Automatic Digital Hologram Denoising by Spatiotemporal Analysis of Pixel-Wise Statistics

M. Leo, C. Distante, Member, IEEE, M. Paturzo, P. Memmolo, M. Locatelli, E. Pugliese, R. Meucci, Member, IEEE, and P. Ferraro, Senior Member, IEEE

Abstract—In this paper, a new technique to reduce the noise in a reconstructed hologram image is proposed. Unlike all the techniques in the literature, the proposed approach not only takes into account spatial information but also temporal statistics associated with the pixels. This innovative solution enables, at first, the automatic detection of the areas of the image containing the objects (foreground). This way, all the pixels not belonging to any objects are directly cleaned up and the contrast between objects and background is consistently increased. The remaining pixels are then processed with a spatio-temporal filtering which cancels out the effects of speckle noise, while preserving the structural details of the objects. The proposed approach has been compared with other common speckle denoising techniques and it is found to give better both visual and quantitative results.

Index Terms—Holography, image denoising, statistical learning, adaptive filters, image segmentation.

I. INTRODUCTION

I

N DIGITAL holography (DH), a CCD camera is used to record the holograms without wet chemical processing which is complex and time-consuming [1]. The continuous spatial distribution of an optically generated hologram is sampled by the discrete sensitive pixels on a CCD array, whose outputs are converted to the digitized signals and stored in an image processing system for numerical evaluation.

Digital Holography uses coherent source and, when the light is diffused by an optically rough surface, speckle noise is generated. It degrades the quality of the rendered images making the accurate viewing of small details very difficult. Moreover, speckle noise is emphasized by the limited CCD pixel size. It follows that it is very important to search for an effective way to reduce the speckle noise in the DH [2] but, unfortunately, its reduction is a challenge because of the badly statistical regularity of the noise.

A variety of approaches have been proposed to overcome this drawback: some of them are based on image processing techniques [12], others on optical techniques [13]. Optical techniques are based on the definition of specific acquisition setup. Instead, digital image processing techniques do not affect the acquisition setup but they try to reduce the noise by numerical processing of the recorded data. Non-adaptive filters (e.g., mean or median filter [12]) have been largely used for this purpose. To reduce the impact of filtering in image areas containing texture and edges, adaptive filters have also been introduced (e.g., Frost filter [16], Lee filter [15], Discrete Fourier Filtering [3], non-local means filtering [14], etc.).

Overall, the techniques in the literature try to identify the contribution of the noise in each pixel through the only information of spatial proximity in a single image. Unfortunately, as evidenced from the literature, this does not always allow to properly separate the noise from the information content of the holographic images.

In this paper a new technique to reduce the noise in a reconstructed hologram image is therefore proposed: the key idea is to extract additional useful information, for noise detection and suppression, through the observation of temporal numerical values at each pixel location in a sequence of images.

The main novelty of the proposed solution is that it combines both spatial information and statistics along the time to detect and suppress the speckle noise as well as to preserve the information content of holograms. This is achieved in two steps: the automatic detection of the areas of the image containing the objects (foreground) is firstly performed. This way, all the pixels not belonging to any objects are directly cleaned up and the contrast between object and background is consistently increased. The remaining pixels are then processed with a spatio-temporal filtering which cancel out the effects of speckle, while preserving the structural details of the objects.

In the full digital holographic process, two kinds of noise have to be handled: the first one is an additive uncorrelated noise that corrupts the observed irradiance, the other one is the speckle noise [17].

In the proposed approach, the first computational step, that detects the region in which the object is contained, removes the contribution of additive noise in the background. The second step, performing spatio-temporal filtering, deals instead with the speckle noise reduction.

The proposed approach has been tested on an holographic sequence of the Perseus statue captured at long wavelength infrared (IR) range (10.6 μm) in off-axis Fourier configuration. The statuette is about 33 cm in height. The recording distance, that is the distance between the object and the camera is 80 cm. As a coherent light source, a 110 W—CW CO2 laser emitting at 10, 6 μm, is used. In this case, the speckle size is about 50 microns. In fact, the speckle size along one direction in the observation plane is proportional to the incident wavelength, and

IEEE Xplore
the distance between the object and the observation plane, is inversely proportional to the maximal illuminated dimension of the object parallel to that direction. Qualitative and quantitative comparisons with the leading approaches of the state of the art demonstrated the effectiveness of the proposed approach.

II. THE PROPOSED APPROACH

The theoretical basis of the proposed approach is the adaptive mixture models theory proposed in [4] for visual tracking purpose.

The mixture models theory can be summarized as follows: at any time \( t \) what is known about a particular pixel \( \{x_0, y_0\} \) is its history:

\[
\{X_1(x_0, y_0), \ldots, X_t(x_0, y_0)\} = \{I(x_0, y_0; i) : 1 \leq i \leq t\}
\]

where \( I \) is the image sequence.

The temporal occurrences of the values of each pixel can obviously be statistically modeled: the simplest way to do this is through a Gaussian distribution centered at the mean pixel value. In order to take into account also noise, a more sophisticated model must be introduced. In other words, a mixture of Gaussian distributions models the recent history of each pixel. In this way the probability of observing the current pixel value in the position \( (x_0, y_0) \) is

\[
P(X_t) = \sum_{i=1}^{k} w_{i,t} \eta(X_t, \mu_{i,t}, \Sigma_{i,t})
\]

where \( k \) is the number of distributions, \( w_{i,t} \) is an estimate of the weight of the \( \mu_{i,t} \) Gaussian in the mixture at time \( t \), \( \mu_{i,t} \) is the mean value of the Gaussian in the mixture at time \( t \), and \( \Sigma_{i,t} \) is the covariance matrix of the Gaussian at time \( t \), and \( \eta \) is a Gaussian probability density function:

\[
\eta(X_t, \mu_{i,t}, \Sigma_{i,t}) = \frac{1}{(2\pi)^{n/2} \left| \Sigma_{i,t} \right|^{1/2}} e^{-\frac{1}{2} \left( X_t - \mu_{i,t} \right)^T \Sigma_{i,t}^{-1} \left( X_t - \mu_{i,t} \right)}.
\]

In this way a mixture of Gaussians (MOG) characterizes the distribution of observed values of each pixel, especially when the pixel is subjected to visit several states as shown below.

Every new pixel value \( X_t \) is then used to update the model. In particular, if among the \( k \) Gaussians there is one having the new pixel value within 2.5 standard deviation of the distribution (matching model), then its parameters are updated as follows:

\[
\mu_{i,t} = (1 - \rho) \mu_{i,t-1} + \rho X_t
\]

\[
\sigma_{i,t}^2 = (1 - \rho) \sigma_{i,t-1}^2 + \rho (X_t - \mu_{i,t})^2 (X_t - \mu_{i,t})
\]

where

\[
\rho = \alpha \eta(X_t, \mu_k, \sigma_k).
\]

The \( \mu \) and \( \sigma \) parameters for the other distributions that do not satisfy the matching condition remain the same.

The prior weights of the \( k \) distributions at time \( t \), \( w_{k,t} \), are adjusted as follows:

\[
w_{k,t} = (1 - \alpha)w_{k,t-1} + \alpha(M_{k,t})
\]

where \( \alpha \) is the learning rate and \( (M_{k,t}) \) is 1 for the model which match and 0 for the remaining models.

If none of the \( k \) distributions matches the value of the current pixel, the least probable distribution is replaced with a distribution having the current value as its mean value, an initially fixed high variance, and low prior weight.

In this paper the above mixture models theory has been adapted for the denoising purpose in the holographic reconstructed image.

In particular, the proposed approach consists of a preliminary phase involving the selection of a small number of pixels not belonging to any object placed within the depth of field of the holographic acquisition setup. Using the terminology of the research area identified as machine learning, these pixels can be considered the reference set \( T = \{X_i\} i = 1 \ldots N \) that allow the system to learn the temporal behaviors of non-object pixels. During the acquisition phase, at the time instant \( t \), a statistical model is built and updated according to (2)–(7) for each pixel of the holographic image.

In addition, the likelihood of the set \( T \) is computed for each pixel according to the probability computation in case of compatible events as follows

\[
P_G(X_t) = \sum_{i=1}^{N} P(X_t^i) - \sum_{i=2}^{N} \prod_{k=1}^{N} P(\sigma_i(X_1^i, \ldots X_t^i))
\]

where \( \sigma_i(x) \) states for the all possible permutations of the elements in \( T \) taken \( i \) at a time. In this way a likelihood image is generated and an adaptive threshold can be used to generate a binary image where pixels most probably belonging to the object under consideration are set to 1 and all the remaining pixels are set to 0.

The resulting binary image is then processed in order to detect connected regions and to discard isolated pixels [5]. The outcome of the region connectivity analysis is a binary mask \( M \). This mask is then superimposed to the original holographic image \( I \) and, in this way, a new intermediate holographic image is obtained:

\[
\tilde{I}(x, y) = \begin{cases} 
I(x, y) & \text{if } M(x, y) = 1 \\
0 & \text{if } M(x, y) = 0 
\end{cases}
\]

The image \( \tilde{I} \) is finally spatial-temporally further filtered in order to smooth the speckle effects on the region belonging to the objects under observation. The final image outcome \( I \) of the proposed denoising approach is obtained in this way. This filtering is done by replacing at time \( t \) the pixel value in \( \tilde{I}(x, y) \) with the value of its temporal statistics averaged with those of the 4-distance neighboring pixels. This is carried out for all non-zero pixels in \( \tilde{I} \) whereas zero pixels remain unchanged. This can be formalized as follows:

\[
\hat{I} = \begin{cases} 
0 & \text{if } \tilde{I}(x, y) = 0 \\
\frac{1}{25} \sum_{i=1}^{4} \sum_{j=1}^{4} \mu_{t}(x + i, y + j) & \text{if } \tilde{I}(x, y) \neq 0
\end{cases}
\]

where \( \mu_t \) is the temporal statistics of the object under consideration.
III. EXPERIMENTAL RESULTS

The proposed approach has been tested on a holographic image sequence of the Perseus statue [9], [10]. The holograms were captured at long wavelength infrared (IR) range (10.6 μm) in off-axis Fourier configuration [10]. A Miricle Thermoteknix 307k, 640 × 480 detector resolution and 25 μm pitch was used.

At first, analyzing over time the statistical distribution of a set of pixels in the image affected by the speckle noise has allowed the choice of a suitable Gaussian mixture model. Fig. 1 shows the distribution of a noisy pixel belonging to the foreground, where it is possible to notice a multimodal behavior of the intensity values for a time window of 100 temporally contiguous frames.

This preliminary study has demonstrated that at least two Gaussians are needed to describe the dynamics of pixel values. However, in order to correctly model also very infrequent values occurrences (that are usually associated to the speckle) a 3-Gaussian model has been used in the following experiments.

Conversely, a greater number of Gaussians has been experimentally proven to be useless since the models generated on the reference points experienced very low values of coefficients associated with the additional Gaussians.

In Fig. 2(a), a photo of the Perseus statue is shown whereas, Fig. 2(b) reports the 100th image of the recorded holographic sequence. Speckle noise is evident and strongly affects the holographic image quality.

In Fig. 3(a) the likelihood image relative to the 100th image obtained by using a set $T$ of 20 reference pixels not belonging to the statue area is shown. In this figure, likelihood values are computed as described in (8) and then rescaled in the range $[0, 255]$ for display purposes: the lightest pixels are related to areas of the background (highest non-object likelihood values) and the darkest pixels are instead related to areas belonging to the statue (lowest non-object likelihood values).

Fig. 3(b) reports the corresponding segmented binary image extracted using a threshold ($th = 0.5$) on the likelihood values.
In Fig. 4(a) the binary mask obtained after topological analysis of the segmented image is shown. Fig. 4(b) presents the new holographic image obtained as described in equation (9).

In Fig. 5(a) the image obtained after the spatiotemporal filtering detailed in (10) is shown.

The outcomes of the proposed approach have then been qualitatively compared with those obtained using latest state of the art spatial speckle denoising techniques. In particular Figs. 5(b)-(f) show the hologram denoised by a median filter (MF) [17], a Discrete Fourier Filter (DFF) [3], a nonlocal means filter (NLM) [14], a Frost filter [16] and a Lee Filter [15] respectively.

In Fig. 5 the different outcomes obtained by: (a) proposed approach; (b) a median filter; (c) a discrete Fourier Filter; (d) a non-local mean filtering; (e) a Frost filter; and (f) a Lee filter.

All in all, from Fig. 5, the superiority of the proposed denoising approach is evident. Its outcome is in fact much clearer than those obtained by using other denoising strategies: on the one hand the noise is completely suppressed in the areas outside the object and strongly attenuated on the regions of the object and, on the other hand, the details of the object appear preserved allowing a high visual quality.

For a qualitative comparison of the various filters, three different metrics have been considered:

1) Speckle Index [8], [9] on (quasi) homogenous areas of the object under consideration;
2) Contrast of the whole image (Global Contrast) [6];
3) Contrast [7] of a small region containing some details of the object (Local Contrast) [7].

Speckle index is a measure of the level of noise in the image: it is computed on a uniform region of the object in order to be sure that every observed deviation from the expected reference value in the pixels is due to noise. That said, it is obvious that the lower the index, the less noisy is the image and in particular, in an ideally uncorrupted image the speckle index value is 0.

Global Contrast indicates how much the object is enhanced with respect to the whole content of the image, whereas, Local Contrast, is a measure of how much small details in the object are preserved in the holographic reconstruction.

Both the global and the local contrast should be kept as high as possible: high values are in fact the numerical evidence that there is variability in the representation of the structures constituting the object under observation.

From the computation point of view, the Speckle index value is computed as

$$S_k = \sqrt{\frac{\sigma}{\mu}}$$  \hspace{1cm} (11)

where $\sigma$ and $\mu$ are the standard deviation and the mean computed in the (quasi) uniform considered region included in the green rectangle in Fig. 6.

Contrast measurements are instead computed as:

$$C = \frac{\sigma}{(\mu)^2}$$  \hspace{1cm} (12)
where \( n \) is a positive number (more precisely \( n = 2 \) in the following), \( \sigma \) is the standard deviation and \( \alpha_4 \) is the kurtosis measure defined by

\[
\alpha_4 = \frac{\mu_4}{\sigma^4}
\]

where \( \mu_4 \) is the fourth moment about the mean and \( \sigma^2 \) is the variance.

As mentioned above, the numerator and denominator in (12) are computed on the whole image for Global Contrast measure and on a small selected region for Local Contrast respectively. The region chosen for Local Contrast computation is included in the red rectangle in Fig. 6: this choice has been made considering that this region contains some details of the face that should be preserved.

A comparison between the six filters in terms of the above metrics is shown in Table I, which provides the numerical evidence that the proposed approach achieved the best results in terms of all the considered metrics.

<table>
<thead>
<tr>
<th></th>
<th>Speckle Index ( S_i )</th>
<th>Global Contrast ( C_g )</th>
<th>Local Contrast ( C_l )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>0.85</td>
<td>1.04</td>
<td>0.75</td>
</tr>
<tr>
<td>Median Filtering</td>
<td>0.69</td>
<td>0.93</td>
<td>0.49</td>
</tr>
<tr>
<td>Discrete Fourier Filter [17]</td>
<td>0.82</td>
<td>1.23</td>
<td>0.43</td>
</tr>
<tr>
<td>NonLocal Mean [14]</td>
<td>0.48</td>
<td>0.84</td>
<td>0.33</td>
</tr>
<tr>
<td>Frost Filter [16]</td>
<td>0.54</td>
<td>0.87</td>
<td>0.39</td>
</tr>
<tr>
<td>Lee Filter [15]</td>
<td>0.58</td>
<td>0.89</td>
<td>0.46</td>
</tr>
<tr>
<td>Intermediate Image ( \tilde{i} ) (eq. 9 in section II)</td>
<td>0.85</td>
<td>1.39</td>
<td>0.83</td>
</tr>
<tr>
<td>Proposed Approach</td>
<td>0.45</td>
<td>1.26</td>
<td>0.49</td>
</tr>
</tbody>
</table>

This result should be placed in relation with the average time per pixel observed for the application of conventional techniques. In particular the median filtering [17] takes on average of \(~3.9 \times 10^{-4} \) s for each pixel, the Discrete Fourier Filtering [3] \(~9.44 \times 10^{-5} \) s, the NonLocal Mean filtering \(~1.1 \times 10^{-3} \) s, the Frost filter \(~1.12 \times 10^{-4} \) and finally the Lee filter \(~1.49 \times 10^{-6} \). Although these data demonstrate the comparability of the computational time of the proposed technique with most of the conventional ones, it should be emphasized that the code used for the proposed technique could be further computationally enhanced (for example, using built-in Matlab functions for multidimensional data handling).

Furthermore the computation required for the proposed technique is largely parallelizable (in particular (8) and (9) are embarrassing parallel since they may be parallel performed for each pixel).

### IV. Conclusion

In this paper a new and innovative technique to reduce the noise in a reconstructed hologram image has been proposed. In order to overcome the limitation of the techniques in the literature, both spatial information and temporal statistics associated with individual pixels are used to detect and suppress the speckle noise as well as to enhance the information content.

Experimental results on an holographic image sequence captured at long wavelength infrared demonstrated that the proposed approach gives better results in both visual quality of the cleaned images and quantitative analysis of numerical outputs in terms of three different metrics.

The analysis and comparison of processing time has also been conducted demonstrating the comparability with the state-of-the-art techniques.

Future works will be addressed to further improve reconstructed hologram image by introducing iterative region growing strategies that will allow overcoming segmentation errors that occurred at the boundaries of the object regions.

Moreover a refinement of the analysis of the pixel statistics will be experimented in order to better preserve details on the reconstructed object.

### References


Marco Leo received the Honors degree in computer science engineering from the University of Salento, Lecce, Italy, in 2001. From May 2001 to August 2012, he was a Researcher with the Institute of Intelligent Systems for Automation (ISSIA-CNR) in Bari, Italy. Since September 2012 he has been a Researcher with the Institute of Optics (INO-CNR) in Lecce, Italy. His main research interests are in the fields of image processing, image analysis, computer vision, pattern recognition, digital signal processing, neural networks, graphical models, wavelet transform, and independent component analysis. He participated in a number of national and international research projects facing automatic video surveillance of indoor and outdoor environments, human attention monitoring, real-time event detection in sport contexts, and nondestructive inspection of aircraft components. He is the author of more than 100 papers in national and international journals and conference proceedings. He is also a co-author of three international patents on visual systems for event detection in sport contexts.

Cosimo Distante received the degree in computer science from the University of Bari, Bari, Italy, in 1997, and the Ph.D. in engineering in 2001 by working in the field of artificial intelligent systems.

Dr. Distante has been Visiting researcher at the Computer Science Department of the University of Massachusetts (Umass at Amherst, MA) where he worked in the context of robot learning. In 1998, he served as a teaching assistant of the Artificial Intelligence course of the Master in Science of the University of Massachusetts. In 2000, he received the Ph.D. in engineering from the University of Salento Thesis: “Intelligent Sensors and Systems”. Since 2003, he is Contract Professor for the courses of Pattern Recognition and Image Processing in Computer Engineering at the University of the Salento. He is expert member for the panel in Innovation Technology of the Ministry of the Economic Development. Dr. Distante is currently with the National Institute of Optics of the CNR where he is responsible for the following research themes: optical devices and ICT methods for industrial applications. His main research interests are in the field of Pattern Recognition and Artificial Intelligence applied to the following sectors: video surveillance, robust estimation, medical imaging and manufacturing.

Melania Paturzo is research scientist at INO-CNR. Her research interest is in the field of Optics. Her research activities concerns the investigation of material properties by means of interferometric techniques, fabrication of periodically poled lithium niobate samples by electric field poling process, development of optical techniques to get super-resolution in digital holographic microscopy, holographic display, quantitative phase-contrast microscopy for cells investigation, optofluorics. She is author and/or coauthor of more than 40 papers in international peer reviewed journals and of more than 50 conference papers.

Pasquale Memmolo was born in Avellino, Italy, in 1982. In 2005, he received the first level degree in computer engineering. In 2008, he received the second level degree in communication engineering, with final mark 110/110, at the “Università degli Studi di Napoli Federico II” (Napoli, Italy), working on numerical and optical multiplexing and de-multiplexing methods in digital holography. In 2011, he completed the Ph.D. in electronic and communication engineering, with a thesis entitled “Compressed Sensing: a new framework for signals recovery and its application in Digital Holography”. Now he works as a Post Doctoral Junior at the Center for Advanced Biomaterials for Health, Care@CRIB, Istituto Italiano di Tecnologia, Napoli, Italy, and he is a collaborator at the CNR—Istituto Nazionale di Ottica, Pozzuoli, Italy. He is also co-author of more of 30 international journal’s publications and proceedings papers.

Massimiliano Locatelli graduated in physics at Florence University, in 2009, with a Thesis on ‘Digital Holography Applications’, developed at INO-CNR. He got a Ph.D. at LENS (European Laboratory for Non-linear Spectroscopy) University of Florence, in 2013, with a thesis on ‘Mid Infrared digital holography and Terahertz imaging’. He is actually a researcher at LENS and INO-CNR of Florence. His research activity is conducted in the field of Optics (in particular Infrared and Terahertz imaging and digital holography), Biophysics and Complex System Physics.

Eugenio Pugliese received the Physics degree from the University of Florence, Italy, with a thesis on Mossbauer Spectroscopy, and is currently a Ph.D. student in nonlinear dynamics and complex systems at the same university. He carries out its research activities in the laboratories of INO-CNR in Florence. His research interests are in the fields of chaotic dynamics of classical and quantum systems, digital holography both in visible and IR range and optical investigation of complex systems.

Riccardo Meucci graduated in physics at the University of Florence in 1982 and received the Ph.D. degree in optics at the same University in 1987. From 1984 to 1987, he served as Researcher in Optics at Istituto di Cibernetica of the Italian National Research Council (CNR) Arco Felice (Naples), thereafter at Istituto Nazionale di Ottica (INO). Actually he is Research Director at INO-CNR. His scientific activity mainly includes: chaotic instabilities in single mode lasers, laser transients, dynamical models for the CO2 laser, control of chaos, spatio-temporal instabilities and delayed systems, polarization instabilities and digital holography in the mid-infrared region.

Pietro Ferraro (M’02–SM’07) is currently Chief Research Scientist at Istituto Nazionale di Ottica del CNR, Pozzuoli, NA, Italy. Previously he worked as Principal Investigator with Alenia Aeronautics. He has published 12 book chapters, 170 papers in journals, more than 150 papers at International Conferences. He Edited two books with Springer. He holds 14 patents. Among his current scientific interests are: holography, interferometry, microscopy, fabrication of nanostructures, ferroelectric crystals, optical fiber sensors, fiber bragg gratings, nano-microfluidics, optofluorics. Dr. Ferraro has chaired many International Conferences. He is in the Editorial Board of Optics and Lasers in Engineering (Elsevier), is Topical Editor of Optics Letters. Dr. Ferraro is Fellow of SPIE and Fellow of OSA.
Automatic Digital Hologram Denoising by Spatiotemporal Analysis of Pixel-Wise Statistics

M. Leo, C. Distante, Member, IEEE, M. Paturzo, P. Memmolo, M. Locatelli, E. Pugliese, R. Meucci, Member, IEEE, and P. Ferraro, Senior Member, IEEE

Abstract—In this paper, a new technique to reduce the noise in a reconstructed hologram image is proposed. Unlike all the techniques in the literature, the proposed approach not only takes into account spatial information but also temporal statistics associated with the pixels. This innovative solution enables, at first, the automatic detection of the areas of the image containing the objects (foreground). This way, all the pixels not belonging to any objects are directly cleaned up and the contrast between objects and background is consistently increased. The remaining pixels are then processed with a spatio-temporal filtering which cancels out the effects of speckle noise, while preserving the structural details of the objects. The proposed approach has been compared with other common speckle denoising techniques and it is found to give better both visual and quantitative results.

Index Terms—Holography, image denoising, statistical learning, adaptive filters, image segmentation.

I. INTRODUCTION

In Digital holography (DH), a CCD camera is used to record the holograms without wet chemical processing which is complex and time-consuming [1]. The continuous spatial distribution of an optically generated hologram is sampled by the discrete sensitive pixels on a CCD array, whose outputs are converted to the digitized signals and stored in an image processing system for numerical evaluation.

Digital Holography uses coherent source and, when the light is diffused by an optically rough surface, speckle noise is generated. It degrades the quality of the rendered images making the accurate viewing of small details very difficult. Moreover, speckle noise is emphasized by the limited CCD pixel size. It follows that it is very important to search for an effective way to reduce the speckle noise in the DH [2] but, unfortunately, its reduction is a challenge because of the badly statistical regularity of the noise.

A variety of approaches have been proposed to overcome this drawback: some of them are based on image processing techniques [12], others on optical techniques [13]. Optical techniques are based on the definition of specific acquisition setup. Instead, digital image processing techniques do not affect the acquisition setup but they try to reduce the noise by numerical processing of the recorded data. Non-adaptive filters (e.g., mean or median filter [12]) have been largely used for this purpose. To reduce the impact of filtering in image areas containing texture and edges, adaptive filters have also been introduced (e.g., Frost filter [16], Lee filter [15], Discrete Fourier Filtering [3], non-local means filtering [14], etc.).

Overall, the techniques in the literature try to identify the contribution of the noise in each pixel through the only information of spatial proximity in a single image. Unfortunately, as evidenced from the literature, this does not always allow to properly separate the noise from the information content of the holographic images.

In this paper a new technique to reduce the noise in a reconstructed hologram image is therefore proposed: the key idea is to extract additional useful information, for noise detection and suppression, through the observation of temporal numerical values at each pixel location in a sequence of images. The main novelty of the proposed solution is that it combines both spatial information and statistics along the time to detect and suppress the speckle noise as well as to preserve the information content of holograms. This is achieved in two steps: the automatic detection of the areas of the image containing the objects (foreground) is firstly performed. This way, all the pixels not belonging to objects are directly cleaned up and the contrast between object and background is consistently increased. The remaining pixels are then processed with a spatio-temporal filtering which cancel out the effects of speckle, while preserving the structural details of the objects.

In the full digital holographic process, two kinds of noise have to be handled: the first one is an additive uncorrelated noise that corrupts the observed irradiance, the other one is the speckle noise [17].

In the proposed approach, the first computational step, that detects the region in which the object is contained, removes the contribution of additive noise in the background. The second step, performing spatio-temporal filtering, deals instead with the speckle noise reduction.

The proposed approach has been tested on a holographic sequence of the Perseus statue captured at long wavelength infrared (IR) range (10.6 μm) in off-axis Fourier configuration. The statuette is about 33 cm in height. The recording distance, that is the distance between the object and the camera is 80 cm. As a coherent light source, a 110 W—CW CO₂ laser emitting at 10, 6 μm, is used. In this case, the speckle size is about 50 microns. In fact, the speckle size along one direction in the observation plane is proportional to the incident wavelength, and
the distance between the object and the observation plane, is inversely proportional to the maximal illuminated dimension of the object parallel to that direction. Qualitative and quantitative comparisons with the leading approaches of the state of the art demonstrated the effectiveness of the proposed approach.

II. THE PROPOSED APPROACH

The theoretical basis of the proposed approach is the adaptive mixture models theory proposed in [4] for visual tracking purpose.

The mixture models theory can be summarized as follows: at any time \( t \) what is known about a particular pixel \( \{ x_0, y_0 \} \) is its history:

\[ X_1^{(x_0, y_0)}, \ldots, X_t^{(x_0, y_0)} \]  

where \( I \) is the image sequence.

The temporal occurrences of the values of each pixel can obviously be statistically modeled: the simplest way to do this is through a Gaussian distribution centered at the mean pixel value. In order to take into account also noise, a more sophisticated model must be introduced. In other words, a mixture of Gaussian distributions models the recent history of each pixel. In this way the probability of observing the current pixel value in the position \( \{ x_0, y_0 \} \) is

\[ P(X_t) = \sum_{i=1}^{k} w_{i,t} \eta \left( X_t, \mu_{i,t}, \Sigma_{i,t} \right) \]  

where \( k \) is the number of distributions, \( w_{i,t} \) is an estimate of the weight of the \( i \) Gaussian in the mixture at time \( t \), \( \mu_{i,t} \) is the mean value of the Gaussian in the mixture at time \( t \), \( \Sigma_{i,t} \) is the covariance matrix of the Gaussian at time \( t \) and, \( \eta \) is a Gaussian probability density function:

\[ \eta \left( X_t, \mu, \Sigma \right) = \frac{1}{(2\pi)^{d/2} |\Sigma|^{1/2}} e^{-\frac{1}{2} (X_t - \mu)^T \Sigma^{-1} (X_t - \mu)} \]  

In this way a mixture of Gaussians (MOG) characterizes the distribution of observed values of each pixel, especially when the pixel is subjected to visit several states as shown below.

Every new pixel value \( X_t \) is then used to update the model. In particular, if among the \( k \) Gaussians there is one having the new pixel value within 2.5 standard deviation of the distribution (matching model), then its parameters are updated as follows:

\[ \mu_{t} = (1 - \rho) \mu_{t-1} + \rho X_t \]  

\[ \sigma_{t}^2 = (1 - \rho) \sigma_{t-1}^2 + \rho \sigma_{t-1}^2 \]  

where

\[ \rho = \alpha \eta(X_t | \mu_k, \sigma_k) \]  

The \( \mu \) and \( \sigma \) parameters for the other distributions that do not satisfy the matching condition remain the same.

The prior weights of the \( k \) distributions at time \( t \), \( w_{k,t} \), are adjusted as follows:

\[ w_{k,t} = (1 - \alpha) w_{k,t-1} + \alpha (M_{k,i}) \]  

where \( \alpha \) is the learning rate and \( \left( M_{k,i} \right) \) is 1 for the model which match and 0 for the remaining models.

If none of the \( k \) distributions matches the value of the current pixel, the least probable distribution is replaced with a distribution having the current value as its mean value, an initially fixed high variance, and low prior weight.

In this paper the above mixture models theory has been adapted for the denoising purpose in the holographic reconstructed image.

In particular, the proposed approach consists of a preliminary phase involving the selection of a small number of pixels not belonging to any object placed within the depth of field of the holographic acquisition setup. Using the terminology of the research area identified as machine learning, these pixels can be considered the reference set \( T = \{ X_i \} \) that allow the system to learn the temporal behaviors of non-object pixels. During the acquisition phase, at the time instant \( t \), a statistical model is built and updated according to (2)–(7) for each pixel of the holographic image.

In addition, the likelihood of the set \( T \) is computed for each pixel according to the probability computation in case of compatible events as follows

\[ P_G(X_t) = \sum_{i=1}^{N} P(X_t) \sum_{k=1}^{N} \prod_{i=1}^{N} P(\sigma_i (\{ X_i \}, X_i|^k)) \]  

where \( \sigma_i (x) \) states for the all possible permutations of the elements in \( T \) taken \( i \) at a time. In this way a likelihood image is generated and an adaptive threshold can be used to generate a binary image where pixels most probably belonging to the object under consideration are set to 1 and all the remaining pixels are set to 0.

The resulting binary image is then processed in order to detected connected regions and to discard isolated pixels [5]. The outcome of the region connectivity analysis is a binary mask \( \hat{M} \). This mask is then superimposed to the original holographic image \( I \) and, in this way, a new intermediate holographic image is obtained:

\[ \hat{I}(x, y) = \begin{cases} I(x, y) & \text{if } M(x, y) = 1 \\ 0 & \text{if } M(x, y) = 0 \end{cases} \]  

The image \( \hat{I} \) is finally spatial-temporally further filtered in order to smooth the speckle effects on the region belonging to the objects under observation. The final image outcome \( \tilde{I} \) of the proposed denoising approach is obtained in this way. This filtering is done by replacing at time \( t \) the pixel value in \( \hat{I}(x, y) \) with the value of its temporal statistics \( \mu_k \) averaged with those of the 4-distance neighboring pixels. This is carried out for all nonzero pixels in \( \hat{I} \) whereas zero pixels remain unchanged. This can be formalized as follows:

\[ \tilde{I} = \begin{cases} 0, & \text{if } \hat{I}(x, y) = 0 \\ \frac{1}{25} \sum_{i=1}^{2} \sum_{j=1}^{2} \mu_{t}(x + i, y + j) & \text{if } \hat{I}(x, y) \neq 0 \end{cases} \]  

where \( \mu_{t}(x, y) \) is the mean of the temporal statistics of the \( t \)th frame of the pixel \( x, y \) and \( \Sigma_{t}(x, y) \) is the covariance matrix of the pixel \( x, y \) of the \( t \)th frame.
III. EXPERIMENTAL RESULTS

The proposed approach has been tested on a holographic image sequence of the Perseus statue [9], [10]. The holograms were captured at long wavelength infrared (IR) range (10.6 μm) in off-axis Fourier configuration [10]. A Miracle Thermoteknix 307k, 640 × 480 detector resolution and 25 μm pitch was used.

At first, analyzing over time the statistical distribution of a set of pixels in the image affected by the speckle noise has allowed the choice of a suitable Gaussian mixture model. Fig. 1 shows the distribution of a noisy pixel belonging to the foreground, where it is possible to notice a multimodal behavior of the intensity values for a time window of 100 temporally contiguous frames.

This preliminary study has demonstrated that at least two Gaussians are needed to describe the dynamics of pixel values. However, in order to correctly model also very infrequent values occurrences (that are usually associated to the speckle) a 3-Gaussian model has been used in the following experiments.

Conversely, a greater number of Gaussians has been experimentally proven to be useless since the models generated on the reference points experienced very low values of coefficients associated with the additional Gaussians.

In Fig. 2(a), a photo of the Perseus statue is shown whereas, Fig. 2(b) reports the 100th image of the recorded holographic sequence. Speckle noise is evident and strongly affects the holographic image quality.

In Fig. 3(a), the likelihood image relative to the 100th image and obtained by using a set $T$ of 20 reference pixels not belonging to the statue area is shown. In this figure, likelihood values are computed as described in (8) and then rescaled in the range $[0, 255]$ for display purposes: the lightest pixels are related to areas of the background (highest non-object likelihood values) and the darkest pixels are instead related to areas belonging to the statue (lowest non-object likelihood values). Fig. 3(b) reports the corresponding segmented binary image extracted using a threshold ($t_h = 0.5$) on the likelihood values.
In Fig. 4(a) the binary mask obtained after topological analysis of the segmented image is shown. Fig. 4(b) presents the new holographic image obtained as described in equation (9).

In Fig. 5(a) the image obtained after the spatiotemporal filtering detailed in (10) is shown.

The outcomes of the proposed approach have then been qualitatively compared with those obtained using latest state of the art spatial speckle denoising techniques. In particular Figs. 5(b)–(f) show the hologram denoised by a median filter (MF) [17], a Discrete Fourier Filter (DFF) [3], a nonlocal means filter (NLM) [14], a Frost filter [16] and a Lee Filter [15] respectively.

All in all, from Fig. 5, the superiority of the proposed denoising approach is evident. Its outcome is in fact much clearer than those obtained by using other denoising strategies: on the one hand the noise is completely suppressed in the areas outside the object and strongly attenuated on the regions of the object and, on the other hand, the details of the object appear preserved allowing a high visual quality.

For a quantitative comparison of the various filters, three different metrics have been considered:

1) Speckle Index [8], [9] on (quasi) homogenous areas of the object under consideration;
2) Contrast of the whole image (Global Contrast) [6];
3) Contrast [7] of a small region containing some details of the object (Local Contrast) [7].

Speckle index is a measure of the level of noise in the image: it is computed on a uniform region of the object in order to be sure that every observed deviation from the expected reference value in the pixels is due to noise. That said, it is obvious that the lower the index, the less noisy is the image and in particular, in an ideally uncorrupted image the speckle index value is 0.

Global Contrast indicates how much the object is enhanced with respect to the whole content of the image, whereas, Local Contrast, is a measure of how much small details in the object are preserved in the holographic reconstruction.

Both the global and the local contrast should be kept as high as possible: high values are in fact the numerical evidence that there is variability in the representation of the structures constituting the object under observation.

From the computation point of view, the Speckle index value is computed as

\[ S_t = \sqrt{\frac{\sigma}{\mu}} \]  

where \( \sigma \) and \( \mu \) are the standard deviation and the mean computed in the (quasi) uniform considered region included in the green rectangle in Fig. 6.

Contrast measurements are instead computed as:

\[ C = \frac{\sigma}{(\mu_t)^n} \]
is a positive number (more precisely $n = 2$ in the following), $\sigma$ is the standard deviation and $\alpha_4$ is the kurtosis measure defined by

$$\alpha_4 = \frac{\mu_4}{\sigma^4} \quad (13)$$

where $\mu_4$ is the fourth moment about the mean and $\sigma^2$ is the variance.

As mentioned above, the numerator and denominator in (12) are computed on the whole image for Global Contrast measure and on a small selected region for Local Contrast respectively. The region chosen for Local Contrast computation is included in the red rectangle in Fig. 6: this choice has been made considering that this region contains some details of the face that should be preserved.

A comparison between the six filters in terms of the above metrics is shown in Table I, which provides the numerical evidence that the proposed approach achieved the best results in terms of all the considered metrics.

<table>
<thead>
<tr>
<th>Filter Type</th>
<th>Speckle Index $S_i$</th>
<th>Global Contrast $C_G$</th>
<th>Local Contrast $C_L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original</td>
<td>0.85</td>
<td>1.04</td>
<td>0.75</td>
</tr>
<tr>
<td>Median Filtering [17]</td>
<td>0.69</td>
<td>0.93</td>
<td>0.49</td>
</tr>
<tr>
<td>Discrete Fourier Filtering [3]</td>
<td>0.82</td>
<td>1.23</td>
<td>0.43</td>
</tr>
<tr>
<td>NonLocal Mean [14]</td>
<td>0.48</td>
<td>0.84</td>
<td>0.33</td>
</tr>
<tr>
<td>Frost Filter [16]</td>
<td>0.54</td>
<td>0.87</td>
<td>0.39</td>
</tr>
<tr>
<td>Lee Filter [15]</td>
<td>0.58</td>
<td>0.89</td>
<td>0.46</td>
</tr>
<tr>
<td>Intermediate Image $i$ (eq. 9 in section II)</td>
<td>0.85</td>
<td>1.39</td>
<td>0.83</td>
</tr>
<tr>
<td>Proposed Approach</td>
<td>0.45</td>
<td>1.26</td>
<td>0.49</td>
</tr>
</tbody>
</table>

TABLE I
VALUES OF QUALITY METRICS

where $n$ is a positive number (more precisely $n = 2$ in the following), $\sigma$ is the standard deviation and $\alpha_4$ is the kurtosis measure defined by

$$\alpha_4 = \frac{\mu_4}{\sigma^4} \quad (13)$$

where $\mu_4$ is the fourth moment about the mean and $\sigma^2$ is the variance.

As mentioned above, the numerator and denominator in (12) are computed on the whole image for Global Contrast measure and on a small selected region for Local Contrast respectively. The region chosen for Local Contrast computation is included in the red rectangle in Fig. 6: this choice has been made considering that this region contains some details of the face that should be preserved.

A comparison between the six filters in terms of the above metrics is shown in Table I, which provides the numerical evidence that the proposed approach achieved the best results in terms of all the considered metrics.

Going into detail, the analysis of the table shows that the proposed approach is able, on the one hand, to best reduce the noise (lowest speckle index), but at the same time it also enhances the relevance of the object with respect to the background (highest Global contrast) and finally it preserves the structural details on the object (highest Local Contrast).

In particular, it is possible to argue that this encouraging result is due to the combined actions of two operational steps involved and described in Section II: on the one hand the initial temporal segmentation process increases the global image contrast (second last row), whereas, on the other hand, the following spatiotemporal filtering reduces speckle on the object regions without heavily affecting object structures (last row). This favorable effect could be exploited in contexts where either higher-contrast or lower speckle index is required.

This section concludes with some details about the processing time of the proposed technique and a quantitative comparison with conventional techniques that only use spatial information.

All techniques have been implemented in Matlab®, version R2012a, running on a commercial Laptop equipped with an Intel Pentium P6200 (with a clock frequency of 2.13 GHz) and 4 GB of RAM. Regarding the proposed technique, the temporal analysis, defined by (8)–(9), takes, for the computation of the likelihood in each pixel, on average of $\sim 3.8 \times 10^{-3}$ sec. The spatial analysis, defined by (10), takes instead, for the computation of the new value in each pixel of the holographic image, an average of $\sim 1.4 \times 10^{-4}$ sec. The final estimate of the time required for the overall processing of each pixel in the image is in the order of $\sim 10^{-5}$ sec.

This result should be placed in relation with the average time per pixel observed for the application of conventional techniques. In particular the median filtering [17] takes on average of $\sim 3.9 \times 10^{-4}$ s for each pixel, the Discrete Fourier Filtering [3] $\sim 9.44 \times 10^{-7}$ s, the NonLocal Mean filtering $\sim 1.1 \times 10^{-3}$ s, the Frost filter $\sim 1.12 \times 10^{-4}$ and finally the Lee filter $\sim 1.49 \times 10^{-6}$. Although these data demonstrate the comparability of the computational time of the proposed technique with most of the conventional ones, it should be emphasized that the code used for the proposed technique could be further computationally enhanced (for example, using built-in Matlab functions for multidimensional data handling).

Furthermore the computation required for the proposed technique is largely parallelizable (in particular (8) and (9) are embarrassing parallel since they may be parallel performed for each pixel).

IV. CONCLUSION

In this paper a new and innovative technique to reduce the noise in a reconstructed hologram image has been proposed.

In order to overcome the limitation of the techniques in the literature, both spatial information and temporal statistics associated with individual pixels are used to detect and suppress the speckle noise as well as to enhance the information content.

Experimental results on an holographic image sequence captured at long wavelength infrared demonstrated that the proposed approach gives better results in both visual quality of the cleaned images and quantitative analysis of numerical outputs in terms of three different metrics.

The analysis and comparison of processing time has also been conducted demonstrating the comparability with the state-of-the-art techniques.

Future works will be addressed to further improve reconstructed hologram image by introducing iterative region growing strategies that will allow overcoming segmentation errors that occurred at the boundaries of the object regions.

Moreover a refinement of the analysis of the pixel statistics will be experimented in order to better preserve details on the reconstructed object.

REFERENCES


Melania Paturzo is research scientist at INO-CNR. Her research interest is in the field of Optics. Her research activities concern the investigation of material properties by means of interferometric techniques, fabrication of periodically poled lithium niobate samples by electric field poling process, development of optical techniques to get super-resolution in digital holographic microscopy, holographic display, quantitative phase-contrast microscopy for cells investigation, optofluidics. She is author and/or coauthor of more than 40 papers in international peer reviewed journals and of more than 50 conference papers.

Pasquale Memmolo was born in Avellino, Italy, in 1982. In 2005, he received the first level degree in computer engineering. In 2008, he received the second level degree in communication engineering, with final mark 110/110, at the “Università degli Studi di Napoli Federico II” (Napoli, Italy), working on numerical and optical multiplexing and de-multiplexing methods in digital holography. In 2011, he completed the Ph.D. in electronic and communication engineering, with a thesis entitled “Compressed Sensing: a new framework for signals recovery and its application in Digital Holography”. Now he works as a Post Doctoral Junior at the Center for Advanced Biomaterials for Health Care@CRIB, Istituto Italiano di Tecnologia, Napoli, Italy, and he is a collaborator at the CNR—Istituto Nazionale di Ottica, Pozzuoli, Italy. He is also co-author of more than 30 international journal’s publications and proceedings papers.

Massimiliano Locatelli graduated in physics at Florence University, in 2009, with a Thesis on ‘Digital Holography Applications’, developed at INO-CNR. He got a Ph.D. at LENS (European Laboratory for Non-linear Spectroscopy) University of Florence, in 2013, with a thesis on ‘Mid Infrared digital holography and Terahertz imaging’. He is actually a researcher at LENS and INO-CNR of Florence. His research activity is conducted in the field of Optics (in particular Infrared and Terahertz imaging and digital holography), Biophysics and Complex System Physics.

Eugenio Pugliese received the Physics degree from the University of Florence, Italy, with a thesis on Mossbauer Spectroscopy, and is currently a Ph.D. student in nonlinear dynamics and complex systems at the same university. He carries out its research activities in the laboratories of INO-CNR in Florence. His research interests are in the field of chaotic dynamics of classical and quantum systems, digital holography both in visible and IR range and optical investigation of complex systems.

Riccardo Meucci graduated in physics at the University of Florence in 1982 and received the Ph.D. degree in optics at the same University in 1987. From 1984 to 1987, he served as Researcher in Optics at Istituto di Cibernetica of the Italian National Research Council (CNR) Arco Felice (Naples), thereafter at Istituto Nazionale di Ottica (INO). Actually he is Research Director at INO-CNR. His scientific activity mainly includes: chaotic instabilities in single mode lasers, laser transients, dynamical models for the CO₂ laser, control of chaos, spatio-temporal instabilities and delayed systems, polarization instabilities and digital holography in the mid-infrared region.

Pietro Ferraro (M’02–SM’07) is currently Chief Research Scientist at Istituto Nazionale di Ottica del CNR, Pozzuoli, NA, Italy. Previously he worked as Principal Investigator with Alenia Aeronautics. He has published 12 book chapters, 170 papers in journals, more than 150 papers at International Conferences. He edited two books with Springer. He holds 14 patents. Among his current scientific interests are: holography, interferometry, microscopy, fabrication of nanostructures, ferroelectric crystals, optical fiber sensors, fiber bragg gratings, nano-microfluidics, optofluidics. Dr. Ferraro has chaired many International Conferences He is in the Editorial Board of Optics and Lasers in Engineering (Elsevier), is Topical Editor of Optics Letters. Dr. Ferraro is Fellow of SPIE and Fellow of OSA.