

Optical Engineering

SPIDigitalLibrary.org/oe

Adaptive nonlinear fringe-adjusted joint transform correlator

Isabelle Leonard
Ayman Alfalou
Mohammad S. Alam
Andreas Arnold-Bos

Adaptive nonlinear fringe-adjusted joint transform correlator

Isabelle Leonard
Ayman Alfalou

ISEN Brest
Département Vision, L@bISEN, 20 rue Cuirassé
Bretagne
CS 42807, 29228 Brest Cedex 2, France
E-mail: ayman.al-falou@isen.fr

Mohammad S. Alam

University of South Alabama
Department of Electrical and Computer
Engineering
SH 4122, 150 Jaguar Drive
Mobile, Alabama 36688-0002

Andreas Arnold-Bos

Thales Underwater Systems
Route de Sainte Anne du Portzic
CS 43814, 29238 Brest Cedex 3, France

Abstract. An optimized technique, based on the fringe-adjusted joint transform correlator architecture, is proposed and validated for rotation invariant recognition and tracking of a target in an unknown input scene. To enhance the robustness of the proposed technique, we used a three-step optimization. First, we utilized the fringe-adjusted filter (H_{FAF}) in the Fourier plane, then we added nonlinear processing in the Fourier plane, and, finally, we used a new decision criterion in the correlation plane by considering the correlation peak energy and the highest peaks outside the desired correlation peak. Several tests were conducted to reduce the number of reference images needed for fast tracking, while ensuring robust discrimination and efficient tracking of the desired target. Test results, obtained using the pointing head pose image database, confirm robust performance of the proposed method for face recognition and tracking applications. Thereafter, we also tested the proposed technique for a challenging application such as underwater mine detection and excellent results were obtained. © 2012 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: [10.1117/1.OE.51.9.098201](https://doi.org/10.1117/1.OE.51.9.098201)]

Subject terms: correlation; nonlinear fringe-adjusted joint transform correlator; face recognition; underwater mine recognition.

Paper 120308P received Mar. 3, 2012; revised manuscript received Jul. 6, 2012; accepted for publication Jul. 13, 2012; published online Sep. 14, 2012.

1 Introduction

Correlation is a robust and efficient pattern recognition technique, which was widely studied in the literature.¹ Correlation based discrimination measures the degree of similarity between a target image (unknown image to be recognized) and a reference image (belonging to a known database). Two families of optical architectures are used by most researchers to implement the correlation technique which are the Vander-Lugt correlator (VLC),² and the joint transform correlator (JTC).³ Unlike the JTC, VLCs require the implementation of a complex-valued filter. The optical implementation of the JTC architecture does not need to align the optical setup, especially in the Fourier domain between the filter and the image target spectrum (VLC). Moreover, the reference images can be updated in near real time.

In this paper, we utilized the JTC-based correlation architecture to realize a near real-time face tracking system. The proposed system must be robust to changes in in-plane rotation of the target, while ensuring accurate discrimination between multiple targets. The main objective, of this paper, is to accurately recognize and track the desired target, such as the face of person in a given scene or to recognize underwater mines from video sequences.

To enhance the performance of the proposed tracking system, we used an optimized version of the JTC architecture, namely the fringe-adjusted JTC (FJTC).⁴ This fringe-adjusted JTC provides excellent correlation discrimination, but the robustness of FJTC may not be sufficient for some applications. Consequently, we added a nonlinearity function to the fringe-adjusted JTC architecture to enhance the

robustness and discrimination capability. In addition, special attention is given to generate an optimized correlation output by applying a decision criterion in the correlation plane. The peak to correlation energy (PCE) criterion is not very suitable for automatic decision making.^{5,6} The presence of autocorrelation peak, and the presence of two correlation peaks, can perturb the decision criterion. To address this limitation, we used an adaptive FJTC architecture to suppress the zero-order autocorrelation peak. Since the correlation plane is symmetric, we work on one half of the correlation plane and modify the criterion to obtain values included in the interval [0 1]. Test results are presented to verify the effectiveness of the proposed technique.

2 Analysis

The JTC technique is one of the most widely used correlation methods reported in the literature.¹ This technique has been adapted and used in many recognition and/or tracking applications.^{1-4,7,8} In this work, we have used this JTC method for a very demanding application of underwater mines recognition. The underwater mines recognition application requires an adaptive JTC method to overcome the various problems caused by the specific conditions of underwater imaging. We begin this section by reviewing the principles of several main JTC architectures. The principle of the JTC correlation architecture has been described in detail in the literature.^{1-4,7,8} In the classical JTC, the known reference image, r , and the unknown input scene, i , (image to be recognized) are displayed side-by-side in the input plane of the correlator, which are separated by a distance, $2x_0$, along the x -axis between the target image and the reference image, to form the input joint image, expressed as

$$f(x, y) = i(x + x_0, y) + r(x - x_0, y). \quad (1)$$

After applying Fourier transform to Eq. (1) such as the input plane, we can record the intensity or the joint power spectrum (JPS), F , using a square law device such as a CCD detector array. In a classical JTC, an inverse Fourier transform of the JPS yields the correlation output in the output plane.³ This correlation output contains two important terms which are a central autocorrelation peak, and two cross-correlation peaks corresponding to the crosscorrelation between the unknown input scene, i , and the known reference image, r . These two cross-correlation peaks are located on either side of the central zero-order peak separated by a distance, $4x_0$.

The classical JTC provides broad correlation peaks and strong side lobes which can complicate the decision making. Several JTC architectures have been developed to solve this problem such as the binary JTC, nonzero-order JTC, and fringe-adjusted JTC.^{4,7,9} In a binary JTC,⁷ the JPS is binarized using a suitable threshold before applying the inverse Fourier transform step. The binarized JPS is defined as

$$F_{\text{binary}}(\mu, \vartheta) = \begin{cases} +1 & \text{if } F(\mu, \vartheta) \geq \text{threshold} \\ -1 & \text{if } F(\mu, \vartheta) \leq \text{threshold} \end{cases}. \quad (2)$$

This architecture provides narrow correlation peaks. However, it is sensitive to noise, rotation and scale variations of the target, and the impact of other artifacts. Moreover, the binarization step requires on-line processing. Another interesting JTC architecture is the nonlinear JTC where a nonlinear threshold is applied to the JPS before applying the inverse Fourier transform to yield the correlation output, expressed as:³

$$F_{\text{NL}} = F^k. \quad (3)$$

In Eq. (3), the user sets the value of k depending on the application. If k is set to 1, we obtain the classical JTC. The nonlinear JTC provides superior discrimination compared to the classical JTC. It is more efficient when the value of k is lower than 0.5.^{10,11}

The last optimization we present, herein, is based on the fringe-adjusted JTC. Alam et al.^{4,10,11} proposed to introduce a real-valued filter, called fringe-adjusted filter (H_{FAF}), in the Fourier plane of the correlator, defined as:

$$H_{\text{FAF}}(\mu, \vartheta) = \frac{B(\mu, \vartheta)}{A(\mu, \vartheta) + |R(\mu, \vartheta)|^2}, \quad (4)$$

where R , denotes the intensity corresponding to the Fourier transform of the known reference image, B , is a constant or a function used to control the gain, and, A , is a function used for avoiding the pole problem and/or reducing the effects of noise. The FJTC is a very efficient architecture, very simple to implement and very convenient for many applications, especially when the noise and background are known or can be estimated. Moreover, many optimized versions of this architecture are proposed and validated.¹ However, the FJTC architecture, based on spectral filtering, makes it very sensitive to noise changes which are a very common problem in underwater images. Examples of noise in underwater images are presented in Fig. 1. Thus, to overcome this problem and to make an efficient architecture, we need to update the noise function used in the spectral filter for each image, which may become a bottleneck for real-time applications. Thus, it is necessary to make optimizations to obtain a method such that it is less sensitive to noise encountered in underwater imagery. To achieve this objective, we introduced a special image preprocessing step and added nonlinearity to the JTC architecture to make it more robust, while ensuring excellent discrimination.

Another class of architecture, called the multi object JTC or nonzero-order JTC, has been found to be very useful for enhancing the correlation discrimination.^{9,11} In this architecture, the reference-only and the input-scene-only power spectra are subtracted from the JPS in the Fourier domain, given by:

$$\begin{aligned} |F_{\text{NZ}}(\mu, \vartheta)|^2 &= |F(\mu, \vartheta)|^2 - |R(\mu, \vartheta)|^2 - |I(\mu, \vartheta)|^2 \\ &= |R(\mu, \vartheta)|^2 - |I(\mu, \vartheta)|^2 \\ &\quad + 2|R(\mu, \vartheta)||I(\mu, \vartheta)| \cos(\phi_i(\mu, \vartheta) - \phi_r(\mu, \vartheta) \\ &\quad + 2ux_0) - |R(\mu, \vartheta)|^2 + |I(\mu, \vartheta)|^2 \\ &= 2|R(\mu, \vartheta)||I(\mu, \vartheta)| \cos(\phi_i(\mu, \vartheta) \\ &\quad - \phi_r(\mu, \vartheta) + 2ux_0), \end{aligned} \quad (5)$$

where I represents the input-scene-only power spectrum. In this technique, the zero-order term and the autocorrelation term, among similar input scene targets or nontarget objects, are automatically eliminated while leaving the desired cross-correlation peaks in the output plane.

In this paper, we propose and validate a nonzero-order fringe-adjusted JTC architecture (NFJTC) to ensure robust discrimination while simplifying the decision making. It may be mentioned, that the nonlinearity increases the robustness of the JTC architecture. Thus, we applied a special

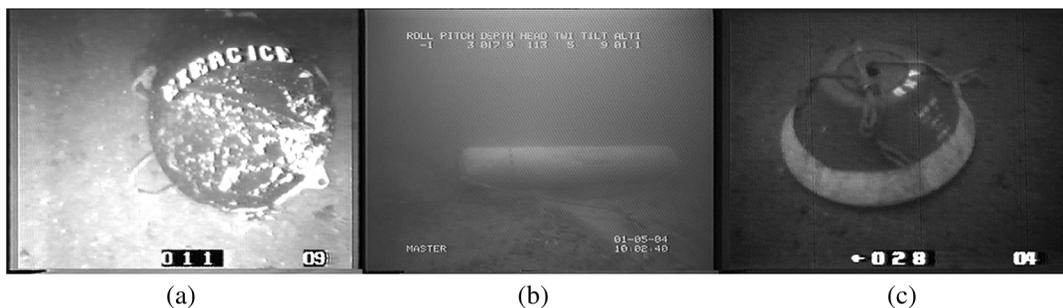


Fig. 1 Grayscale images representing three studied targets that come from underwater videos: a spherical mine (a), a cylindrical mine (b) and a Manta mine (c).

nonlinear function to the NFJTC to increase the robustness of the architecture.

To investigate the performance of the proposed technique, let us consider two practical applications. The first application involves recognizing and tracking a human face that changes according to different rotation angles. The system proposed, in this paper, allows robust tracking and discrimination of a face turning in different rotation angles ranging from -90 deg to $+90$ deg. Facial images contain information as well as noise. In general, the noise is known, and is quite similar, from one image to another in the sequence. Once our system is validated, we applied it for underwater mine detection and recognition.¹² The main problem, in this application, is the poor visibility which can limit the visibility to only a few meters or less in turbid waters.

3 Challenges of Underwater Imaging

A generic underwater image formation model is shown in Fig. 2.¹³ In Fig. 2, light source (1) transmits light (2) towards the object (6). On the trajectory to and from the object, light is intercepted by undesired underwater particles (3), and light rays (2), (5), (7) and (9) become exponentially dimmer due to direct attenuation. Part of the light is also reflected towards the camera by medium (3), resulting in the backscattered component (4), which corresponds to the image of the water volume. Finally, some light is slightly deviated by medium (3) on its way back to the camera, resulting in the forward scattered component (8) which adds blurring to the image. Underwater images are very noisy images. Moreover, for underwater imaging, this noise changes from one image to another. This limits the efficiency of image recording medium and requires a noise model to reduce its impact. Detection and recognition can be significantly perturbed by this noise component.

To address this challenge, we introduced a three-stage preprocessing step as shown in Fig. 3. The first stage removes the moiré effect which is generated due to the use of analog cameras. It is modeled by a sine function following the procedure outlined in Ref. 14. Weak contrast often reduces the efficiency of recognition methods. In order to improve the underwater image contrast, we adapted the preprocessing technique presented in Ref. 15. The overall algorithm involved in the implementation of the second stage is summarized in Fig. 4. The output image with enhanced contrast is calculated as:

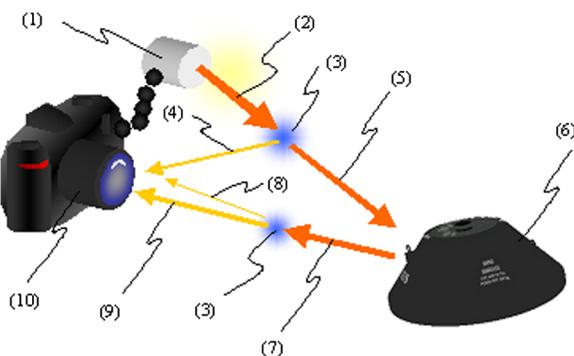


Fig. 2 Diagram explaining the light propagation in an underwater medium.

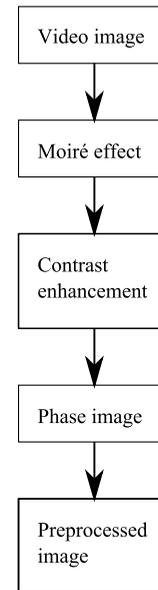


Fig. 3 Preprocessing steps to improve image quality.

$$I = \frac{I_{LP}I_s + \frac{I_d}{I_{LP}}}{I_{LP} + \frac{1}{I_{LP}}}, \quad (6)$$

where I_{LP} , represents the original image filtered with a FGaussian low-pass filter, I_s represents the result of the subtraction of the I_{LP} from the original image, and I_d represents the result of the division of the original image by I_{LP} . Details of this algorithm can be found in Ref. 15. Finally, in the third stage, we calculate the phase image and we reduce noise with a band pass filter created with a Gaussian wavelet as depicted in Fig. 5.¹²

4 Analytical Modeling

To investigate the robustness of the proposed technique, we selected several existing algorithms and compared their performance by using three images as shown in Figs. 6 to 8. The first two images came from one video sequence acquired by a diver and represents an upside down Manta mine (MN 103). These two image frames are closely located in the video (8 frames apart) where the mine-like characteristic did not change significantly. These two images were used to test the robustness of the proposed and alternate architectures. Since these two images were shot within a small time-frame and are similar, so the output of the filter should also be similar. The discrimination of our system is tested by applying JTC architectures on a mine image and the Lena image as shown in Fig. 8.

At first, we compared the classical JTC and the nonzero-order JTC on the manta mine image as shown in Fig. 6. We used image frame 1082 of the test video as the reference image and as the target image and the results are shown in Fig. 9. From Fig. 9, it is evident that the suppression of the autocorrelation peak simplifies the decision making. Moreover, this suppression reduces the effect of noise in the correlation plane.

Next, we investigated the performance of the nonzero-order JTC for the above mentioned images as shown in Figs. 6 to 8. We used PCE as the decision criterion in the nonzero-order JTC correlation plane, defined as:⁵

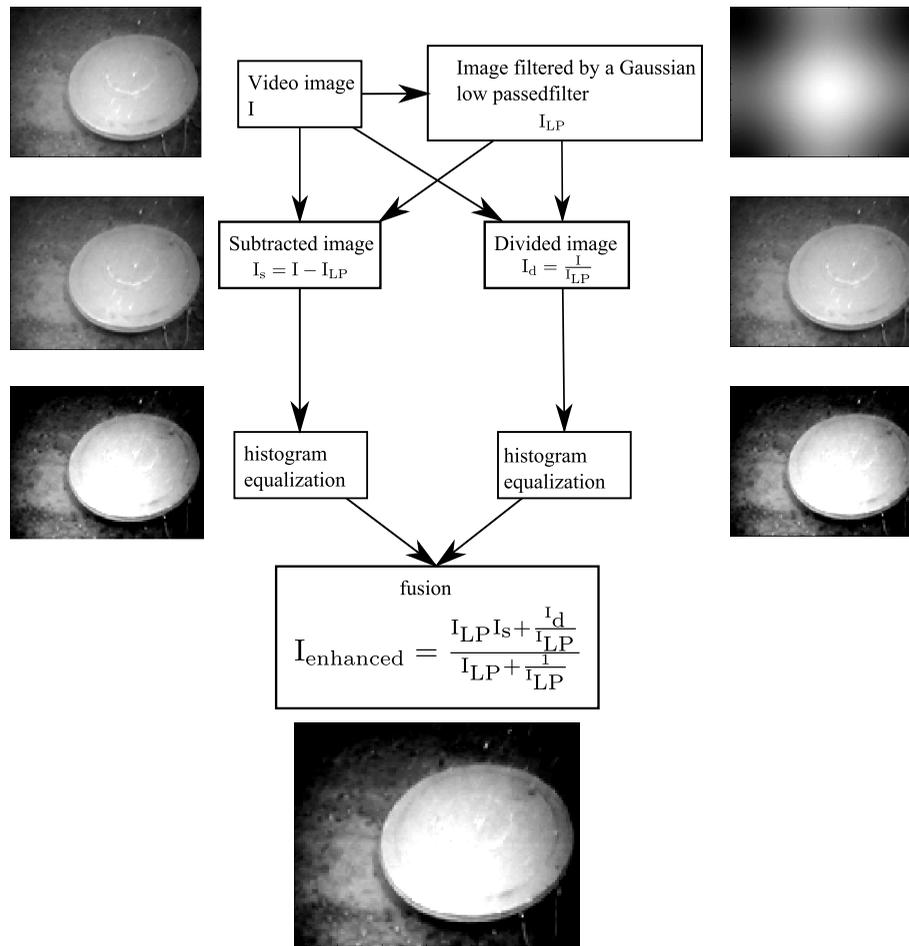


Fig. 4 Image enhancement algorithm.¹⁵

$$PCE = \frac{\sum_{x=x_0-t}^{x=x_0+t} \sum_{y=y_0-t}^{y=y_0+t} |C(x,y)|^2}{\sum_{x=1}^N \sum_{y=1}^M |C(x,y)|^2}, \quad (7)$$

where (x_0, y_0) denote the position of the correlation peak, $C(x, y)$ denotes the value of the correlation plane at point (x, y) , t is set to the number of neighboring pixels used, and (N, M) denote the size (in pixels) of the correlation plane. The corresponding results are shown in Fig. 10.

Figure 9 shows that the nonzero-order JTC architecture yields helpful results for decision making. Figure 10 shows that the fringe adjusted JTC architecture observes these constraints on PCE values better than the binary

JTC architecture. The nonlinear JTC architecture is a robust one, such that, PCE values are close between Figs. 9 and 10 as shown in the left columns. However, it does not provide robust discrimination as shown in Fig. 10. PCE values are close for the cases of (i) two mines on the left, and (ii) one mine and the Lena.

In this paper, our idea is to combine the attractive features, of fringe-adjusted JTC and nonlinear JTC, in order to obtain a robust architecture for correlation discrimination. In the proposed architecture, called nonlinear fringe-adjusted JTC (NNJTC), the JPS is modified to obtain the nonlinear fringe-adjusted JPS. The mathematical steps needed to obtain this nonlinear fringe-adjusted JPS are summarized in Eq. (8):

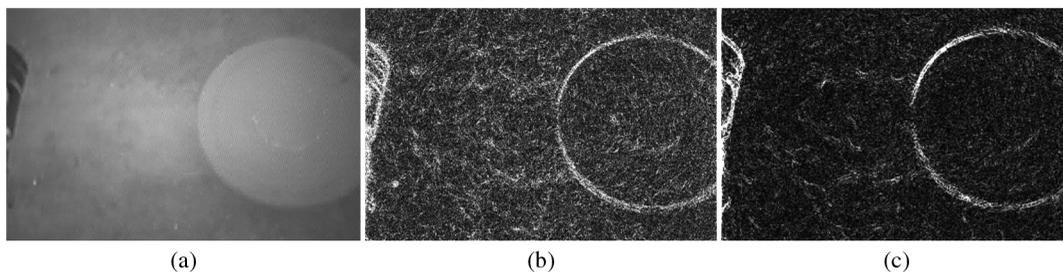


Fig. 5 Initial image, phase image, phase and wavelet image.



Fig. 6 Manta mine, image no. 1082.



Fig. 7 Manta mine, image no. 1090.

$$\begin{aligned}
 f(x, y) &= i(x + x_0, y) + r(x - x_0, y)FF(\mu, \vartheta) \\
 &= |I(\mu, \vartheta)| \exp(\phi_i(\mu, \vartheta)) \exp(jux_0) + \dots \\
 &\quad + |R(\mu, \vartheta)| \exp(\phi_r(\mu, \vartheta)) \exp(-jux_0)G(\mu, \vartheta) \\
 &= |F(\mu, \vartheta)|^2 = |I(\mu, \vartheta)|^2 + |R(\mu, \vartheta)|^2 \\
 &\quad + 2|I(\mu, \vartheta)||R(\mu, \vartheta)| \cos(\phi_i(\mu, \vartheta) \dots \\
 &\quad - \phi_r(\mu, \vartheta) + 2ux_0)|F_{NZ}(\mu, \vartheta) \\
 &= 2|I(\mu, \vartheta)||R(\mu, \vartheta)| \cos(\phi_i(\mu, \vartheta) - \phi_r(\mu, \vartheta) \\
 &\quad + 2ux_0)|G1(\mu, \vartheta) \\
 &= F_{NZ} \cdot H_{FAF}(\mu, \vartheta)G1(\mu, \vartheta) \\
 &= F_{NZ} \frac{B(\mu, \vartheta)}{A(\mu, \vartheta) + |R(\mu, \vartheta)|^2} G_{NN-FAF}(\mu, \vartheta) \\
 &= (G1(\mu, \vartheta))^k, \tag{8}
 \end{aligned}$$

where G , represents the classical JPS, F_{NZ} represents non-zero JPS, H_{FAF} represents the FAF filter, $G1$ is the nonzero-order fringe-adjusted JPS, G_{NN-FAF} is the nonlinear and the



Fig. 8 Lena.

nonzero-order fringe-adjusted JPS and k denotes the degree of nonlinearity. The value of k has been selected, after several tests, to provide the best compromise between robustness and discrimination. Before applying the NNJTC, we investigated the performance of the binary JTC and fringe-adjusted JTC for underwater images discussed in the previous section.

5 Binary JTC and Fringe-Adjusted JTC Performance Comparison

To compare the performance of nonzero-order binary JTC and the nonzero-order fringe-adjusted JTC, we consider the Lena image, Fig. 8, as the known reference image as well as the unknown target image, which is a scenario of perfect match between the reference and the target. To mimic the effects of underwater imaging, we modified the target image with underwater-like noise. At first, we added blur as if the image is perturbed by forward scattered light. Then we added nonuniform luminosity, which is a problem when the underwater robot uses artificial light. The next step was to combine the effects of blur and nonuniform light into the final target image. For the last test, we manually removed the background of the reference image. The target image is the Lena image corrupted with blur and nonuniform light.

The correlation output results are presented in Table 1. It is evident that the nonzero-order, fringe-adjusted JTC yields better results compared to the binary JTC. The FJTC is robust against different type of noise, such as underwater noise depicted in Table 1. According to these results and for underwater videos, the FJTC can provide better results compared to the binary JTC.

6 Test Results

6.1 Face Recognition

At first, we applied the fringe-adjusted JTC for face recognition and tracking application. The objective is to detect and track a person via face recognition in a video sequence as shown in Fig. 11. Figure 12 presents the results obtained with fringe-adjusted JTC where only one reference image is used to construct the fringe-adjusted filter H_{FAF} . It should be noted that the face of the subject shown in Fig. 11 without rotation. The reference image is introduced on one side of the input plane. The background is defined as information outside a square around the face in the reference image considered. Then we display the target image on the other side of

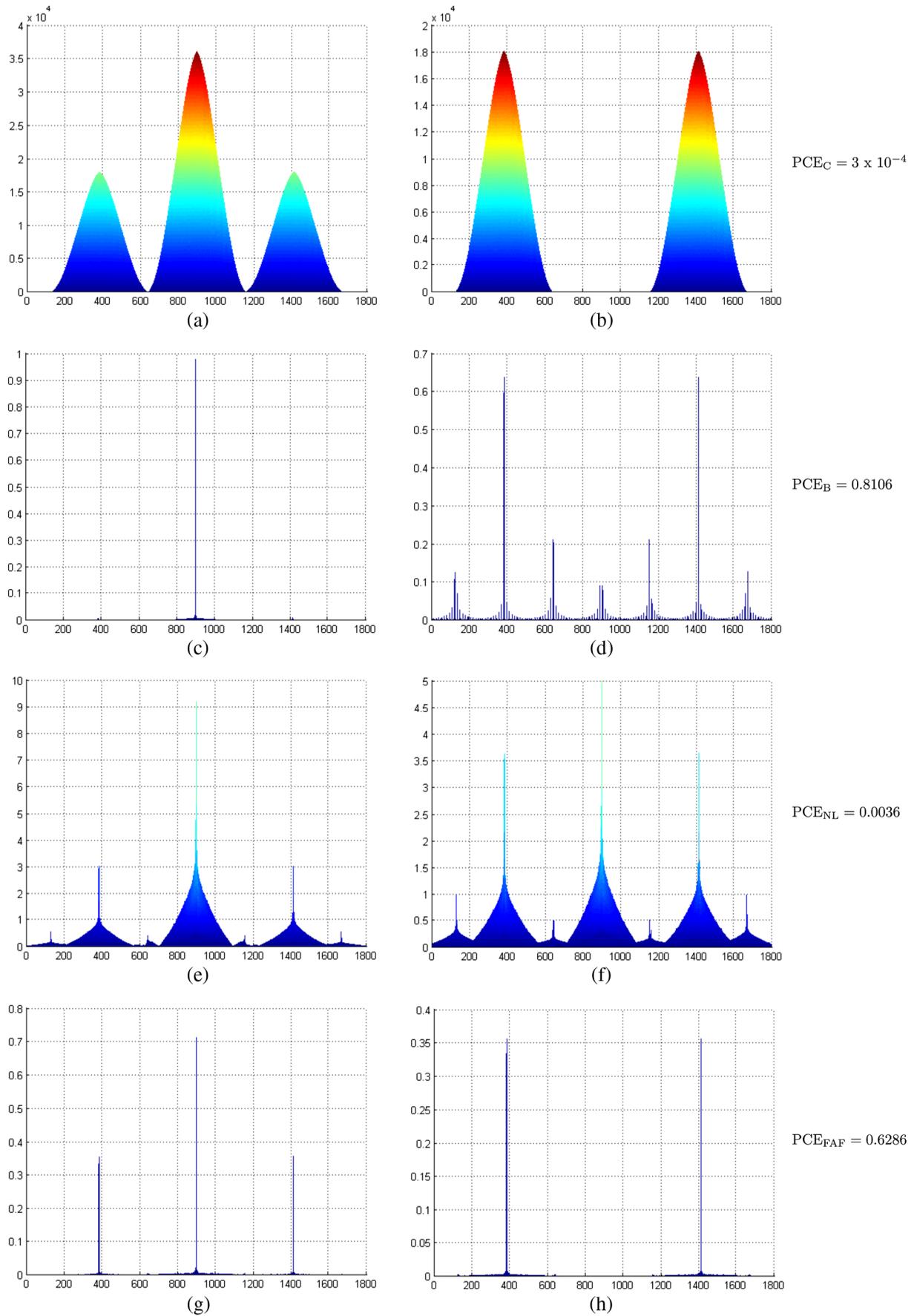


Fig. 9 Comparison of JTCs with zero order (left column) and nonzero JTCs (right column). (a) and (b): classical JTC. (c) and (d): binary JTC. (e) and (f): nonlinear JTC. (g) and (h): fringe adjusted JTC.

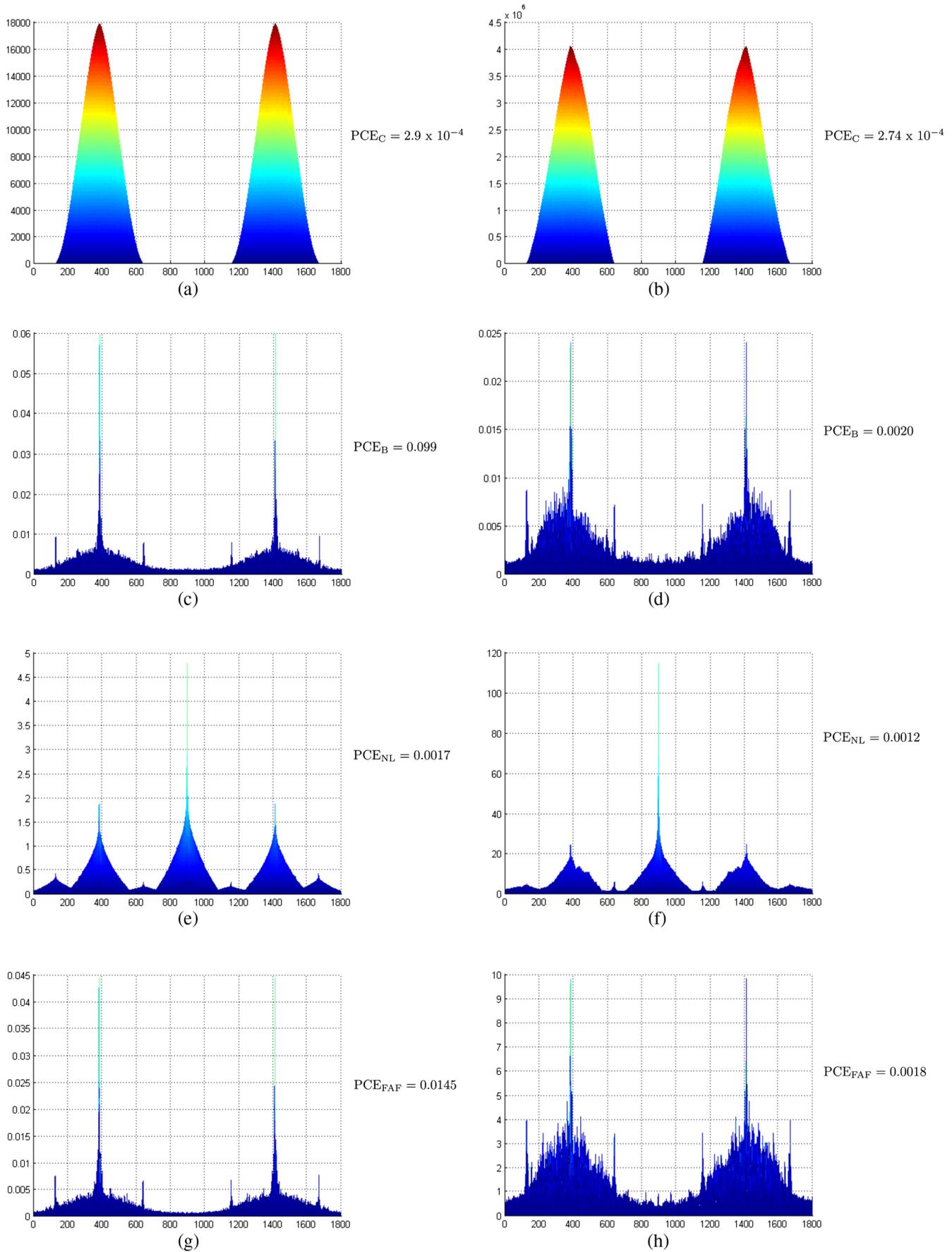


Fig. 10 Results of the robustness and the discrimination tests. Left column: comparison of the 2 mine images. Right column: comparison of the mine image and Lena image. (a) and (b): classical JTC. (c) and (d): binary JTC. (e) and (f): nonlinear JTC. (g) and (h): fringe adjusted JTC.

Table 1 Comparison of the binary JTC and the fringe adjusted JTC on images with underwater noises. a: Lena-Lena. b: Lena-blurred Lena. c: Lena-Lena with non uniform light. d: Lena-Lena with blurred and nonuniform light. e: Modified Lena- Lena with blurred and nonuniform light.

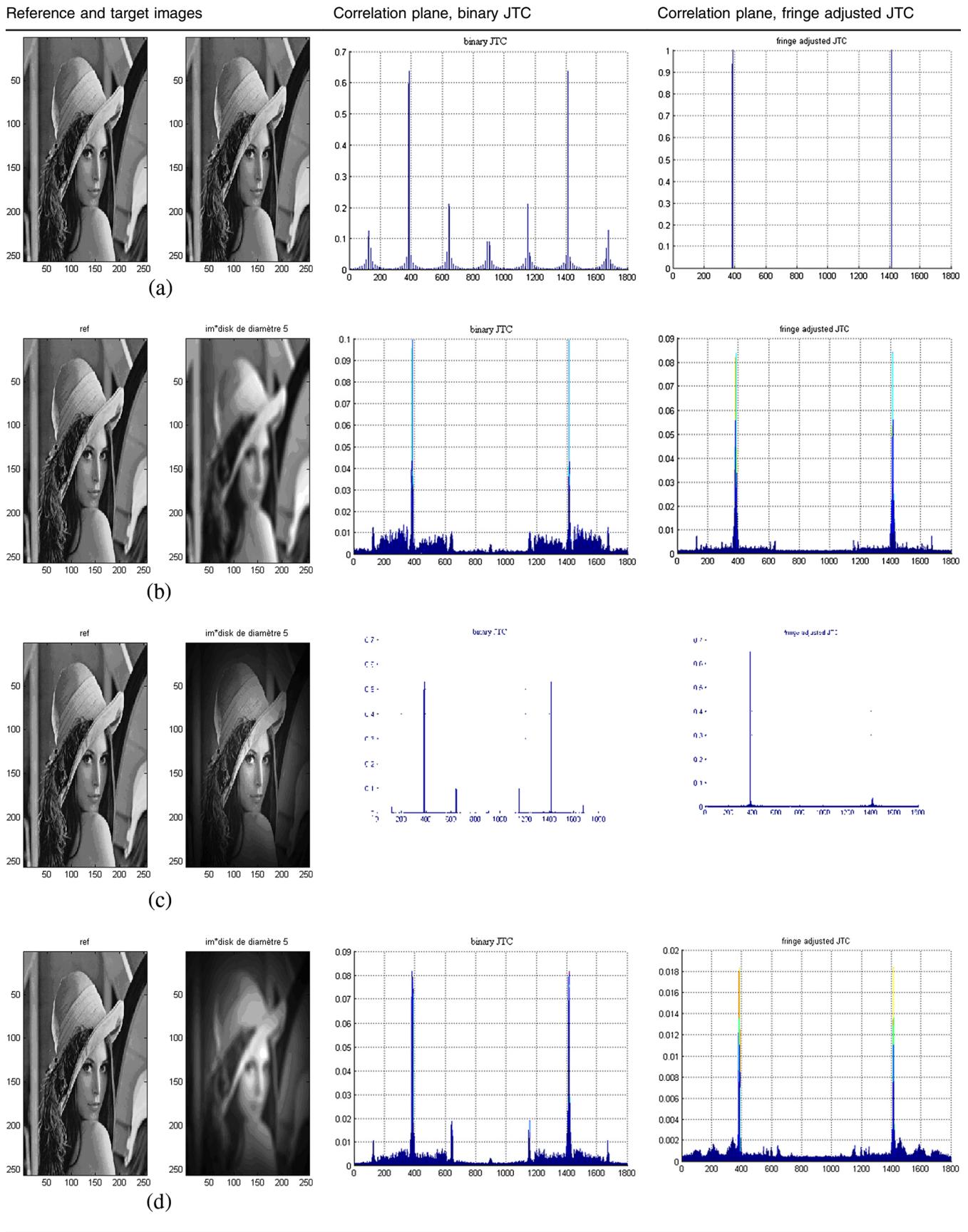
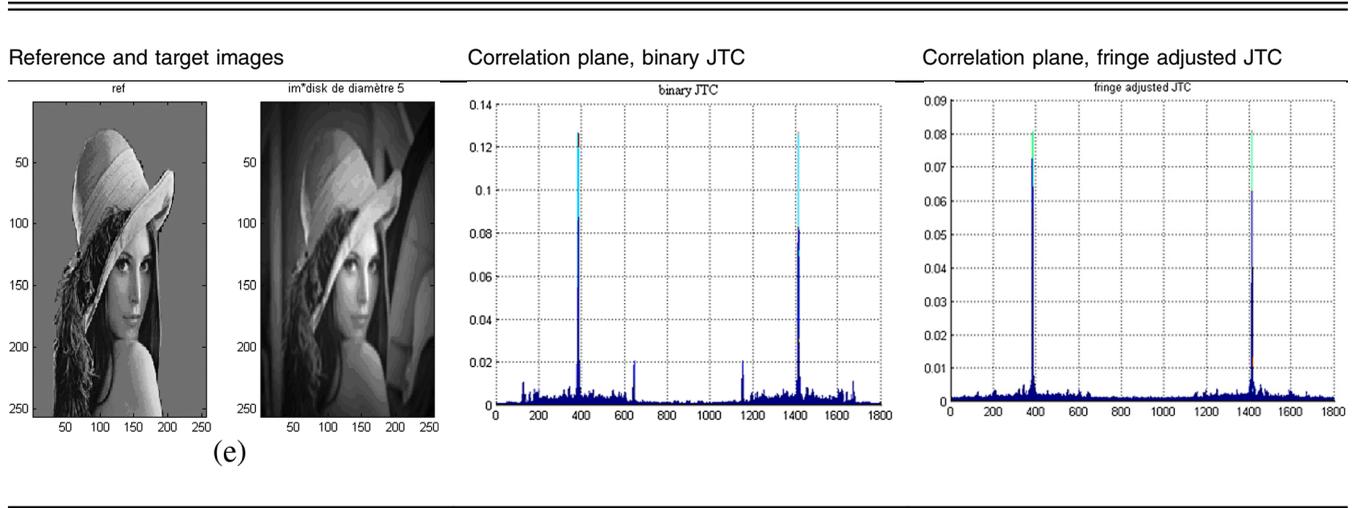


Table 1 (Continued).

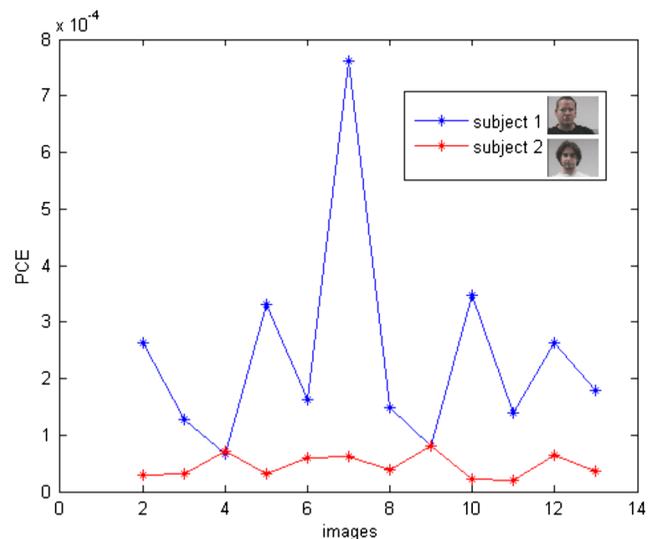
Fig. 11 Examples of facial rotation (-90 deg to $+90$ deg). Subject number (1), Base PHPID.¹⁶

the input plane. The target faces are obtained from the database presented in Fig. 11, one at a time starting from -90 deg to $+90$ deg in a sequential manner.¹⁶

We initialize the system with a known reference image, such as the subject's face (number 1), to be tracked so that initial position of subject's face is known. By comparing this reference face with the target image, we ascertain the presence or absence of the subject (number 1) in the target image and the position of the subject's face in the target image. Thereafter, only the information around the position of the face is selected in this target image. To achieve this, we multiply the target image, with a filter equal to 1, around the position of the face found and 0 elsewhere. Afterwards, we introduce the preprocessed image as the reference image in the fringe-adjusted JTC input plane. Then a new target image, subject number 2, is introduced in the input plane and the process is repeated until all images from the database are processed. After various tests performed by using this algorithm, we have identified a tolerance rotation angle equal to 15 deg between the target image and reference image in which the correlation remains robust. Beyond this value between the reference and target images, the correlation becomes less robust and may lose track of the subject's face. With a tolerance angle equal to 15 deg, the tracking problem of a person, with in-plane rotation between -90 and $+90$ deg, consists of correlating 13 pairs of images: $(-90, -75), (-75, 60), \dots, (-15, -0), (0, +15) \dots (+60, +75)$, and $(+75, +90)$.

The tracking results of subject number (1) are shown in Fig. 12 (blue curve). The PCE criterion was used for decision

making in the correlation plane.⁵ Figure 12 also shows the results when subject number 2, such as an undesirable subject present in the input scene (false alarm: red curve). The results, depicted in Fig. 12, show that the proposed fringe-adjusted JTC yields good results. However, for some cases, the values of the two cases are too close for making a robust system. A robust system is a system where the values of

Fig. 12 Results obtained with a FA-JTC using only one reference image position without any rotation in the H_{FAF} Filter to define the background noise.

recognition are close to 1. A discriminated system is a system where the false alarms are close to 0, with significant difference between true and false alarms.

To optimize the performance of the fringe-adjusted JTC system, we propose to utilize an adaptive fringe-adjusted filter (H_{FAF}) in the Fourier plane.¹⁷ In this technique, the background noise is calculated for each reference. For this purpose and after finding the position of head in the target image at time t_0 , we multiply the target image with a filter that has a value of 0 around the position of the face and 1 elsewhere. One can also replace the 0 value with the average of the values around the 0 value zone.

The target image obtained at time, t_0 , becomes the reference image at time, t_1 , in the proposed tracking system. Thereafter, we introduce the preprocessed image as the background noise in the fringe-adjusted filter formulation. The results obtained after the first iteration are shown in

$$P = \frac{\left(\sum_{i=i_0-t}^{i=i_0+t} \sum_{j=j_0-t}^{j=j_0+t} |C(i, j)|^2 \right) - \left\{ \sum_{k=1}^5 \frac{\left[\left(\sum_{i=i_k-t}^{i=i_k+t} \sum_{j=j_k-t}^{j=j_k+t} |C(i, j)|^2 \right) \right]}{5} \right\}}{\sum_{i=i_0-t}^{i=i_0+t} \sum_{j=j_0-t}^{j=j_0+t} |C(i, j)|^2}, \quad (9)$$

where (i_0, j_0) denote the position of the correlation peak, $C(i, j)$ denotes the value of the correlation plane at point (i, j) , t is set to the number of neighboring pixels used around the considered peak, (i_k, j_k) are the positions of five highest peaks excluding the correlation peak.

Step 2: Apply additional nonlinearity in the Fourier plane of FJTC similar to Eq. (3).¹⁰ After many trial runs in our laboratory, we set the value of k to 0.75.

Figure 14 shows the different test results obtained with our tracking system based on the optimized fringe-adjusted JTC. By comparing the results presented in Figs. 12 and 13, it is evident that the proposed technique yields better discrimination. Indeed, the difference between the two curves (red and blue) is much larger. The robustness has also improved as shown by the shorter range of values of the blue curve compared to the previous tests.

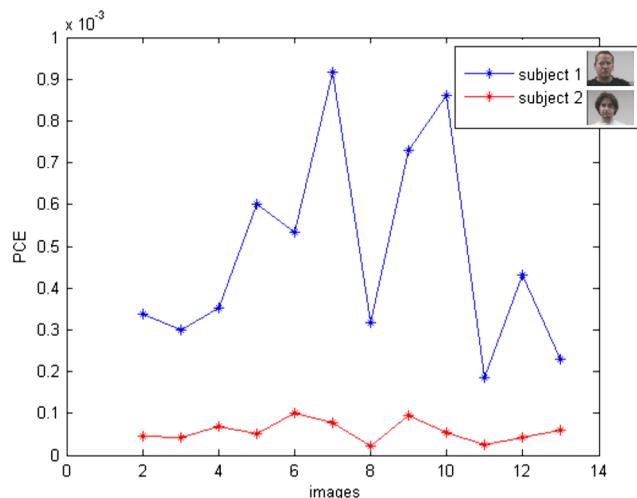


Fig. 13 Results obtained using the adaptive fringe-adjusted JTC.

Fig. 13 which clearly shows that the adaptive fringe-adjusted JTC improves discrimination performance considerably, when compared to the first case. Figure 13 shows a system that has a better robustness and discriminatory, but we think that we can find a system and a criterion that increase robustness and discriminatory.

To further increase the performance of the proposed system, we propose to incorporate a decision criterion in the correlation plane which involves the following optimizations steps:

Step 1: Enhanced PCE decision criterion (P): we propose to take into account the correlation peak energy as well as the five highest peaks excluding the highest correlation peak. The Enhanced PCE decision (P) criterion is defined as:

For noisy images, we searched for a robust criterion that provides values close to 1 in case of recognition and values close to 0 in case of false alarm. An idea is to compare the highest peak of the correlation plane to other high peaks. Consequently, we remove the autocorrelation peak in order to compute a criterion only with inter-correlation peaks and noise. Since the correlation plane is symmetric, we can work on one half of the correlation plane. Our objective is to find the maximum value, M , of the new correlation plane. We set a first threshold, M_{-3dB} to the value of the correlation peak M divided by $\sqrt{2}$, and n this value corresponds to 3 dB and to 50% of the correlation peak energy. We compute the cumulative sum (S1) of the pixels around the correlation peak and with values higher than M_{-3dB} . Next we compute the sum (S2) of all the pixels excluding the highest correlation peak with values higher than $0.75 \times M_{-3dB}$, as shown in Fig. 15. Finally, we introduce a distance metric defined as:

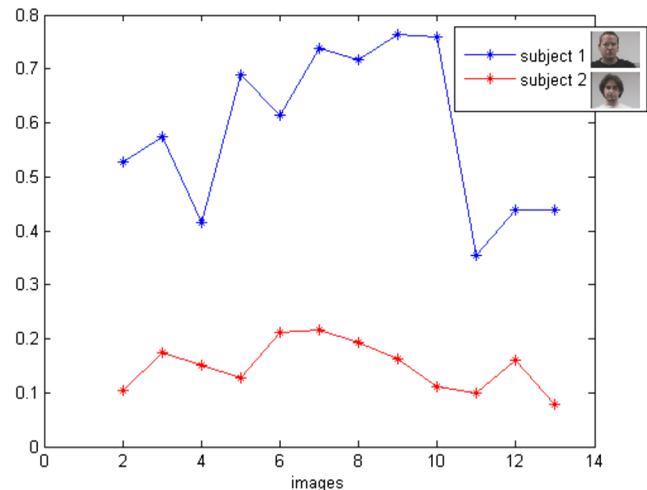


Fig. 14 Results obtained with a nonlinearity in the Fourier plane of the FA-JTC and the new criterion.

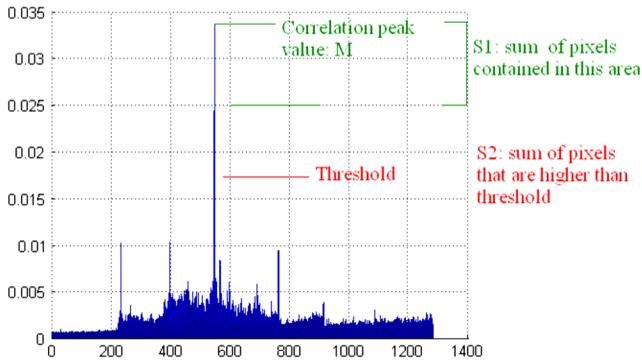


Fig. 15 Computation of our new criterion.

$$C = \frac{S1 - S2}{S1}. \quad (10)$$

The distance criterion yields values in the interval $[0, 1]$. In case of recognition, we obtain a high correlation peak with negligible noise, therefore, the value $S2$ will be very low. In case of false alarm, the correlation plane will be very noisy and the highest correlation peak will be surrounded by high peaks and the value $S2$ will be very close to the value $S1$.

We applied this criterion to the previous test and the corresponding results are shown in Fig. 16. This criterion provides recognition values close to 1 and false alarm values close to 0. Thus, the robustness and discrimination ability of the proposed technique is very high. Using the aforementioned procedure, users can set a threshold for automatic decision making.

6.2 Underwater Mine Recognition

In the previous section, we validated the proposed architecture and criteria for face recognition applications. In this section, we applied the proposed technique for underwater mine detection. As there are a lot of images, we used composite filter techniques to fuse close references and reduce the size of the database.¹² For this application, we selected five filters to recognize mines in a specific region of an underwater video sequence. As explained earlier, underwater

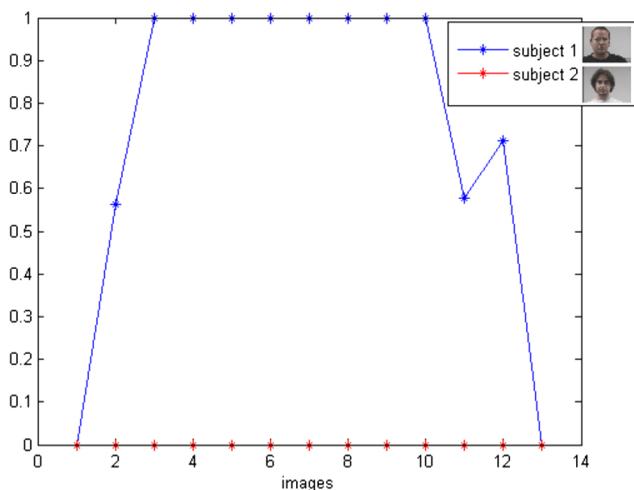


Fig. 16 Results obtained with a nonlinearity in the Fourier plane of the FA-JTC and the last criterion.

images are perturbed by the medium and need preprocessing. Without preprocessing, correlation results will not provide desired output since target related contours will not stand out. The preprocessing step consists of resizing the image to reduce computation time. Then, we enhance the contrast of images using the procedure described in Ref. 14.

To obtain the contour image, we chose to compute the phase of the image. In fact, edges correspond to sharp phase changes.¹⁸ To achieve this, we compute the Fourier transform of the image where the image can be expressed as $\text{Spetrum} = \rho \cdot e^{i\theta}$. The amplitude information is contained in ρ and the phase information is contained in the exponential function. To keep only the phase information, all values are divided by their amplitude ρ yielding $\text{Spetrum} = e^{i\theta}$. Next, we use a band pass filter, constructed with a Gaussian wavelet, to reduce the remaining noise. The results of this preprocessing step are shown in Fig. 5.

From this contour image database, we conducted five composite filters for mine recognition in the mine-presence zones as illustrated in Fig. 17. Each filter is constructed with five reference contour images. With these five composite filters and preprocessed images, we tested the performance of the nonzero-order fringe adjusted JTC and the nonzero-order nonlinear fringe adjusted JTC. We set the nonlinearity factor k to 0.85 and corresponding correlation outputs are shown in Table 2. From Table 2, it is obvious that the nonzero-order fringe adjusted JTC provides a high false alarm rate while the nonzero-order nonlinear fringe adjusted JTC provides moderate recognition rate.

To enhance the decision performance further by increasing the recognition rate and decreasing the false alarm rate, we decided to add an optimization step. We assume that for 10 images before the present target image if we have five correlations, it means that a mine should be detected in the present target image. Knowing that the mine is fixed

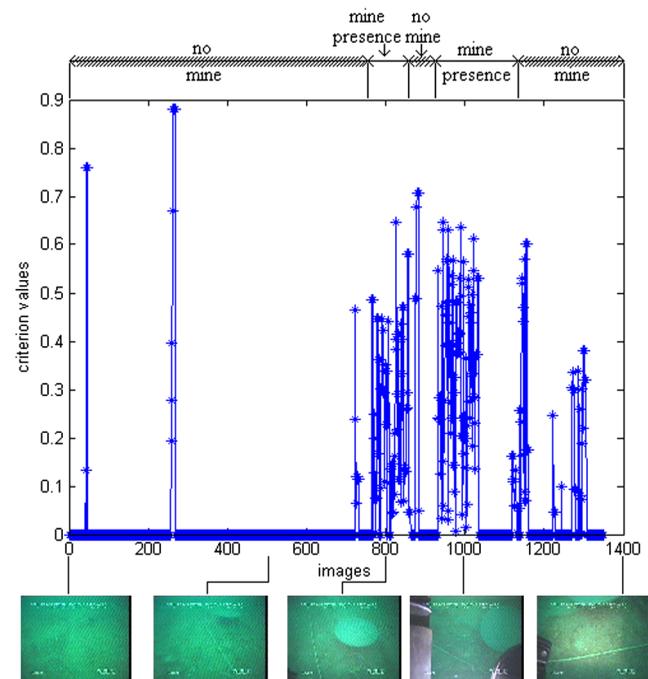


Fig. 17 Results of the optimized nonlinear fringe-adjusted JTC, $k = 0.85$.

Table 2 Results of the different JTCs for mine recognition.

Method	Recognition (%)	False alarm (%)
Fringe adjusted JTC	93.10	56.44
Nonlinear fringe adjusted JTC, $k = 0.85$	52.07	16.29

Table 3 Results of the nonlinear fringe-adjusted JTC with optimization.

Method	Recognition (%)	False alarm (%)
Optimized nonlinear fringe adjusted JTC, $k = 0.85$	65.52	10.13

and knowing the navigation parameters, a mine, which is detected in one image frame, is highly unlikely to disappear in the following image frame. Our choice of 10 images has been made after several tests in order to find the right compromise between robustness and computing time. Using this concept, our optimization consist in considering only the last 10 raw decisions, coming from last 10 images, for a given present target image. Using last 10 nonoptimized results leads us to decrease the false alarm rate. In fact, if we use the optimized results, the detector never stops detecting after five detections, even if the five detections are false. However if we use our raw data, this optimization leads to a stop of detection, five images after the last image where a mine has been detected. On the test results, we can see that the recognition rate has improved while the false alarm rate has decreased, as illustrated in Table 3. In Fig. 17, we see regions where a mine is supposed to be detected and regions where the algorithm is supposed to find nothing. The results presented Table 3 and Fig. 17 shows that we managed to increase the recognition rate to 65.52 percent and we decrease the false alarm rate to 10.13 percent.

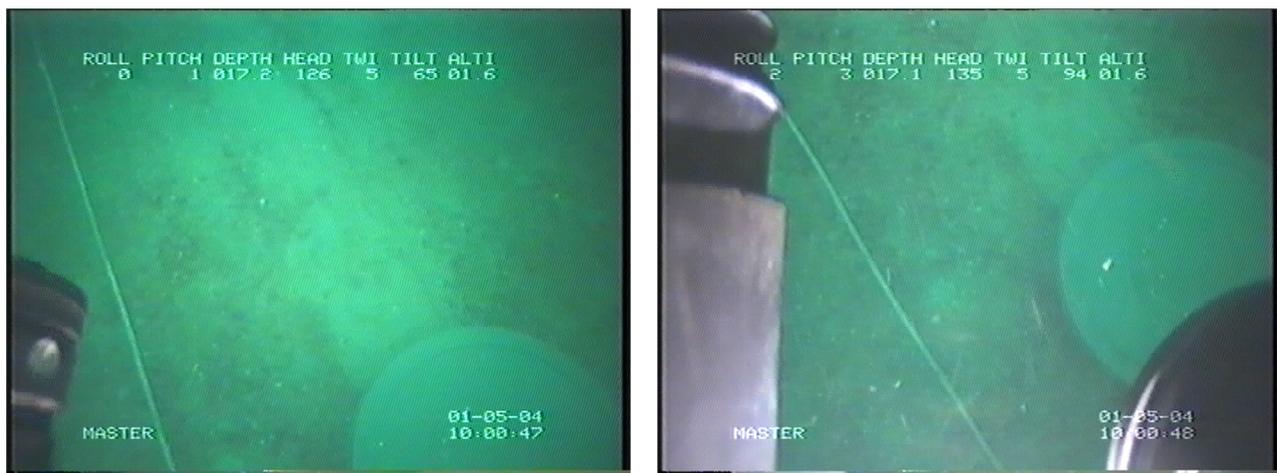
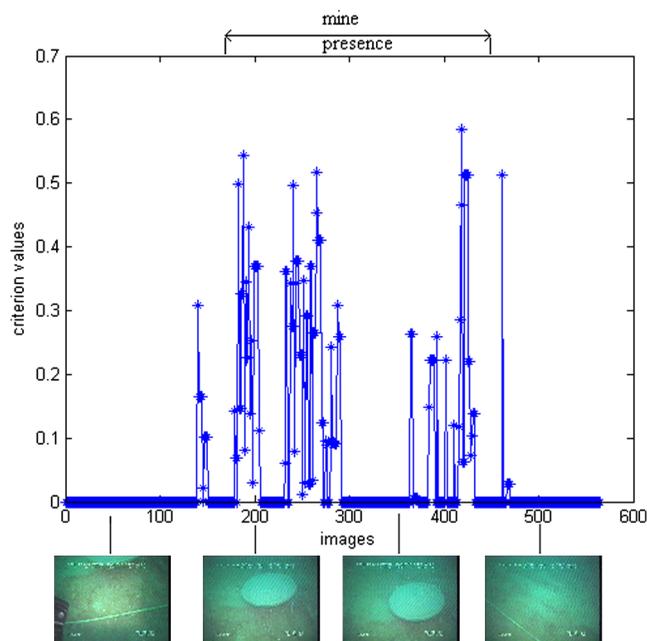
**Fig. 18** Images with presence of the camera support.**Fig. 19** Difference between mine size in the 2 video sequences.

Table 4 Robustness results of the different JTCs.

Method	Recognition (%)	False alarm (%)	True to false alarm ratio (%)
Fringe-adjusted JTC	88.52	59.93	1.48
Nonlinear fringe-adjusted JTC, $k = 0.85$	34.07	12.33	2.76
Optimized nonlinear fringe-adjusted JTC, $k = 0.85$	41.11	5.48	7.5

Nevertheless, our optimized algorithm detected some objects, in the last region of the video sequence, where there is no mine. This can be explained by the presence of the camera support in target images as shown in Fig. 18. We did not add a mask to suppress this man-made object because there are images with no camera support and images with camera support, but it did not occupy the same proportion of the image as shown in Fig. 18, however, this constraint can be avoided when recording the images.

To show the robust performance of our proposed approach, several tests were conducted and the results were compared with nonoptimized methods. To do this, we selected a video sequence where the same mine is recorded, but the distance between the viewer and the mine is shorter in this case compared to the previous case shown in Fig. 19. Results obtained, in terms of recognition and false alarm probabilities, are summarized in Table 4. The visualization of the robustness of the optimized NNJTC is shown in Fig. 20. The nonzero-order fringe adjusted JTC still provides a high false alarm rate while the nonzero-order nonlinear fringe adjusted JTC presents good recognition rate. The

**Fig. 20** Results of the optimized nonlinear fringe-adjusted JTC, $k = 0.85$.

optimization allows an increase of the recognition rate and a decrease of the false alarm rate, when compared to the nonlinear fringe adjusted JTC.

7 Conclusion

In this paper, we proposed and validated the first version of a robust recognition and tracking system based on the fringe adjusted JTC architecture. In the proposed technique, we introduced nonlinearity to the fringe adjusted JTC, and also used an adaptive correction criterion in the correlation plane.

On the face recognition application, we show that a rotation angle of 15 deg between the target image and the reference image does not disable the recognition ability of the system, while considerably reducing the number of correlations necessary and, therefore, increasing the processing speed of the system. Moreover, the adaptive decision criterion, used in the output plane, minimized the number of false correlation peaks in the correlation plane thus reducing the false alarms. Finally, the addition of nonlinearity in the Fourier plane makes it possible to adapt or find a compromise between the robustness and discrimination of our tracking system by adjusting the degree of nonlinearity (k). For underwater mine recognition application, the proposed NNJTC architecture yields promising results.

Acknowledgments

This paper complements our paper "Nonlinear fringe-adjusted JTC-based face tracking using an adaptive decision criterion".¹⁹ The raw mine images used in this publication were provided by the GESMA (Groupe d'Etudes Sous Marines de l'Atlantique) under the TOPVISION project coordinated by Thales Underwater Systems SAS. This project is related to Techno-Vision Program, which was launched by the French Ministry of Defense. Detailed information can be found in the website <http://topvision.gesma.fr>. This work is also partly supported by the Regional Council of Brittany.

References

1. A. Alfalou and C. Brosseau, "Understanding correlation techniques for face recognition: from basics to applications," *Face Recognition*, M. Oravec, Ed., ISBN: 978-953-307-060-5, INTECH. Available from: <http://sciendo.com/articles/show/title/understanding-correlation-techniques-for-face-recognition-from-basics-to-applications>.
2. V. VanderLugt, "Signal detection by complex spatial filtering," *IEEE Trans. Info. Theory* **10**(2), 139–145 (1964).
3. C. S. Weaver and J. W. Goodman, "A technique for optically convolving two functions," *Appl. Opt.* **5**(7), 1248–1249 (1966).
4. M. S. Alam and M. A. Karim, "Fringe-adjusted joint transform correlation," *Appl. Opt.* **32**(23), 4344–4350 (1993).
5. J. L. Horner, "Metrics for assessing pattern-recognition performance," *Appl. Opt.* **31**(2), 165–166 (1992).
6. I. Léonard, A. Alfalou, and C. Brosseau, "Spectral optimized asymmetric segmented phase-only correlation filter," *Appl. Opt.* **51**(14), 2638–2650 (2012).
7. B. Javidi and C. J. Kuo, "Joint transform image correlation using a binary spatial light modulator at the Fourier plane," *Appl. Opt.* **27**(4), 663–665 (1988).
8. L. Guibert et al., "On-board optical joint transform correlator for real-time road sign recognition," *Opt. Eng.* **34**(1), 135–143 (1995).
9. C. T. Li, S. Yin, and F. T. S. Yu, "Nonzero-order joint transform correlator," *Opt. Eng.* **37**(1), 58–65 (1998).
10. M. S. Alam, "Fractional power fringe-adjusted joint transform correlation," *J. Opt. Eng.* **34**(11), 3208–3216 (1995).
11. M. S. Alam and M. A. Karim, "Multiple target detection using a modified fringe-adjusted joint transform correlator," *J. Opt. Eng.* **33**(5), 1610–1617 (1994).
12. I. Leonard, A. Arnold-Bos, and A. Alfalou, "Interest of correlation-based automatic target recognition in underwater optical images: theoretical justification and first results," *Proc. SPIE* **7678**, 767800 (2010).
13. S. Q. Duntley, "Light in the sea," *JOSA*, **53**(2), 214–233 (1963).

14. D. N. Sidorov and A. C. Kokaram, "Suppression of moiré patterns via spectral analysis," *Proc. SPIE*, **4671**, 895 (2002).
15. A. Arnold-Bos, J.-P. Malkasse, and G. Kervern, "Towards a model-free denoising of underwater optical images," in *Proc. of the IEEE conference on Ocean (Europe)*, IEEE, Europe (2005).
16. N. Gourier, D. Hall, and J. L. Crowley, "Estimating face orientation from robust detection of salient facial features," in *Proc. of Pointing 2004, ICPR, International Workshop on Visual Observation of Deictic Gestures*, Cambridge, United Kingdom (2004).
17. M. S. Alam and A. Bal, "Dynamic target tracking with fringe-adjusted joint transform correlation and template matching," *Appl. Opt.* **43**(25), 4874–4881 (2004).
18. A. Oppenheim and J. Lim, "The importance of phase in signals," *Proc. IEEE* **69**(5), 529–541 (1981).
19. I. Léonard, A. Alfalou, and M. S. Alam, "Nonlinear fringe-adjusted JTC-based face tracking using an adaptive decision criterion," *Proc. SPIE* **8134**, 81340M (2011).

Isabelle Leonard received the engineering diploma in optronic from the French engineering school Polytech'Paris-Sud in 2009. Since 2009, she is a PhD student at Institut Supérieur d'Electronique et du numérique (ISEN). Her research interests deal with image processing and underwater medium. She has published over 7 refereed journal articles or conference papers.



Ayman Alfalou received his PhD in telecommunications and signal processing from École nationale supérieure des télécommunications de Bretagne (ENSTB)-France and the University of Rennes 1 in 1999. He held a one-year post-doctoral position at ENSTB during which he designed and realized a compact and high-rate optical correlator. Since 2000, he has been a professor of telecommunications and signal processing at ISEN-Brest. At ISEN, he founded the Vision-L@bISEN. His research interests deal with optical engineering, optical information processing, signal and image processing, telecommunications and optoelectronics. He has supervised several PhD, MSc, and engineering school students. He has published over 110 refereed

journal articles or conference papers and special sessions. He is a senior member of SPIE, OSA, IEEE, and a member of IoP.



Mohammad S. Alam is a professor and Chair at the University of South Alabama (USA). His research interests include ultra fast computing architectures and algorithms, image processing, pattern recognition and tracking, biometric recognition, infrared imaging systems, and smart energy management and control. He is the author or co-author of over 475 refereed journals, conference publications, or project reports, 15 book chapters, and a book on IPTV (IEC Press).

He received numerous excellences in research, teaching and service awards. He served or serves as the PI or Co-PI of many research projects totaling over \$14M, supported by NSF, FAA, DoE, ARO, AFOSR, SMDC, NASA, WPAFB, BP and ITT industry. He has organized and chaired many international conferences and serves as a Guest Editor for several professional journals. He is a Fellow of OSA, SPIE, IoP, IET, a Life Fellow of the Bangladesh Computer Society (BCS), and of the Institution of Engineers Bangladesh (IEB), and a Senior Member of IEEE.

Andreas Arnold-Bos graduated from École nationale supérieure de l'aéronautique et de l'espace (SUPAERO) [(now Institut Supérieur de l'aéronautique et de l'espace (ISAE)], Toulouse, France, in 2004, with the engineering degree of aerospace engineering, as well as a master's degree in Signal and Image Processing. From 2004 to 2007 he worked towards a PhD thesis on radar engineering at the École Nationale Supérieure des Ingénieurs des Études et Techniques d'Armement in Brest, France. In 2007, he joined Thales Underwater Systems as a full-time R & D engineer while simultaneously pursuing his PhD thesis, which he successfully defended in 2010. His past and current interests involve radar and sonar engineering, underwater optics and autonomous systems.