT) is subjected to the same pre-processing steps as the input images. To overcome face rotation problems, the composite VLC filter is used[19, 20]. Several images with different poses (front-facing as well as pointing left, right, up, and down at an angle between 5 and 10 degrees) are processed and their phase information is summed into one frequency-domain image. This image is normalized and saved in the database to be used against target images for comparison and identification. Filter saturation is avoided by using high bitrate containers (matrix "type" in OpenCV) which do not significantly slow down the implementation. The resulting pre-processing method is a succession of fast and simple steps to focus on and enhance facial features.

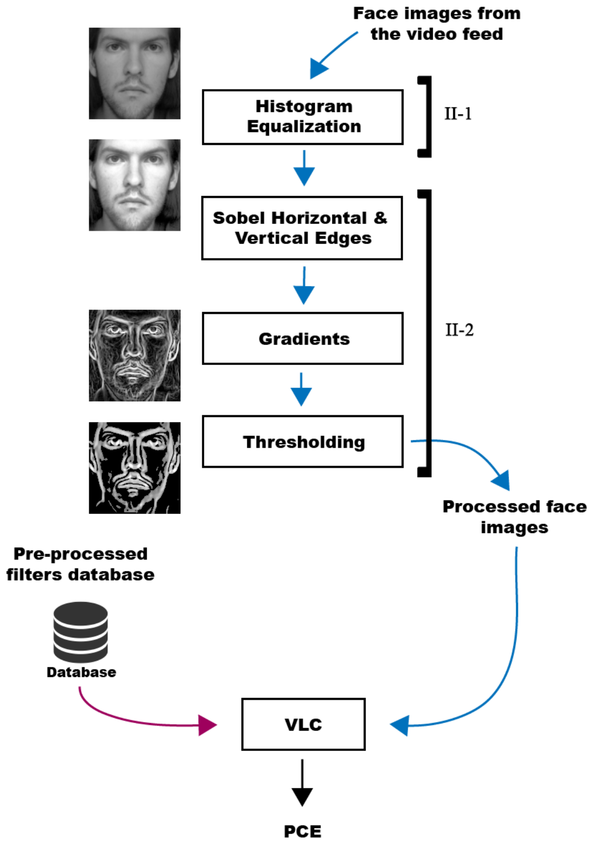

Figure 7 - *Pre-processing algorithm synoptic.*

III. Results

III-1 Validation

Our pre-processing method was first tested on the Extended Yale Face Database B with phase-only (POF) VLC and then on our own database for real-time identification using composite phase-only VLC. In order to validate our pre-processing method for illumination and enhanced discrimination, 10 Yale database images representative of the various illumination possibilities were selected and used to create 10 references.

Multiple references were used to avoid biased results from a single filter and to give the reader a general idea of the performance on this database. VLC identification against those 10 references was independently performed on all images in the database. Pre-processing was performed on all images. The PCE decision threshold was raised from the minimum to the maximum value to generate identification results for ROC curves. Figure 2 shows the ROC curves for VLC identification on the Yale database with no pre-processing, Figure 8 shows the results after our pre-processing method was applied according to the algorithm synoptic (see Figure 7).

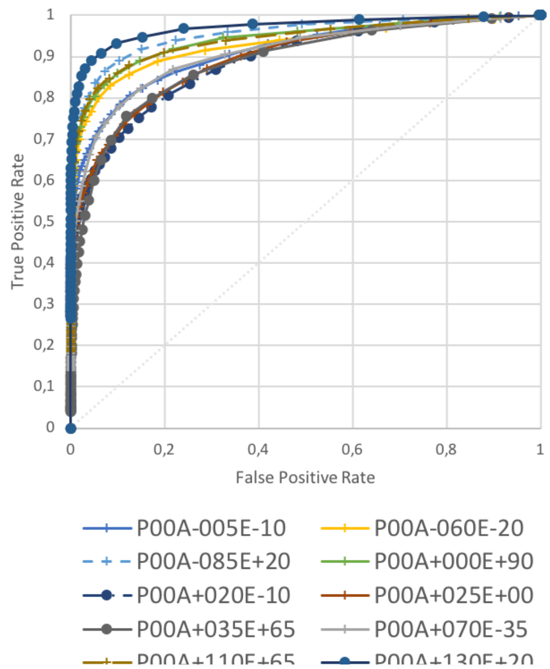

Figure 8 - 10 *ROC curves generated using ten different lighting conditions, describing VLC identification performance on the Yale database with pre-processing.*

Figure 9 takes the best and worst ROC curves with and without pre-processing on the Yale database to allow quick comparison. Our pre-processing method significantly raises true positive rates and gives more robust identification for faces with challenging illumination.

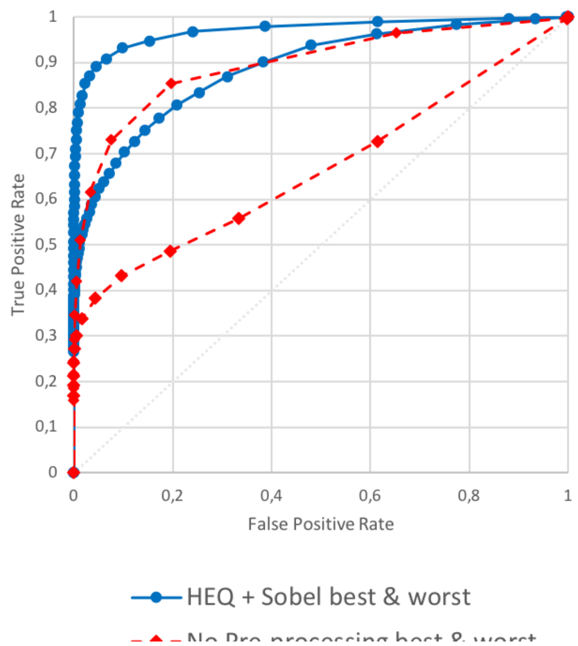

Figure 9 - *ROC curves describing the best and worst VLC identification performance on the Yale database with and without per-processing.*

III-2 Experimental Results

To quantify the effectiveness of our method on our real-case identification system, a video clip (640x480 @ 25 FPS) of a person facing the camera and moving around was recorded. The person's face was automatically detected, tracked, and compared to a 23-person face database for identification every frame. The results in Figure 10 are the effectiveness of the VLC identification with 4 different pre-processing methods, namely, histogram equalization, optimized phase contrast [21], optimized histogram equalization + gradients (HEQ+S), and no pre-processing to give a starting point to the comparison.

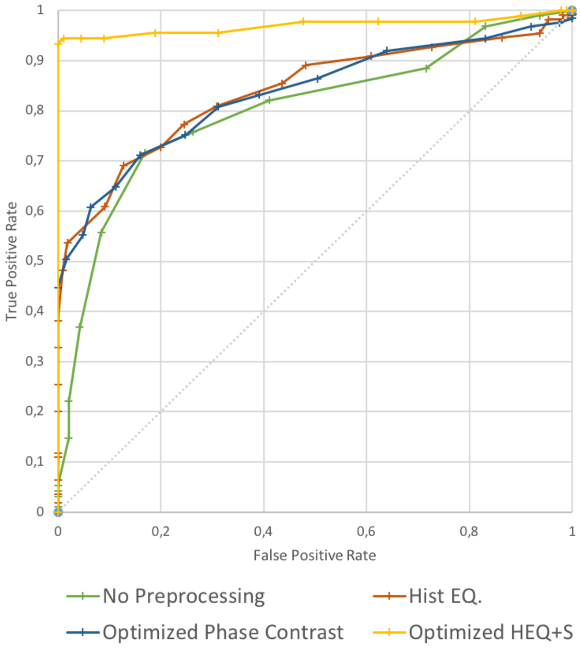

Figure 10 *- ROC curves of VLC identification results for four different pre-processing methods applied to live video feeds.*

The *No Pre-processing* curve demonstrates that simple VLC identification is not robust and is prone to false positives while histogram equalization and optimized phase contrast provide a higher true positive rate at a very low false positive rate. Our method out-performs the other three by a significant margin, providing a 95% true positive rate with close to no false positives. This was designed to answer our contract's requirement that no person not registered in the database should be allowed inside the perimeter. All this while keeping a high frame-rate, relative to human perception, on an embedded system.

IV. Experimental implementation

In this section, we go over the implementation of our pre-processing method for VLC identification on an Nvidia Jetson embedded system. The Jetson TX1 card was chosen for its processing power and small form factor as well as its ability to handle multiple cameras simultaneously. A MIPI CSI-interface camera from Leopard Imaging was used to capture the live footage. This interfacing allows for separate hardware management of the video feed and avoids over-burdening the CPU as is the case with USB-interface cameras. All processing is done locally and either displayed live via the HDMI port or saved in various text files. The image processing methods are implemented in C++ using the OpenCV computer vision library. To allow for multiple target tracking and identification, a multi-thread architecture was developed. The main thread displays the live video feed and various information while the other threads work jointly to 1) detect 2) track 3) identify. Both the pre-processing and VLC identification happen in the identification thread. Pre-processed images of authenticated persons stored locally are used in the VLC. The system runs in real-time compared to human perception, the mean frame-rate stands around 20 FPS.

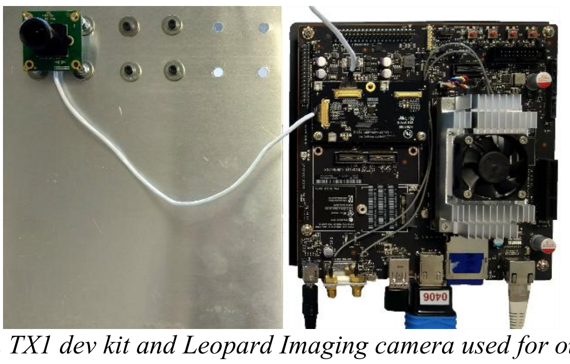

Figure 11 *- Nvidia Jetson TX1 dev kit and Leopard Imaging camera used for our embedded smart video surveillance system.*

A video clip of the real-time experimental results is available as supplementary material and at the following link (https://youtu.be/Iox4ikN6ZHo). This video was recorded via a screen capture of our embedded system's output. Figure 12 shows sample frames from the supplementary material.

***Figure 12 -** Sample frames from our supplementary material.*

In this video, one of the researcher's face and eyes are automatically detected and tracked with simple optical flow methods. The tracked area is then subjected to VLC identification against a local database. The decision (person's name) is displayed when the identity is confirmed. The researcher moves so as to test cases for rotation, scale and illumination variation. At the end of the video, because of a sudden light change, the tracker drifts, the face is detected again and the ID is set to unknown before it is confirmed once more.

V. Conclusion

We presented a pre-processing algorithm to optimize VLC identification robustness against false positives while increasing true positive rate. This work exports face recognition techniques from simulations to a real-life application. Histogram equalization alone was first tested but showed no significant improvement to identification robustness. Gradient computation via Sobel convolution kernels was then introduced to extract useful information from the images, gradient magnitudes were thresholded to discard small feature changes and to focus on strong traits. The combination of the two techniques proved its efficiency on both a well-known face database and a real-life application. This method is simple and comes at a low processing cost, it is therefore suitable for undersized systems and low power applications. The concept of using descriptors such as gradients (or their magnitude, here) instead of image intensities to draw conclusions from an image follows the ongoing trend in computer vision research that suggests that using higher-level information is more efficient than using raw pixel values.In experimental implementation, our pre-processing method raises true positive rate by a factor of two and very significantly increases discrimination when compared to the other assessed methods. A video submitted as supplementary material demonstrates the effectiveness and speed of our method on a live video feed (https://youtu.be/Iox4ikN6ZHo).

References

[1]. A. Quaglia, C.M. Epifano, eds., Face recognition: methods, applications and technology, Nova Science, New York, 2012.

[2]. Q. Wang, A. Alfalou, C. Brosseau, New perspectives in face correlation research: a tutorial, Advances in Optics and Photonics. 9 (2017) 1. doi:10.1364/AOP.9.000001.

[3]. A. Alfalou, A. Kouas, M. Elbouz, Y. Ouerhani, C. Brosseau, A. Fakhfakh, F. Ghozzi, A new correlation criterion based on energy distribution: face recognition application, in: M.S. Alam (Ed.), Pattern Recognition and Tracking XXIX, SPIE, Orlando, United States, 2018: p. 42. doi:10.1117/12.2312500.

[4]. M. Jridi, T. Napoléon, A. Alfalou, One lens optical correlation: application to face recognition, Applied Optics. 57 (2018) 2087. doi:10.1364/AO.57.002087.

[5]. J. L. Tribillon, Corrélation Optique, Teknéa, Toulouse, 1999.

[6]. A.A.S. Awwal, What can we learn from the shape of a correlation peak for position estimation?, Applied Optics. 49 (2010) B40. doi:10.1364/AO.49.000B40.

[7]. A.A.S. Awwal, M.A. Karim, S.R. Jahan, Improved correlation discrimination using an amplitude-modulated phase-only filter, Applied Optics. 29 (1990) 233. doi:10.1364/AO.29.000233.

[8]. A.V. Lugt, Signal detection by complex spatial filtering, IEEE Transactions on Information Theory. 10 (1964) 139–145. doi:10.1109/TIT.1964.1053650.

[9]. J.L. Horner, P.D. Gianino, Phase-only matched filtering, Applied Optics. 23 (1984) 812. doi:10.1364/AO.23.000812.

[10]. M. Savvides, B.V.K.V. Kumar, P.K. Khosla, Eigenphases vs eigenfaces, in: Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., IEEE, Cambridge, UK, 2004: p. oo810–oo813 Vol.3. doi:10.1109/ICPR.2004.1334652.

[11]. Lu X, Image analysis for face recognition. Personal notes, May:36 (2003).

[12]. A.S. Georghiades, P.N. Belhumeur, D.J. Kriegman, From few to many: illumination cone models for face recognition under variable lighting and pose, IEEE Transactions on Pattern Analysis and Machine Intelligence. 23 (2001) 643–660. doi:10.1109/34.927464.

[13]. Sécurisation des villes intelligentes, Vipress. (2018). http://www.vipress.net/securisation-des-villes-intelligentes-109-millions-deuros-pour-le-projet-safecity/.

[14]. S.M. Pizer, E.P. Amburn, J.D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J.B. Zimmerman, K. Zuiderveld, Adaptive histogram equalization and its variations, Computer Vision, Graphics, and Image Processing. 39 (1987) 355–368. doi:10.1016/S0734-189X(87)80186-X.

[15]. B.V.K. Vijaya Kumar, L. Hassebrook, Performance measures for correlation filters, Applied Optics. 29 (1990) 2997. doi:10.1364/AO.29.002997.

[16]. A. Alfalou, C. Brosseau, P. Katz, M.S. Alam, Decision optimization for face recognition based on an alternate correlation plane quantification metric, Optics Letters. 37 (2012) 1562. doi:10.1364/OL.37.001562.
[17]. A.V. Oppenheim, J.S. Lim, The importance of phase in signals, Proceedings of the IEEE. 69 (1981) 529–541. doi:10.1109/PROC.1981.12022.
[18]. M. Juneja, P.S. Sandhu, Performance Evaluation of Edge Detection Techniques for Images in Spatial Domain, International Journal of Computer Theory and Engineering. (2009) 614–621. doi:10.7763/IJCTE.2009.V1.100.
[19]. B.V.K. Vijaya Kumar, Tutorial survey of composite filter designs for optical correlators, Applied Optics. 31 (1992) 4773. doi:10.1364/AO.31.004773.
[20]. A. Al Falou, G. Keryer, J.-L. de Bougrenet de la Tocnaye, Optical implementation of segmented composite filtering, Applied Optics. 38 (1999) 6129. doi:10.1364/AO.38.006129.
[21]. C.S. Yelleswarapu, S.-R. Kothapalli, D.V.G.L.N. Rao, Optical Fourier techniques for medical image processing and phase contrast imaging, Optics Communications. 281 (2008) 1876–1888. doi:10.1016/j.optcom.2007.05.072.

IOSR Journal of Computer Engineering (IOSR-JCE) is UGC approved Journal with Sl. No. 5019, Journal no. 49102.

Nicolas Gaborel. " Implementation of a VLC correlator for face identification in a real-life application." IOSR Journal of Computer Engineering (IOSR-JCE) 21.2 (2019): 39-48.