



HAL
open science

Adaptive nonlinear fringe-adjusted joint transform correlator

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

► **To cite this version:**

Isabelle Leonard, Ayman Alfalou, Mohammad S. Alam, Andreas Arnold Bos. Adaptive nonlinear fringe-adjusted joint transform correlator. *Optical Engineering*, 2012, pp.098201-1 098201-14. 10.1117/1.OE.51.9.098201 . hal-00781755

HAL Id: hal-00781755

<https://hal.science/hal-00781755>

Submitted on 28 Jan 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Adaptive Nonlinear Fringe-adjusted Joint Transform Correlator

Isabelle Leonard^a, Ayman Alfalou^{a*}, Mohammad Showkat Alam^b, Andreas Arnold-Bos^c

^a ISEN Brest, Département Vision, L@bISEN, 20 rue Cuirassé Bretagne, CS 42807,
29228 Brest Cedex 2, France

^b Department of Electrical and Computer Engineering, EEB 75, University of South Alabama
6001 USA South Dr., Mobile, AL 36688-0002

^c Thales Underwater Systems, Route de Sainte Anne du Portzic, CS 43814, 29238 BREST CEDEX
3, France

* **ayman.al-falou@isen.fr**

ABSTRACT

An optimized technique based on the fringe-adjusted JTC (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 by: (1) utilizing the fringe-adjusted filter (H_{FAF}) in the Fourier plane, (2) adding nonlinear processing in the Fourier plane, and (3) using 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 PHPID (Pointing Head Pose Image Database) data base confirm robust performance of the proposed method for face recognition and tracking applications. Thereafter, we also tested the proposed technique for a challenging application i.e. underwater mine detection and excellent results were obtained.

Keywords: correlation, nonlinear fringe-adjusted JTC, face recognition, underwater mine recognition.

1. INTRODUCTION

Correlation is a robust and efficient pattern recognition technique 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: (1) the Vander Lugt Correlator (VLC) [2], and (2) the Joint Transform Correlator (JTC) [3]. Unlike the JTC, VLCs requires the implementation of a complex-valued filter. The optical implementation of the JTC architecture doesn't 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.

This paper complements our paper "Nonlinear fringe-adjusted JTC-based face tracking using an adaptive decision criterion" [4]. We utilized the JTC-based correlation architecture to realize a face tracking system. The JTC architecture can be used for near real time operation because the reference images can be updated in near real time. In addition, the proposed system must be robust to changes in in-plane rotation of the target while ensuring accurate discrimination between multiple targets. In this paper, our objective 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 [5]. This fringe-adjusted JTC is very discriminatory but its robustness may not be sufficient for some applications. For that, we add a nonlinearity function to the fringe-adjusted JTC architecture to obtain a robust and discriminatory architecture. Indeed, the nonlinear function increases the classical JTC discriminatory.

Special attention is given to generate an optimized correlation output based on a decision criterion. The peak to correlation energy (PCE) criterion is not very suitable for automatic decision making [6-7]. The presence of autocorrelation peak and the presence of two correlation peaks can perturb the decision criterion. For that we used an adapted JTC architecture to suppress the zero-order

autocorrelation peak. Since the correlation plane is symmetric, we work on the half 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. ANALYTICAL MODELING

The JTC correlation architecture has been described in detail in the literature, but we briefly review the principle of JTC architecture in this paper. For more information, we invite the reader to consult the many works proposed and validated in this area by several researchers working in this field [1-3, 5,8-9].

In a 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. This plane can be represented mathematically by function " f " described in Eq. (1). To simplify its mathematical representation, we consider only a shift distance " x_0 " along the x axis between the target image and the reference image.

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

After applying Fourier transform to Eq. (1) *i.e.*, the input plane, we can record the intensity or the joint power spectrum (JPS) using a square law device such as a CCD detector array (we note F this JPS). In a classical JTC [3], an inverse Fourier transform of the JPS yields the correlation output in the output plane. This correlation output contains two important terms: (1) a central autocorrelation peak, and (2) two cross-correlation peaks corresponding to the correlation between image " i " and image reference " r ". These two cross-correlation peaks are located on either side of the central zero-order peak at a distance proportional to " x_0 ".

The classical JTC provides large correlation peaks. These large peaks and the presence of autocorrelation peak can perturb the decision. Several JTC architectures have been developed to solve this problem such as the binary JTC, non-zero-order JTC, and fringe-adjusted JTC [8, 10, 5]. In a binary JTC [8], the JPS is binarized using a suitable threshold before applying the inverse Fourier transform step. The binarized JPS is defined as

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

This architecture provides narrow correlation peaks. However, it is sensitive to noise, rotation and scale changes. Moreover, the binarization step requires on-line processing. Another interesting JTC architecture is the nonlinear JTC [11] where a nonlinear threshold is applied to the JPS as shown below:

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

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. If k is set to 0, we obtain the binary JTC. The discrimination of the nonlinear JTC is superior than the discrimination of the classical JTC. It is more efficient when the value of k is lower than 0.5 [11- 12].

The last optimization we present herein is the fringe-adjusted JTC. Alam et al. [5, 11-12] proposed to introduce a real-valued filter, called fringe-adjusted filter (H_{FAF}), in the Fourier plane of the correlator, defined as

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

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. Another class of architecture, called the multiobject JTC or nonzero-order JTC, has been found to be very useful for enhancing the correlation discrimination [10, 12]. In this architecture, the reference-only and the input-scene-only power spectra are subtracted from the JPS in the Fourier domain, given by

$$\begin{aligned}
|F_{NZ}(\mu, \nu)|^2 &= |F(\mu, \nu)|^2 - |R(\mu, \nu)|^2 - |I(\mu, \nu)|^2 & (5) \\
|F_{NZ}(\mu, \nu)|^2 &= |R(\mu, \nu)|^2 + |I(\mu, \nu)|^2 + 2|R(\mu, \nu)I(\mu, \nu)|\cos(\phi_i(\mu, \nu) - \phi_r(\mu, \nu) + 2u x_0) \dots \\
&\quad - |R(\mu, \nu)|^2 - |I(\mu, \nu)|^2 \\
|F_{NZ}(\mu, \nu)|^2 &= 2|R(\mu, \nu)I(\mu, \nu)|\cos(\phi_i(\mu, \nu) - \phi_r(\mu, \nu) + 2u x_0)
\end{aligned}$$

In this technique, the zero-order term and the autocorrelation term among similar input scene targets or non-target objects are automatically eliminated while leaving the desired crosscorrelation peaks in the output plane.

In this paper, we first propose and validate a nonzero-order fringe-adjusted JTC architecture (NZ_FAJTC). Then we applied a special non linear function to the NZ_FAJTC. Before discussing the results obtained with this optimized fringe-adjusted JTC architecture, we present the two applications considered in this paper.

Our first application consists in 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° to $+90^\circ$. Face images contain information as well as noise. We labeled noise related continuous information like background. This noise is “known” and is quite similar from one image to another in the sequence.

Once our system is validated, we apply it for underwater mine detection and recognition [13]. The main problem in this application is the poor visibility, which can cut the visibility down to only a couple of meters, if not less, in turbid waters.

3. CHALLENGES OF UNDERWATER IMAGING

The underwater image formation model (see for instance [14] for reference) is explained in Fig. 1. In Fig. 1, Light source (1) transmits light (2) towards the object (6). On the trajectory to and from the object, light is intercepted by particles (3), so the light rays (2), (5), (7) and (9) become exponentially dimmer due to the direct attenuation. Part of the light is also reflected towards the camera by the medium (3), resulting in the backscattered component (4), which corresponds to the “image” of the water volume. Finally, some light is slightly deviated by the medium (3) on its way back to the camera, resulting in the forward scattered component (8), which adds blur 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 investigate the performance of the proposed technique, we chose three images (Fig. 2, Fig. 3 and Fig. 4) and several algorithms were applied for comparison purposes. The first two images came from one video sequence acquired by a diver and represents an upside down MN 103 “Manta” mine. 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. Indeed, two images shot within a small timeframe are similar,

so the output of the filter should also stay similar. The discrimination of our system is tested by applying JTC architectures on a mine image and the Lena image (Fig. 4).

At first, we compared the classical JTC and the non-zero-order JTC on the manta mine image as shown in Fig. 2. We used the image frame 1082 of the test video as the reference image and as the target image. The results are in Fig. 5 and 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 non-zero-order JTC for the above mentioned images (Fig. 2, Fig. 3 and Fig. 4). We used the PCE criterion as decision criterion on the non-zero-order JTC correlation plane as [6]

$$PCE = \frac{\sum_{x=x_0-t}^{x=x_0+t} \sum_{y=y_0-t}^{y=y_0+t} |C(x, y)|^2}{\sum_{x=1}^{x=N} \sum_{y=1}^{y=M} |C(x, y)|^2} \quad (6)$$

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. 6.

Figure (5) shows that the nonzero-order JTC architecture yields helpful results for decision making. Figure (6) shows that the fringe adjusted JTC architecture observes these constrains on PCE values better than the binary JTC architecture. The nonlinear JTC architecture is a robust one i.e. PCE values are close between Fig. 5 and Fig. 6 as shown in the left columns. However, it does not provide robust discrimination as shown in Fig. 6 i.e. PCE values are close for the cases of (i) two mine on the left, and (ii) one mine and the Lena.

In this paper, our idea is to combine fringe adjusted JTC and nonlinear JTC in order to obtain robust architecture for correlation discrimination. This architecture, called nonlinear fringe-adjusted JTC, where the JPS is modified to obtain the nonlinear fringe-adjusted JPS, defined as

$$F_{FAF-NL}(\mu, \nu) = \left(\frac{B(\mu, \nu)}{A(\mu, \nu) + |R(\mu, \nu)|^2} \cdot F \right)^k \quad (7)$$

Where k denotes the nonlinear degree, we set the value of k depending on the application. If k is set to 1, we obtain the classical FA_JTC. Before applying the nonlinear fringe-adjusted JTC, we will compare the results of the binary JTC and the fringe-adjusted JTC on underwater images discussed in the previous section.

4. Binary JTC and Fringe-adjusted JTC Performance Comparison

To compare the performance of non-zero-order binary JTC and the non-zero-order fringe-adjusted JTC, we consider the Lena image (Fig. 4) as the known reference image as well as the unknown target image i.e. 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 brings artificial light. The next step was to combine blurred and nonuniform light into the final target image. For the last test, we manually remove the background of our reference image. The target image is the Lena image corrupted with blur and non uniform light.

The correlation output results are presented in Table 1. It is evident that the non-zero-order fringe-adjusted JTC yields better results compared to the binary JTC. The FA_JTC is robust against different type of noise considered in table 1 i.e. underwater noise. According to these results and for underwater videos, the FA_JTC can provide better results compared to the binary.

5. TEST RESULTS

5.1. FACE RECOGNITION

At first, we applied the fringe-adjusted JTC for face recognition. The objective is to detect and track a person via face recognition in a video sequence (see Fig.7). Fig. 8 presents the results obtained with fringe-adjusted JTC where only one reference image is used to construct the fringe-adjusted filter H_{FAF} (the face of the subject shown in Fig. 7 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 the input plane. The target faces are obtained from the database presented in Fig. 7, one at a time starting from -90° to $+90^\circ$ in a sequential manner [15].

We initialize the system with a known reference image, *i.e.*, 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 do 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. With the various tests performed by using this algorithm, we have identified a tolerance rotation angle equal to 15° 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° , the tracking problem of a person with in-plane rotation (between -90 and $+90^\circ$), consists of correlating "13" pairs of images: $(-90, -75)$, $(-75, 60)$, ..., $(-15, -0)$, $(0, +15)$... $(+60, +75)$, and $(+75, +90)$.

The tracking results of subject number (1) are shown in Fig. 8 (blue curve). The PCE criterion was used for decision making in the correlation plane [6]. Fig. 8 also shows the results when subject number 2 i.e., undesirable subject is present in the input scene (false alarm: red curve). The results depicted in Fig. 8 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. That we call a robust system is a system where the value of recognition are close to 1. A discriminated system is a system where the false alarm are close to 0 and the distance between recognition and false alarm is important. The presented results do not fulfill these criterions.

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 [16]. 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 equal which 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_l 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 Fig. 9 which clearly shows that the adaptive fringe-adjusted JTC improved the discrimination performance when compared to the first case.

Fig. 9 shows a system that have 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. Moreover, we applied the k -th power nonlinearity to the fringe-adjusted JPS, where k is a real and positive number. The optimization procedure is described in the following section.

To further increase the performance of the proposed system, we utilized two additional optimizations steps:

- 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

$$P = \frac{\left(\sum_{i=i_0-t}^{i=i_0+t} \sum_{j=j_0-t}^{j=j_0+t} |C(i, j)|^2 \right) - \left(\sum_{k=1}^5 \left(\sum_{i=i_k-t}^{i=i_k+t} \sum_{j=j_k-t}^{j=j_k+t} |C(i, j)|^2 \right) / 5 \right)}{\sum_{i=1}^{i=i_0+t} \sum_{j=1}^{j=j_0+t} |C(i, j)|^2} \quad (8)$$

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.

- 2) The second optimization performed incorporated in the fringe-adjusted filter involves adding nonlinearity in the Fourier plane of the correlator [11] such as presented in equation 3. After many tests in our laboratory, we set the value of k to 0.75.

Fig. 10 shows the different test results obtained with our tracking system based on an optimized fringe-adjusted JTC. By comparing the results presented in Figs. 8 and 9, it is evident that the proposed technique yields better discrimination. Indeed the difference between the two curves (red and blue) is larger. The robustness has been increased to: the distances between the values of the blue curve are shorter than in the previous tests.

In case of noisy images, users may want a more robust criterion. So we searched for a 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 intercorrelation peaks and noise. Since the correlation plane is a symmetric one, 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 $\sqrt{2}$ (this value corresponds to 3dB 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 $0.75 \times M_{-3dB}$ (see Fig.11). Finally, we compute a distance metric defined as

$$C = \frac{S1 - S2}{S1} \quad (9)$$

The distance criterion yields values in the interval $[0, 1]$. In case of recognition, as we have a high correlation peak with negligible noise i.e. 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 results are shown in Fig. 12. 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. Thanks to this discrimination, users can set a threshold for automatic decision making.

5.2. UNDERWATER MINE RECOGNITION

In the previous section, we validated the proposed architecture and criteria for face recognition applications. We will now apply it 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 [13]. For this application, we selected five filters to recognize mines in a specific region of an underwater video sequence. As explained earlier, underwater images are perturbed by the medium and need preprocessing. Without preprocessing, correlation results will not provide desired 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 with the processing described in Ref. 17. To obtain the contour image, we chose to compute the phase of image. In fact, edges correspond to sharp phase change [18]. To achieve this, we compute the Fourier transform of the image where the image can be expressed as $I = r.e^{i\theta}$. The amplitude information is contained in r and the phase information is contained in the exponential function. To keep only the phase information, all values are divided by their amplitude r yielding $I = e^{i\theta}$. Next we use a band pass filter constructed with a Gaussian wavelet to reduce the remaining noise. The results of this preprocessing step is shown on Fig. 13.

From these contour images database, we conducted five composites filters for mine recognition in the mine-presence zones (Fig.14). Each filter is constructed with five reference contour images. With these five composite filters and preprocessed images, we tested the nonzero order fringe adjusted JTC, and the nonzero order nonlinear fringe adjusted JTC. We set the nonlinear factor k to 0.85 and corresponding correlation outputs are shown in Table 2. The nonzero order fringe adjusted JTC provide a high false alarm rate while the nonzero order nonlinear fringe adjusted JTC provides moderate recognition rate.

To enhance the decision performance further (increase the recognition rate and decrease the false alarm rate), we decided to add an optimization to our set-up. We assume that for 10 images (before the target image) if we have five correlations a mine should be detected in the considered target present image. A mine which is present in one image frame is highly unlikely to disappear from the following image frame. To do this optimization, we consider only the no optimized results. In fact, if we used the optimized results, the detector will never stop detecting after five detections! But this

optimization leads to a stop of detection five images after the last image where a mine has been detected. This can lead to an increase of the false alarm rate especially when the last image where a mine had been detected is the last image of the mine presence region. On the test results, we can see that the recognition rate has improved while the false alarm rate has decreased (see Table 3). On Fig. 14, 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. 14 shown that we managed to increase the recognition rate (65.52 %) and we decrease the false alarm rate (10.13 %).

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. 15. 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 (see Fig. 15). But this constraint can be avoided when making 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 test than in the previous (Fig.16). Results obtained, in terms of recognition and false alarm probabilities are given in Table 4. The visualization of robustness of the optimized nonlinear fringe-adjusted JTC is shown in Fig. 17. The nonzero order fringe adjusted JTC still provides a high false alarm rate while the nonzero nonlinear fringe adjusted JTC presents a medium recognition rate. The optimization allows an increase of the recognition rate and a decrease of the false alarm rate, in comparison with the nonlinear fringe adjusted JTC.

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. For that we adapted a nonlinear method to the Fringe adjusted JTC. The we used a adapted correction criterion.

On the face recognition application, we show that a rotation angle of 15° 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 changing the value of parameter k .

With the underwater mine recognition application, we apply these novel techniques to a real and complex problem. Results show that the JTC architecture gives promising results. There is still work to do to obtain architecture and a criterion optimized to real problems.

ACKNOWLEDGEMENTS

The raw mine images contained in this publication are courtesy of the French State. The data was provided by the GESMA (Groupe d'Etudes Sous Marines de l'Atlantique) within the TOPVISION project coordinated by Thales Underwater Systems SAS. This project is related to Techno-Vision Program launched by the French Ministry of Defense. More detailed information can be found on <http://topvision.gesma.fr> site.

This work is supported in part by the Regional Council of Brittany.

REFERENCES

- [1] Alfalou A., Brosseau C., "Understanding Correlation Techniques for Face Recognition: From Basics to Applications," *Face Recognition*, Milos Oravec (Ed.), ISBN: 978-953-307-060-5, INTECH. Available from:
<http://sciyo.com/articles/show/title/understanding-correlation-techniques-for-face-recognition-from-basics-to-applications>
- [2] VanderLugt V., "Signal detection by complex spatial filtering," *IEEE Trans. Info. Theory*, IT-10, 139-145 (1964).
- [3] Weaver C.S., Goodman J.W., "A Technique for Optically Convolvering Two Functions," *Applied Optics*, 5, 1248-1249 (1966).
- [4] Léonard I., Alfalou A., Alam M., "Nonlinear fringe-adjusted JTC-based face tracking using an adaptive decision criterion," *Proceedings of SPIE 8134* (2011).

- [5] Alam M.S., Karim M.A., "Fringe-adjusted joint transform correlation," *Applied Optics*, 32, 4344-4350 (1993).
- [6] Horner J.L., "Metrics for assessing pattern-recognition performance," *Applied Optics*, 31, 165-166 (1992).
- [7] Léonard I., Alfalou A., Brosseau C., "Spectral optimized asymmetric segmented phase-only correlation filter", *Applied Optics* (2012).
- [8] Javidi B., Kuo C.J., "Joint transform image correlation using a binary spatial light modulator at the Fourier plane," *Applied Optics*, 27, 663-665 (1988).
- [9] Guibert L., Keryer G., Serval A., Attia M., MacKensie H.S., Pellat-Finet P. and de Bougrenet de la Tocnaye J.-L., "On-board optical joint transform correlator for real-time road sign recognition", *Opt. Eng.*, 34, 1, 135-143 (1995)
- [10] Li C.T., Yin S., Yu F.T.S., "Nonzero-order joint transform correlator," *Opt. Eng.*, 37, 1, 58-65 (1998).
- [11] M. S. Alam, "Fractional power fringe-adjusted joint transform correlation," *Journal of Optical Engineering*, Vol. 34, p. 3208-3216, 1995
- [12] M. S. Alam and M. A. Karim, "Multiple target detection using a modified fringe-adjusted joint transform correlator," *Journal of Optical Engineering*, Vol. 33, p. 1610-1617, 1994
- [13] Leonard I., Arnold-Bos A., Alfalou A., "Interest of correlation-based automatic target recognition in underwater optical images: theoretical justification and first results," *Proceedings of SPIE Ocean Sensing and Monitoring II*, 7678 (2010).

- [14] Duntley S. Q. "Light in the Sea", JOSA, Vol. 53, Issue 2, pp. 214-233 (1963)
- [15] Gourier N., Hall D. and Crowley J. L., "Estimating Face Orientation from Robust Detection of Salient Facial Features," Proc. of Pointing 2004, ICPR, International Workshop on Visual Observation of Deictic Gestures (2004).
- [16] Alam M.S. and Bal A., "Dynamic target tracking with fringe-adjusted joint transform correlation and template matching," Applied Optics, 43, 4874-4881 (2004).
- [17] Arnold-Bos A., Malkasse J.P., Kervern G., "Towards a model-free denoising of underwater optical images," in Proceedings of the IEEE conference on Ocean (Europe) (2005).
- [18] Oppenheim A., Lim J., "The importance of phase in signals," in Proceedings of the IEEE, 69,5,529-541 (1981).

LIST OF TABLES

Table 1: Comparison of the binary JTC and the fringe adjusted JTC on images with underwater noises

Table 2: Results of the different JTCs on mine recognition

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

Table 4: Robustness results of the different JTCs

TABLES

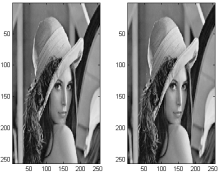
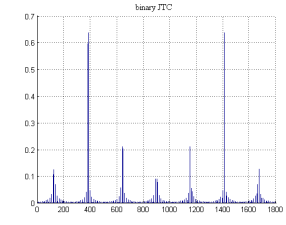
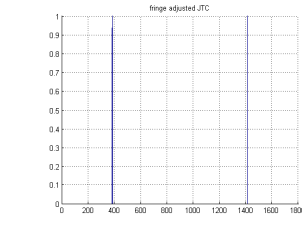
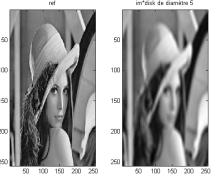
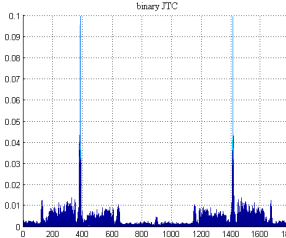
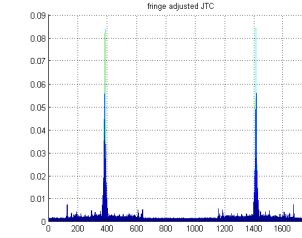
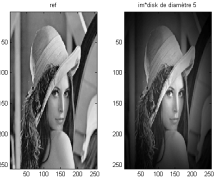
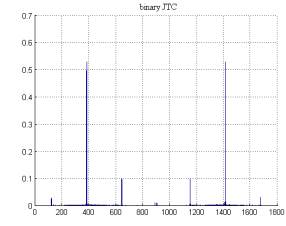
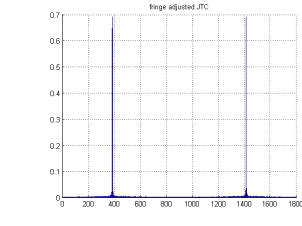
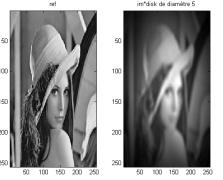
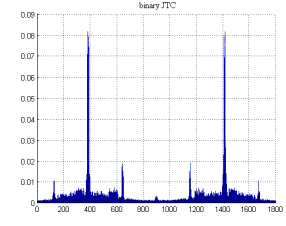
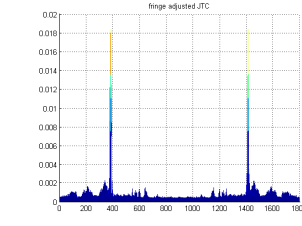
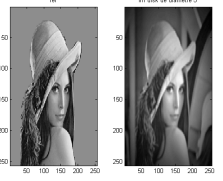
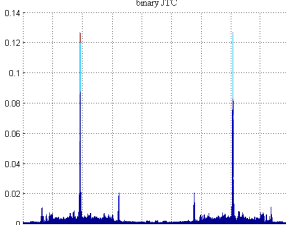
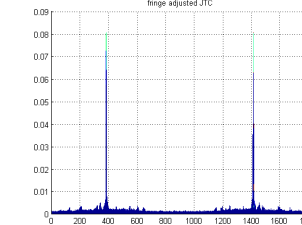
Reference and target images	Correlation plane, binary JTC	Correlation plane, fringe adjusted JTC
 <p>a</p>		
 <p>b</p>		
 <p>c</p>		
 <p>d</p>		
 <p>e</p>		

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 non-uniform light. e: Modified Lena- Lena with blurred and non-uniform light

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

Table 2: Results of the different JTCs on mine recognition

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

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

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

Table 4: Robustness results of the different JTCs

LIST OF CAPTIONS

Fig. 1: Diagram explaining the light propagation in an underwater medium.

Fig. 2: Manta mine, image no. 1082

Fig. 3: Manta mine, image no. 1090

Fig. 4: Lena

Fig. 5: Comparison of JTCs with zero order (left column) and non zero 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. 6: 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

Fig. 7: Examples of facial rotation (-90° to $+90^\circ$). Subject number (1), Base PHPID [15].

Fig. 8: 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.

Fig. 9: Results obtained using the adaptive fringe-adjusted JTC

Fig. 10: Results obtained with a non-linearity in the Fourier plane of the FA-JTC and the new criterion

Fig. 11: Computation of our new criterion

Fig. 12: Results obtained with a non-linearity in the Fourier plane of the FA-JTC and the last criterion

Fig. 13: Initial image, phase image, phase and wavelet image

Fig. 14: Results of the optimized nonlinear fringe-adjusted JTC, $k=0.85$

Fig. 16: Difference between mine size in the 2 video sequences

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

FIGURES

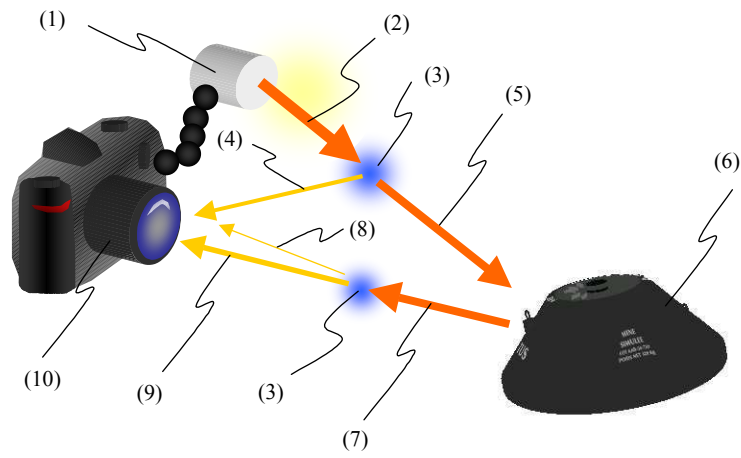


Fig. 1: Diagram explaining the light propagation in an underwater medium.



Fig. 2: Manta mine, image no. 1082



Fig. 3: Manta mine, image no. 1090

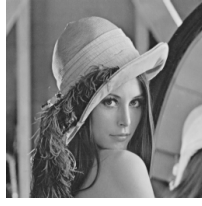
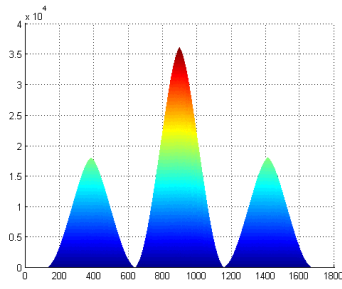
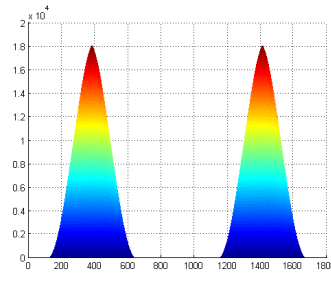


Fig. 4: Lena

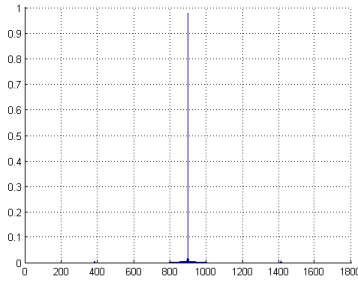


(a)

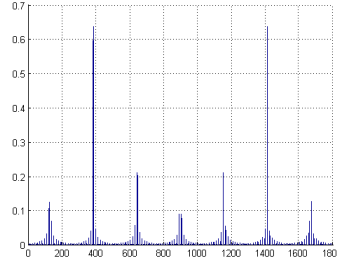


(b)

$$PCE_C=3 \times 10^{-4}$$

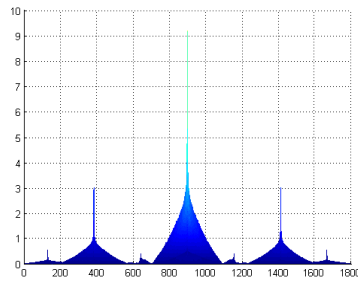


(c)

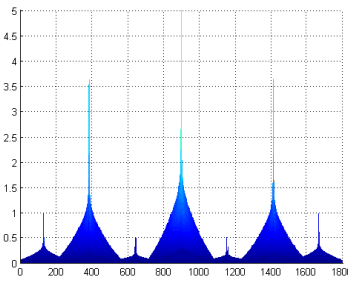


(d)

$$PCE_B=0.8106$$

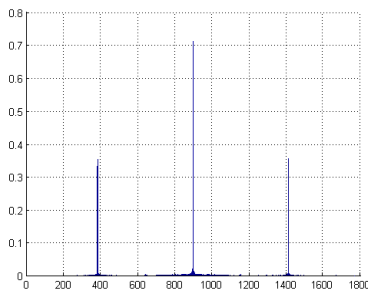


(e)

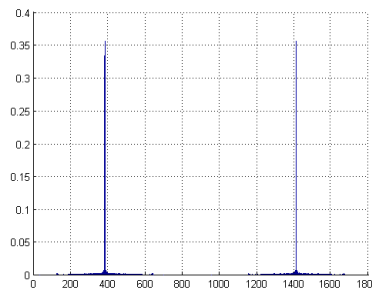


(f)

$$PCE_{NL}=0.0036$$



(g)

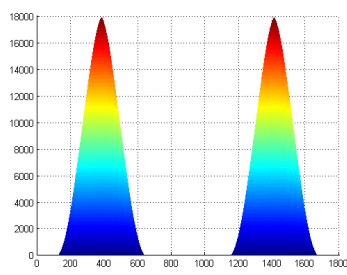


(h)

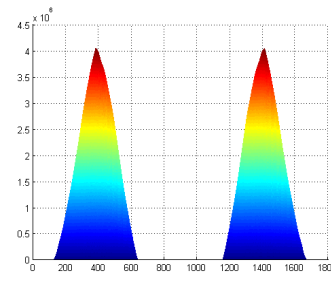
$$PCE_{FAF}=0.6286$$

Fig. 5: Comparison of JTCs with zero order (left column) and non zero JTCs (right column).

(a) and (b): classical JTC. (c) and (d): binary JTC. (e) and (f): nonlinear JTC. (g) and (h): fringe adjusted JTC



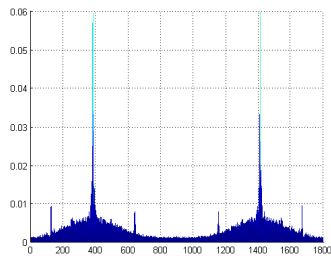
$$PCE_C=2.9 \times 10^{-4}$$



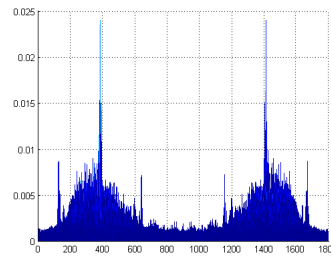
$$PCE_C=2.74 \times 10^{-4}$$

(a)

(b)



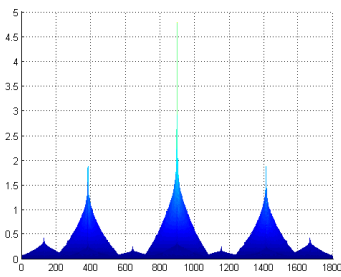
$$PCE_B=0.099$$



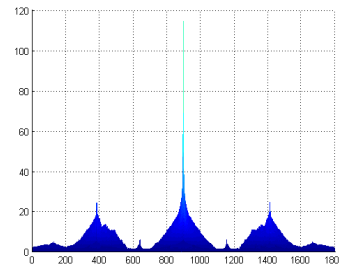
$$PCE_B=0.0020$$

(c)

(d)



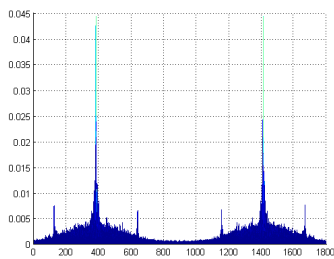
$$PCE_{NL}=0.0017$$



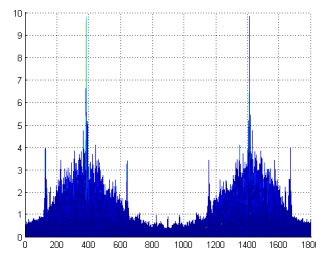
$$PCE_{NL}=0.0012$$

(e)

(f)



$$PCE_{FAF}=0.0145$$



$$PCE_{FAF}=0.0018$$

(g)

(h)

Fig. 6: 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



Fig. 7: Examples of facial rotation (-90° to $+90^\circ$). Subject number (1), Base PHPID [15].

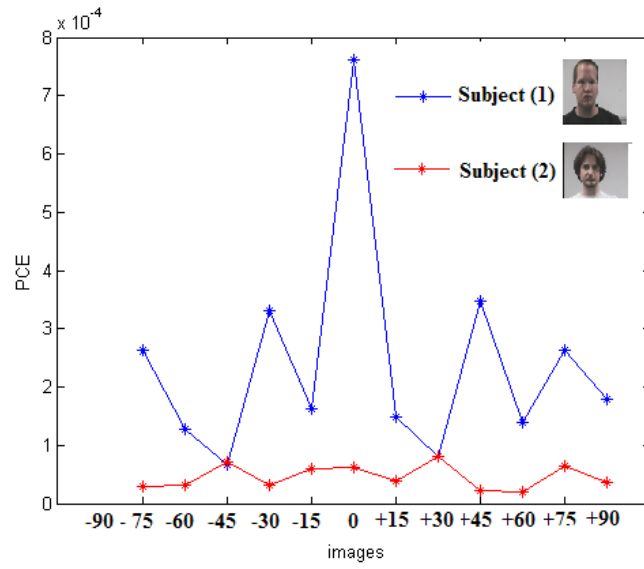


Fig. 8: Results obtained with a FA-JTC using only one reference image position without any rotation in the H_{FAF} Filter to define the background noise.

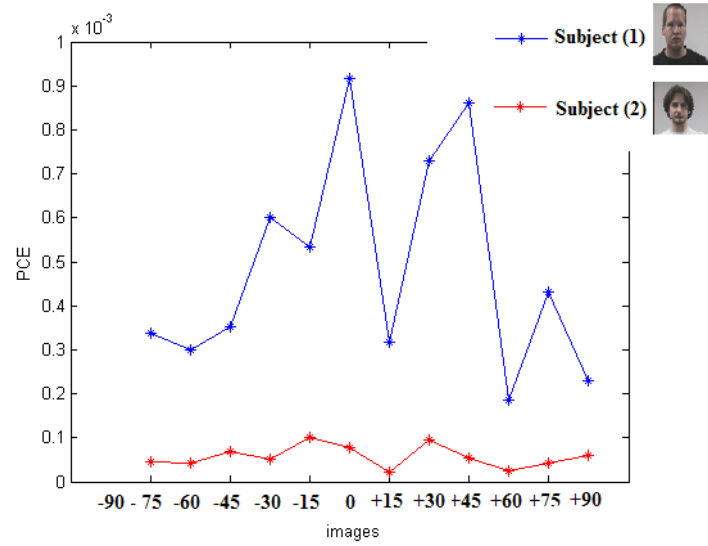


Fig. 9: Results obtained using the adaptive fringe-adjusted JTC

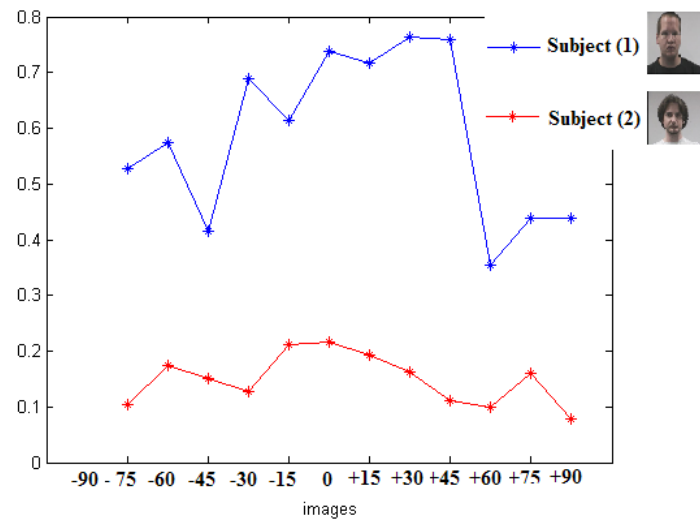


Fig. 10: Results obtained with a non-linearity in the Fourier plane of the FA-JTC and the new criterion

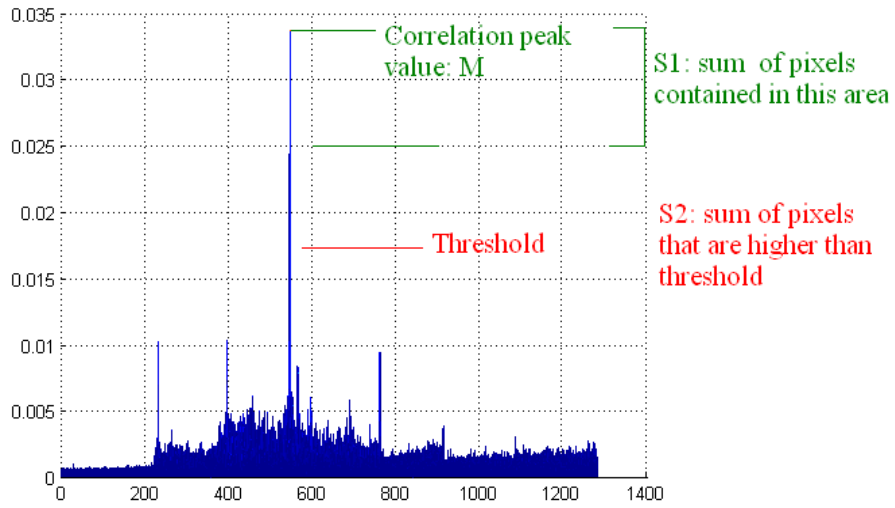


Fig. 11: Computation of our new criterion

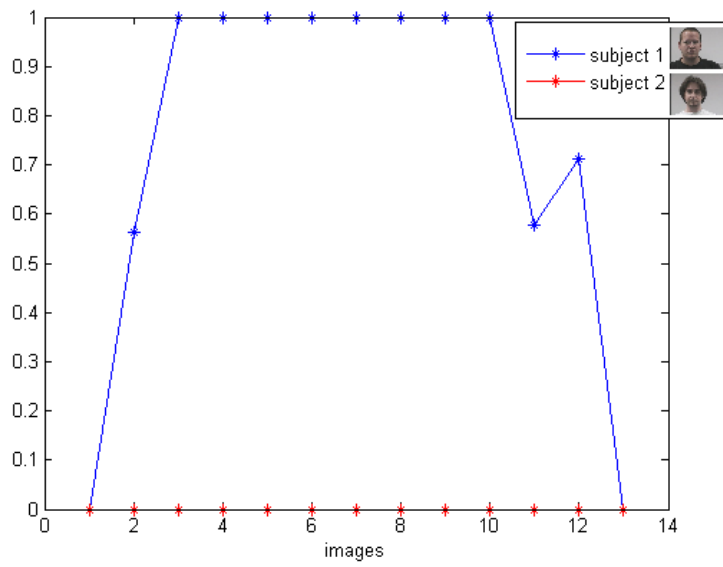


Fig. 12: Results obtained with a non-linearity in the Fourier plane of the FA-JTC and the last criterion

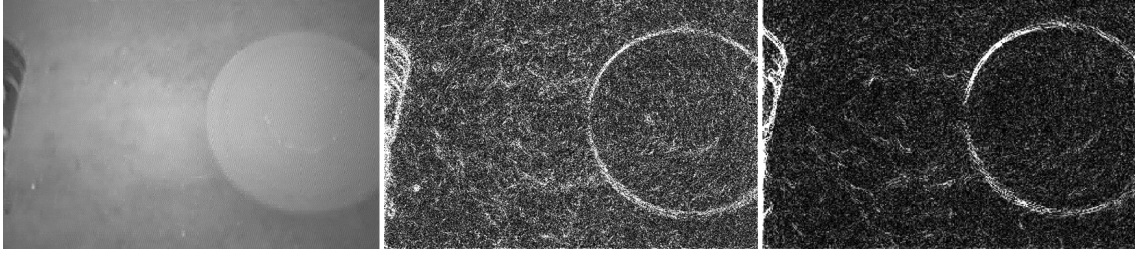


Fig. 13: Initial image, phase image, phase and wavelet image

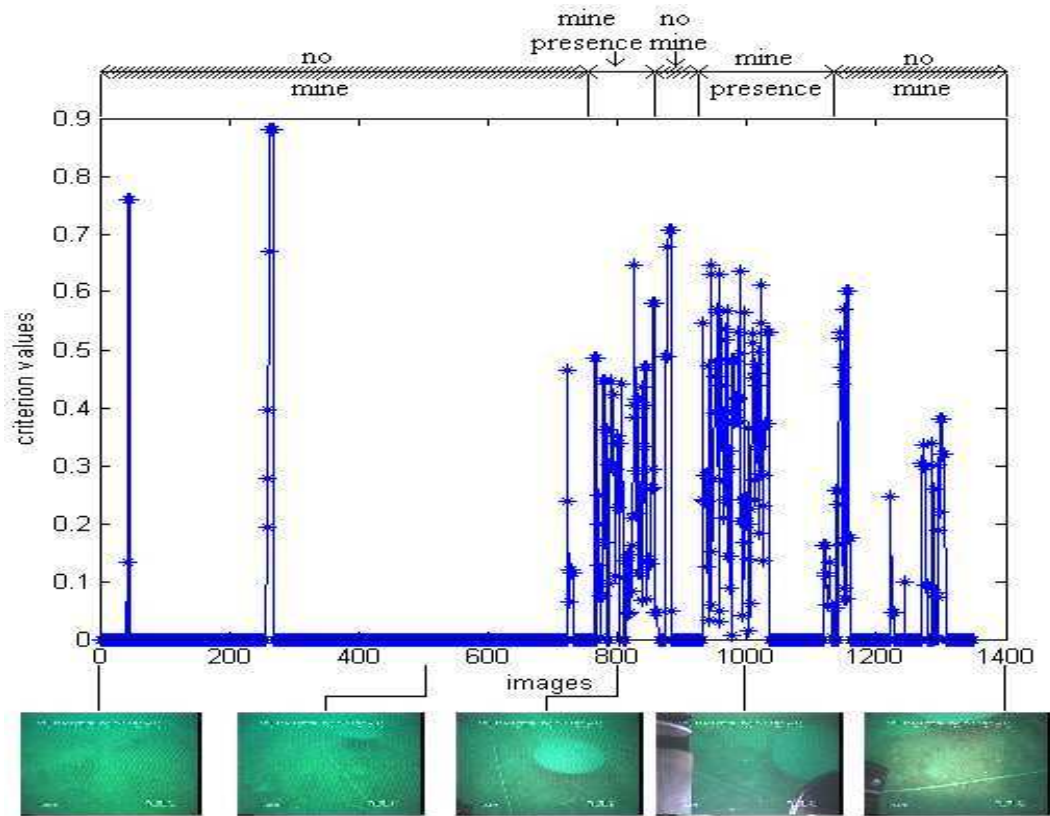


Fig. 14: Results of the optimized nonlinear fringe-adjusted JTC, $k=0.85$



Fig. 15: Images with presence of the camera support



Fig. 16: Difference between mine size in the 2 video sequences

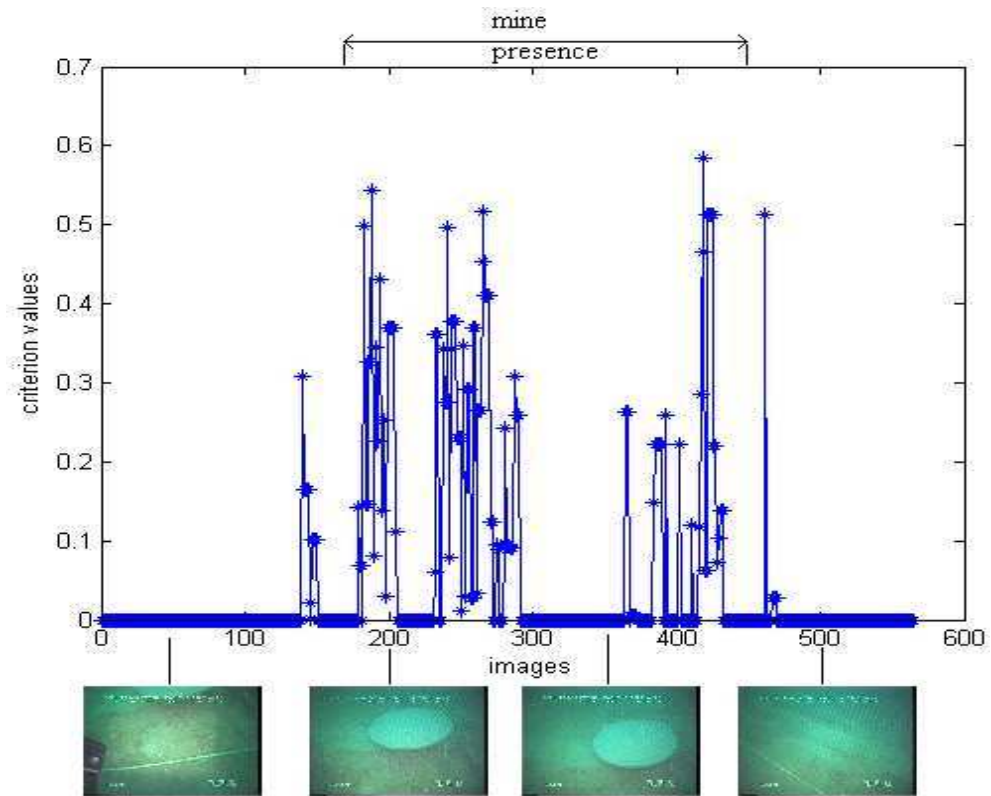


Fig. 17: Results of the optimized nonlinear fringe-adjusted JTC, $k=0.85$