

Multifocus image fusion through pseudo-Wigner distribution

Salvador Gabarda

Gabriel Cristóbal

Instituto de Óptica "Daza de Valdés"
(CSIC)

Serrano 121, Madrid-28006
Spain

Abstract. We describe a new image fusion paradigm that provides an enhanced image from a set of source images that present regions with different spatial degradation patterns. The fusion procedure is based on the use of a new defocusing pixel-level measure. Such a measure is defined through a 1-D pseudo-Wigner distribution function (PWD) applied to nonoverlapping N -pixel window slices of the original image. The process is repeated to cover the full image size. By taking a low-resolution image as a reference image, which can be defined by blurring and averaging the two source images, a pixel-level distance measure of the defocus degree can be obtained from the PWD of each image. This procedure makes it possible to choose from a focusing point of view the in-focus pixels from each one of the given source images. The method is illustrated with different pairs of images of the same scene, which are partly focused and partly defocused in different regions. The image fusion approach that we propose here can work for any source of images available, and the comparison using evaluation measures such as mean square error or percentage of correct decisions shows that our framework can outperform the current approaches for the analyzed cases. One additional advantage of the present approach is its reduced computational cost when compared with other methods based on a full 2-D implementation of the PWD. © 2005 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.1881412]

Subject terms: Wigner distribution; image fusion; multifocus; image enhancement.

Paper 040043 received Jan. 23, 2004; revised manuscript received Jul. 27, 2004; accepted for publication Oct. 25, 2004; published online Mar. 30, 2005.

1 Introduction

Image blurring can be regarded as a consequence of some imperfect image formation process, which limits the image bandwidth. Hence, blurring can be produced by different causes, i.e., camera or object motion, defocused optical systems, and atmospheric turbulence. Sometimes these degradations can be modeled to recover the original image, but in many cases this information is not available. The recovering of the original image from a degraded input has been the subject of many papers in the area of image restoration that can be indexed with the keywords of image deblurring or image deconvolution.¹ When image restoration is accomplished without any *a priori* knowledge about the degradation cause, we are dealing with blind image deconvolution methods.² If the blurring is not homogeneously distributed, the defocusing process will affect different regions of the image with different strength. This scenario is referred to as space-variant blurring.³ A special case of space-variant deblurring occurs when more than one degraded image is available and therefore image fusion methods^{4,5} can be applied. We present a fusion algorithm based on the use of the pseudo-Wigner distribution, which is suitable for blind space-variant deblurring. One additional advantage of the current method, which is based in a

1-D PWD sliding window implementation, is its reduced computational cost compared with other methods based on full 2-D PWD implementation.

This work is structured as follows. In Sec. 2, the discrete PWD is briefly outlined and its usefulness for spatial filtering is described. In Sec. 3 a pixel-based distance scheme between PWD vectors is described as the basis of the fusion procedure. Image fusion results, using the PWD, are given in Sec. 4, and a quantitative fusion quality assessment is presented that considers some ground-truth test images. Finally, conclusions are drawn in Sec. 5.

2 Wigner Distribution

Blurring can be described as the convolution of the original image with a 2-D point-spread function (PSF). If such a function is not a function of the spatial location under consideration, it is said to be spatially invariant (essentially this means that the image is blurred exactly in the same way at every spatial location). Therefore, if the PSF is a known function, it is possible to remove the blur by means of an inverse operation in the spectral domain called deconvolution. Importantly, modeling and restoring images degraded by spatially varying blurs are more challenging problems that still require further attention to achieve a successful solution. The use of a spatial/spatial-frequency representation is greatly appropriated for addressing a spatially varying blurring problem because of the local character of the degradation process. Our fusion method is based on the use

of many input realizations of the same object at the same time. Deblurring is obtained by comparing the PWD of the degraded image with the PWD of a worst-case prototype. Hence, we approach the spatially varying blur case by means of one of the conjoint spatial/spatial-frequency representations. A general formulation of the spatial/spatial-frequency representations was given by Cohen^{6,7} and later generalized to N -D signals by Jacobson and Wechsler.⁸ It has been claimed that the Wigner distribution function (WD) has the best properties to be used in image processing, against other representations of this kind. In fact, some researchers have considered the WD as a “masterform” distribution from which the rest of the distributions can be obtained.^{8,9} Thus it has the best resolution, which is matched to that of the image in both domains. It overcomes the resolution tradeoffs that traditionally have limited the utility of windowed power spectrum analysis. Besides that, the WD is a joint bilinear representation, very close to positive, invariant within linear transformations, and it contains all the image information. It presents two main advantages with respect to other local representations. First, the WD is a real-valued function and encodes directly the Fourier phase information. Second, the election of the appropriate window size, which depends on the kind of analyzed information, is not required for the computation of the WD. However, the use of other Cohen’s class distributions appears to be feasible following this approach, but their performance needs to be validated in future experiments.

The Wigner distribution (WD) is a bilinear (quadratic) signal representation introduced by Wigner.¹⁰ A comprehensive discussion of the WD properties can be found in a series of classical articles by Claasen and Mecklenbräuker.¹¹ Originally, the WD was applied to continuous variables as follows. Let us consider an arbitrary 1-D function $z(x)$. The WD of $z(x)$ is given by

$$W(x, u) = \int_{-\infty}^{\infty} z\left(x + \frac{\alpha}{2}\right) z^*\left(x - \frac{\alpha}{2}\right) \exp[-i(u \cdot \alpha)] d\alpha, \quad (1)$$

where $*$ denotes complex conjugation.

By considering the shifting parameter α as a variable, Eq. (1) represents the Fourier transform (FT) of the product $z(x + \alpha/2)z^*(x - \alpha/2)$, where u denotes the spatial-frequency variable and, hence, the WD can be interpreted as the local spectrum of the signal $z(x)$. The Wigner distribution satisfies many desirable mathematical properties.¹¹ Although the PWD was initially defined for continuous variable functions, Claasen and Mecklenbrauker proposed at the beginning of the 1980s a first definition for discrete variable functions. However, some attempts to extend definitions of the WD to discrete signals have not been completely successful.¹² For this application, we have selected the following discrete Wigner distribution, similar to the one proposed by Claasen and Mecklenbräuker and also by Brenner¹³

$$W(n, m) = 2 \sum_{k=-N/2}^{N/2-1} z(n+k) z^*(n-k) \exp\left[-2i\left(\frac{2\pi m}{N}\right)k\right]. \quad (2)$$

This equation is limited to a spatial interval $[-N/2, N/2 - 1]$, and it has to be considered as a pseudo-Wigner distribution (PWD). In Eq. (2), n and m represent the spatial and space-frequency discrete variables, respectively, and k is the shifting parameter, which is also discrete. Equation (2) can be interpreted as the discrete Fourier transform (DFT) of product $r(n, k) = z(n+k)z^*(n-k)$. It is worth noting that the discretizing process implies a loss of some properties of the continuous WD. One important property preserved within the definition given by Eq. (2) is the inversion property, which is a very desirable feature for the recovering of the original signal, and which allows local filtering operations on the images under consideration.

According to the inversion property,¹¹ the even samples can be recovered from Eq. (2) from the following expression¹⁴

$$z(2n)z^*(0) = \sum_{m=-N/2}^{N/2-1} W(n, m) \exp\left[-2i\left(\frac{2\pi m}{N}\right)n\right], \quad (3)$$

and the odd samples can be recovered from

$$z(2n-1)z^*(1) = \sum_{m=-N/2}^{N/2-1} W(n, m) \exp\left[-2i\left(\frac{2\pi m}{N}\right)n\right]. \quad (4)$$

To get the original values, we have to perform an inverse DFT for recovering of the function $r(n, k)$, which gathers all the information included in the original discrete signal. Equation (3) is obtained by writing $n=k$ in $r(n, k)$, which implies that the values $z(2n)z^*(0)$ are on the main diagonal of the matrix $r(n, k)$. Similarly, Eq. (4) results from taking $k=n-1$ in the product function $r(n, k)$, thus the odd sample values are located above the main diagonal. To recover the exact values of the samples, we have to divide the diagonal values by $z^*(0)$ and $z^*(1)$, respectively.¹⁴ Unless the sign of the samples are undetermined due to the product sign rule, they can always be considered positive, because we are dealing with digital images of real positive gray-value levels. From Eqs. (3) and (4), it can be shown that n varies in the interval $[-N/4, N/4-1]$, due to the factor 2, which affects the left side of both equations. The 2-D PWD can be considered as a generalization of the 1-D case, but to reduce the computational cost to a minimum, the 1-D PWD will be used as defined before, by taking $N=8$, i.e., a 1-D window of 8 pixels.

3 Pixel-Based Distance Between PWD Vectors

The 1-D, N -pixels windowed PWD, as described earlier, can be used to process the data of images, with the only requirement being that the gray levels must be defined by real and positive values. The method proposed herein involves scanning the original image by nonoverlapping sliding window blocks of N pixels in a row and assigning to each pixel of the row its N -component PWD vector. In such a way, we are gathering both spatial and spatial-frequency information of every pixel of the image. If we have multiple input information, i.e., two images partly defocused in complementary regions, it is possible, after the theoretical

considerations that follow, to propose a method for fusing both images and to obtain an enhanced one, which can be considered as the unknown original.

Our method is based on the idea that every pixel on the image has a different N -component PWD vector associated with it; thus, a measure can be taken for determining their greatest distance to a reference image. This approach is closely related with shape matching.¹⁵ The selection of a similarity measure is required to take an accurate decision, since dissimilarity can be coupled with the notion of distance. Because we are dealing with a distance metric, it is required to set a reference value. It is well known that the more defocused an image is, the more the high frequencies will be diminished and consequently its PWD will be affected. Therefore, by taking as a reference a blurred average from the two originals given by $z_r = (z_1 * h + z_2 * h) / 2$, z_1 and z_2 represent the two input images and h represents a box filter whose values are detailed later. The main idea behind this reference is to have a worst case that allows the establishing of a distance measurement at pixel level from the PWD of each image,

$$d_i(\zeta) = \|W_i(n, m) - W_r(n, m)\|_{n=\zeta}, \quad (5)$$

where ζ represents an arbitrary pixel of the image, subindices i and r stand for input and reference images, respectively, and n, m are defined as in Eq. (2). $W_i(n, m)$ indicates the PWD for image i , and $W_r(n, m)$ is the N -component PWD for the reference image. The normal operator in Eq. (5) is defined as usual,

$$\|\xi\| = \left[\sum_{i=-N/2}^{N/2-1} \xi^2(i) \right]^{1/2}, \quad (6)$$

where ξ represents any arbitrary real function. The selected distance [Eq. (5)] is an Euclidean distance, which belongs to the more general family of Minkowsky distances.¹⁵

The rationale of using Eq. (5) can be formulated as follows. Let us consider an original image z and a blurring process given by $g = z * h_0$, where $*$ denotes the convolution operation. After applying the PWD to $z(n)$ and $g(n)$ separately, their respective representations W_z and W_g are obtained. From a filtering point of view, it can be helpful to consider that the PWDs of the functions $z(n)$ and $g(n)$ are related by the product

$$W_g(n, m) = W_z(n, m)H_1(n, m), \quad (7)$$

where $H_1(n, m)$ can be interpreted as a pseudo-filtering function, since g might not exist.^{14,16} For a thorough discussion see Ref. 17. If the Fourier transformation is represented by $\mathcal{F}[\cdot]$, Eq. (7) can be rewritten as

$$\mathcal{F}[r_z(n, k) * h_1(n, k)] = \mathcal{F}[r_z(n, k)] \cdot \mathcal{F}[h_1(n, k)], \quad (8)$$

where the convolution affects only variable k , and the relationship $r_g(n, k) = r_z(n, k) * h_1(n, k)$ between their respective product functions holds, after the convolution property of the Fourier transform, where $z(n)$, $g(n)$, $r_z(n, k)$, and $r_g(n, k)$ all represent real non-negative functions. On the other hand, according with the energy conservation prin-

ciple $h_0(n)$, $h_1(n, k)$ will not present negative coefficients. Consequently, the following conditions are satisfied

$$\sum_{n=-N/2}^{N/2-1} h_0(n) = 1 \quad \text{and} \quad \forall n / h_0(n) \geq 0 \Rightarrow \|h_0(n)\| \leq 1, \quad (9)$$

$$\sum_{k=-N/2}^{N/2-1} h_1(n, k) = 1 \quad \text{and} \quad \forall n, k / h_1(n, k) \geq 0 \Rightarrow \|h_1(n, k)\|_{n=\zeta} \leq 1. \quad (10)$$

Then, by applying the properties of the Euclidean distance to Eq. (8), the following expression is achieved

$$\|\mathcal{F}[r_z(n, k) * h_1(n, k)]\|_{n=\zeta} \leq \|\mathcal{F}[r_z(n, k)]\|_{n=\zeta}. \quad (11)$$

If both images to be fused have the same energy, Parseval's theorem tells us that equality holds in Eq. (11). Nevertheless, differences between PWDs belonging to different degraded versions of the same image will not necessarily nullify. Consequently, these differences can be used for defining an Euclidean distance. Thus, by means of Eq. (5) and by taking the blurred averaged image as a reference, distances d_1 and d_2 can be obtained for the same pixel ζ from input images $z_1(n)$ and $z_2(n)$. Theoretically, the largest Euclidean distance corresponds to the pixel belonging to the less locally defocused image. Then, by comparing distances d_1 and d_2 , an activity map can be obtained [see Figs. 1(d) and 1(e)] that provides an image mask of regions of high resolution (in-focus) and low resolution (out of focus). Therefore, the pixels from the two input images can be assigned to two different classes: one class corresponding to the set of pixels with a better in-focus measure, and the other class for the set of pixels with the worst in-focus measure. Thus, fusion of images z_1 and z_2 can be attained, and therefore an all-in-focus enhanced version of the original image is produced. The suitability of the method is based on two features. The first one is the advantage of the local frequency information provided by the PWD, and the second one is the use of a metric measurement provided by the Euclidean distance applied to the PWDs of the input signals. The combination of these two features leads us to a way of selecting the best choice for the value of a pixel in a local context of the image.

4 Experimental Results

Several examples are presented for quantitative evaluation of the method, based on the use of artificially generated space-variant registered input images. This method uses a pixel-based neighborhood and requires prior registration of the images to be considered. Otherwise, the resulting fusion procedure introduces a blurring effect that is more noticeable for severe misregistration. The scheme described here is not a deconvolution method, and its main feature is for detecting the best focused pixels among a set of occurrences provided by different input images. If a sharp view of a given area is not present in any of the inputs, a blurred spot will result in the output, as the method cannot provide better results than the least defocused image in a local context. Summarizing, there are no constraints in the blurring affecting the input images to the process except the energy

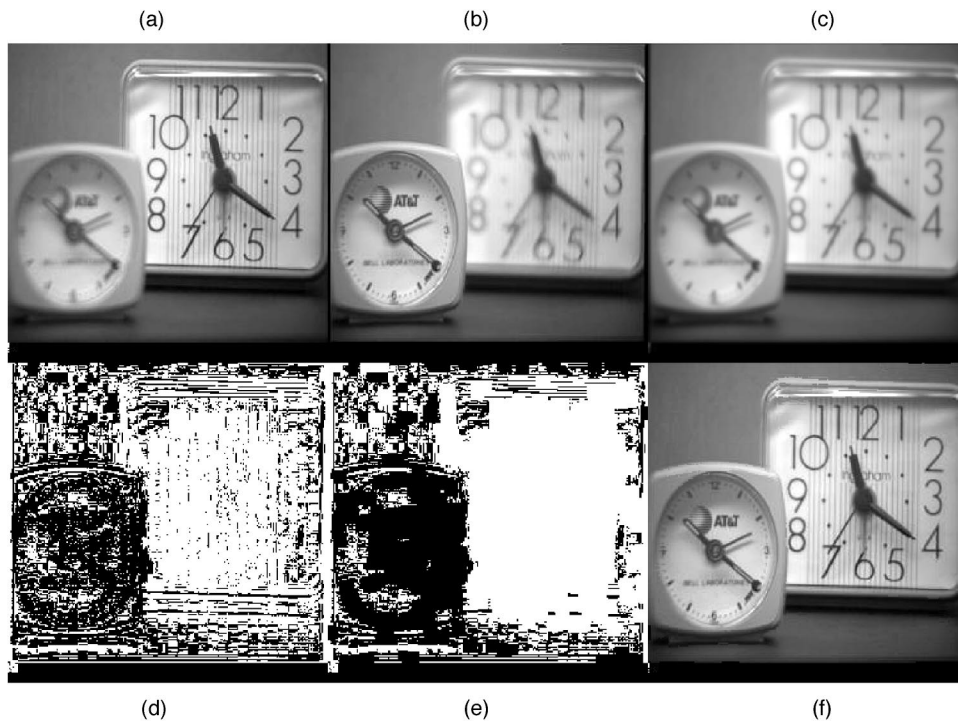


Fig. 1 (a) and (b): Two complementary space-variant defocused images. (c) Mixed reference as described in the text. (d) Binary activity map from PWD distance measures at pixel level. (e) Smoothing step for eliminating isolated pixels. (f) Resulting enhanced image.

limitations described in Sec. 3. Nevertheless, accurate registration of the input images and local blur-free information in at least one of the input images are requirements to be fulfilled for a correct performance of the method. The worst-case reference image can be obtained by any suitable energy preserving convolution kernel but is spatially limited to the window size used for computing the PWD.

Figures 1(a) and 1(b) show two images that present in-focus regions located in different (and complementary) parts of the image. As mentioned before, they can be processed by comparing them with the mixed reference image shown in Fig. 1(c), which was obtained as previously described using a blurring convolution kernel $h_0 = (1/16) \times [1\ 4\ 6\ 4\ 1]$, applied by rows and then by columns. The fusion procedure can be summarized as follows. First, each image row is divided into 8 pixel arrays, and then a 1-D PWD of each segment of 8 pixels is computed according to Eq. (2). The small value of such an analysis mask ($N=8$) permits higher computational savings and a localized spectral analysis. As result of such computation, an 8×8 matrix of data is obtained, where each column provides the PWD of each pixel. This procedure is repeated for every image row to cover the full image. Then, the Euclidean distance to the mixed reference image is obtained pixel-by-pixel using Eq. (5). The highest distance according to such criteria allows extracting the in-focus pixels of the referred images. Since the input images [Figs. 1(a) and 1(b)] can be considered complementary versions of an unknown ground-truth image, the activity map shown in Fig. 1(d) allows computing of the fused image [Fig. 1(f)]. This activity map is a binary image whose elements are related one-to-one with the pixels from the images. It has a value of 1 when

$d_1(\zeta) > d_2(\zeta)$, and a value of 0 otherwise. During this process, a smoothing binary filter is used to eliminate isolated pixels from the binary activity map. Despite that different criteria can be used for this purpose, i.e., median filters, morphological erosion, etc., a new algorithm is used as described in Appendix A in Sec. 6. Figure 1(e) shows the result of applying such an algorithm to Fig. 1(d) for removing isolated errors. Figures 2 and 3 illustrate, with synthetic blurred versions of the Cameraman and Lena images, the performance of the method. Results and comments are similar to those given for Fig. 1. From a qualitative point of view, the results shown in Figs. 1, 2, and 3 are satisfactory. Nevertheless, to provide a quantitative assessment, some objective quality measures have been used.⁵ With this goal we tested several ground-truth images of 256×256 pixels with 8-bit resolution. Two complementary degraded images were generated by blurring, both horizontally and vertically, in predefined regions of each original with the function $h_0 = (1/16)[1\ 4\ 6\ 4\ 1]$, but similar results were achieved using other energy preserving kernels. Thus, each fused image can be compared with its original ground-truth image. As quality indicators, we used the mean squared error (MSE), the peak signal-to-noise ratio (PSNR), and the percentage of correct decisions (P_c) as defined in Appendix B in Sec. 7. The results from processing the well-known Cameraman and Lena images are summarized in Table 1. This new scheme appears to provide equal or better performance than the best existing methods in the literature⁵ when comparing similarly generated inputs. On the other hand, the computational complexity of the scheme proposed is bounded by the FFT, which is $O(N \log_2 N)$,

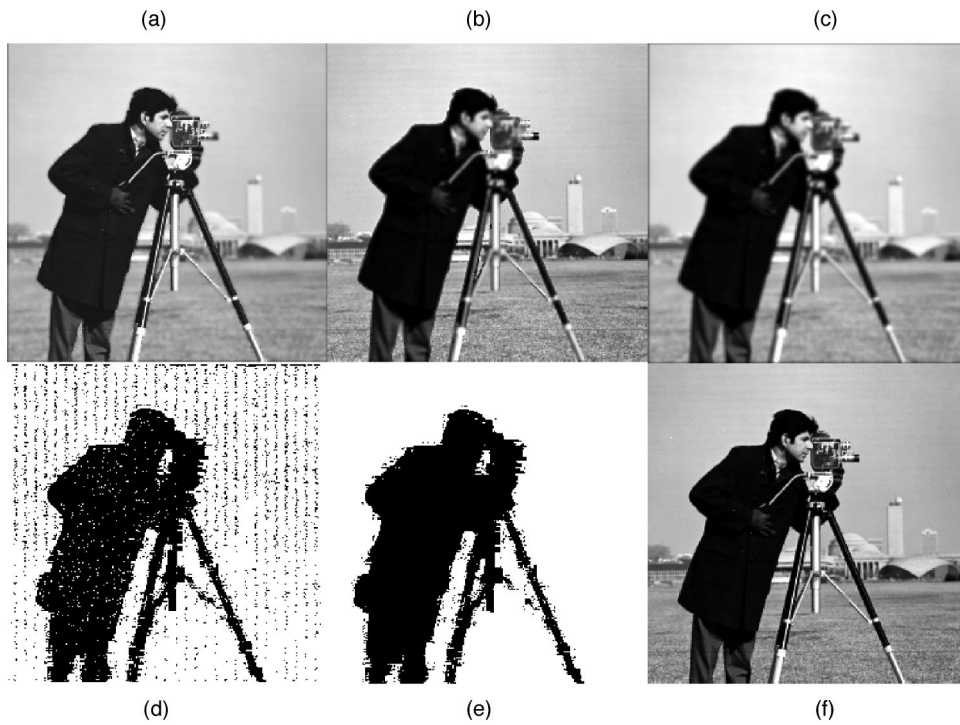


Fig. 2 (a) and (b): Two complementary space-variant defocused images. (c) Mixed reference as described in the text. (d) Binary activity map from PWD distance measures at pixel level. (e) Smoothing step for eliminating isolated pixels. (f) Resulting enhanced image.

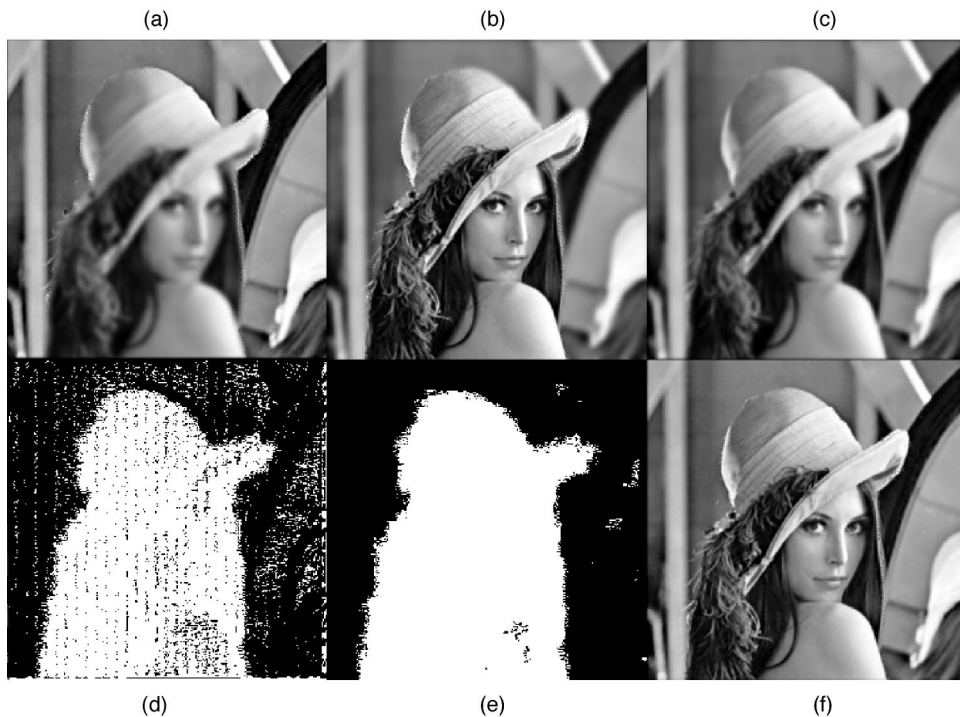


Fig. 3 (a) and (b): Two complementary space-variant defocused images. (c) Mixed reference as described in the text. (d) Binary activity map from PWD distance measures at pixel level. (e) Smoothing step for eliminating isolated pixels. (f) Resulting enhanced image.

Table 1 Quality measurements.

| Cameraman | | | |
|-----------|-----------------------------------|-----------------------------------|--------------------------------|
| Measure | Blurred version 1 versus original | Blurred version 2 versus original | Enhanced image versus original |
| PSNR | 28.86dB | 27.36dB | 46.92dB |
| RMSE | 9.20 | 10.93 | 1.15 |
| P_c | 58.25 | 58.47 | 98.55 |
| Lena | | | |
| Measure | Blurred version1 versus original | Blurred version2 versus original | Enhanced image versus original |
| PSNR | 36.32dB | 35.79dB | 49.75dB |
| RMSE | 3.89 | 4.14 | 0.83 |
| P_c | 63.32 | 58.01 | 97.33 |

where N is the number of pixels involved in the operation. The PWD requires a FFT for each pixel of the image, therefore the computational complexity can be estimated as $O(MN \log_2 N)$, where M is the number of pixels in the image and N is the window size used for implementing the PWD. The benefits of using a 1-D over a 2-D PWD processing scheme can be better illustrated by comparing window sizes. Let us consider N the length of the window in the 1-D case, and a $K \times K$ window for the 2-D case. The comparison factor between both schemes can be determined as $O(N \log_2 N / K^2 \log_2 K^2)$, i.e., if we take $N=K$, the complexity of a 1-D compared with a 2-D scheme is reduced by a factor of $2N$. As this method is based on selecting a pixel out of a set of input values, only a noisy free scenario can be considered here. That means that the noise-free regions can be obtained at least from one of the input

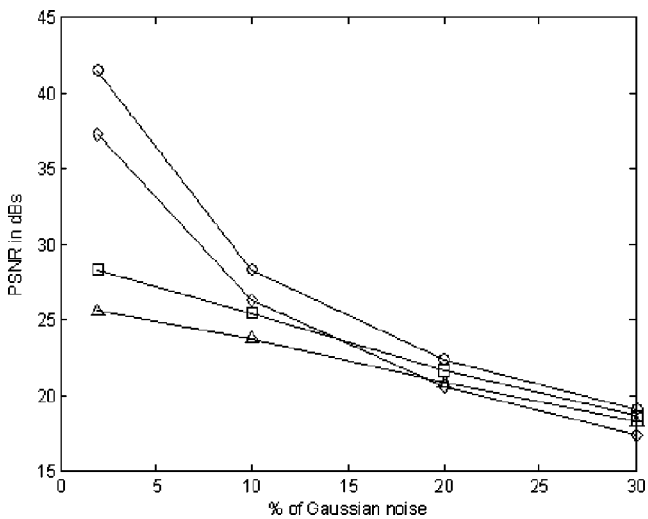


Fig. 4 PSNR values for different degraded realizations by adding Gaussian noise versus original image. Circles: noisy copy of the original image. Triangles: noisy version A. Squares: noisy version B. Diamonds: fused image from corresponding noisy A and B versions.

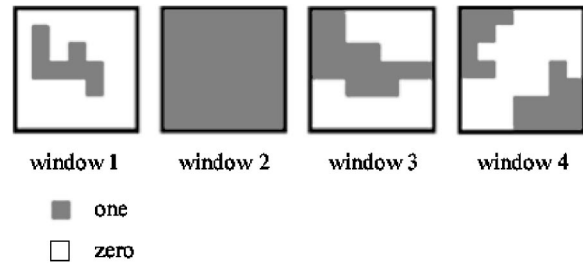


Fig. 5 Examples of different binarized pixels inside an analysis window.

images to be fused. Hence, this method is not able to perform denoising, and implies that if all input images are noisy, the output will be noisy as well. The Lena image was taken as an example for testing the noise influence on the image fusion process. This was accomplished by adding to the input images different amounts of Gaussian noise. Figure 4 shows a plot of the results for different levels of noise degradation. The two different degraded versions A and B of the original image are represented by squares and triangles, respectively. Diamonds represent the fused image and the circles represent a noisy original with the same amount of Gaussian noise as their respective A and B inputs. From such graphics, it can be observed that the fused image approximates the noisy copy of the original, but it appears noisier. As more Gaussian noise is added, the differences between input and output images are less relevant. For a small amount of additive noise, there is a noticeable gain in the signal-to-noise ratio of the fused image, but for higher levels of noise the performance of the method degrades fast.

5 Conclusions

A new fusion method based on the use of PWD is presented and applied to spatial-variant defocused images. Quality measures show that the results provide enhanced versions of the input images with equal or better performance than the best existing methods. This scheme can also be extended when more than two images are available. Another advantage of the present scheme is its reduced computational cost and simultaneous performance of a localized spectral analysis. The use of a short window for PWD analysis greatly decreases computational time. In addition, a new noise cleaning edge-preserving method is presented as an auxiliary tool for removing isolated error pixels from the activity maps. Despite the fact that the Lena and Cameraman images were artificially blurred for quantitative assessment, this method can be easily extended for dealing with arbitrary images with spatial-variant blur.

6 Appendix A

In a binary segmentation process, small isolated areas or single pixels are generally considered as noise and have to be eliminated. Therefore, the binary-segmented image must be postprocessed by a denoising algorithm. This process can be achieved in different ways.¹⁵ Here we have developed a method that can be classified among the fuzzy digital topology methods.^{18,19} By defining a characteristic function (probability) for the degree of connectedness (DOC),

isolated points can be removed from the object while selecting their size and keeping the edges of the remaining areas (object) unaffected.

When exploring a binary image with a window of $w \times w$ pixels, different configurations of 1's and 0's are obtained (Fig. 5). The window boundary divides the image into two sets: inside and outside window pixels. The inside of the window is connected with the outside on the topology of the 1's and 0's around the boundary of the window. To keep a consistency, we require the central pixel of the window to be a 1. A measure of the number of pixels close to the boundary is taken, and a DOC value is assigned to the central pixel of the window as a probability value $P \in [0,1]$. Then, the suppression of the 1's inside the window is decided by a threshold value. Moreover, a small DOC value provides a strong argument to delete the inside of the window. This uncertainty in connectedness, which is not due to randomness but to the ambiguity about the situation of 1's and 0's on the image, provides a solution to this problem from a fuzzy image processing approach.

As stated earlier, the criterion in which the algorithm is based refers to the connectivity of the pixels inside and outside the window through a probability measure. For this analysis, the inner pixels (I) and the outer pixels (E) next to the window boundary are taken. Next, 0's and 1's are compared in both sets to evaluate the probability P of region I being connected to region E . The fuzzy rule imposed to delete all 1's inside a window is as follows: "Given a binary image and a window centered in a pixel, whose value is 1, if the degree of connectedness P between inside (I) and outside (E) window regions is smaller than a certain value P_0 , then the 1's inside such a window must be changed to 0's." A given region of the image limited by a square window must be considered for estimating the probability P required to make a decision (Fig. 5). The inbound pixels of the window are connected with the outbound pixels under a 4 neighborhood connectivity criterion.¹⁵ Accordingly, two pixels are considered neighbor pixels if they are adjacent in the same line (row or column). Thus, the inner region of the window is connected with the outer region if there are at least two neighboring pixels of value 1, and, one of the pixels is inbound while the other one is outbound the window.

Let us consider the assumption that Q is the total number of pixels for set E , with q 1's and $Q-q$ 0's. Similarly, the set I has $Q-4$ pixels, of which p are 1's and the rest are 0's. After straightforward combinatory calculations, the probability of regions I and E to be connected is given by

$$P = 1 - \frac{\sum_{k=0}^{\min(4,p)} \binom{Q-8}{p-k} \binom{4}{k} \binom{Q-p-k}{q}}{\binom{Q}{q} \binom{Q-4}{p}}, \quad (12)$$

where

$$\binom{Q-8}{p-k} = \begin{cases} \binom{Q-8}{p-k} & \text{if } Q-8 \geq p-k \\ 0 & \text{otherwise} \end{cases},$$

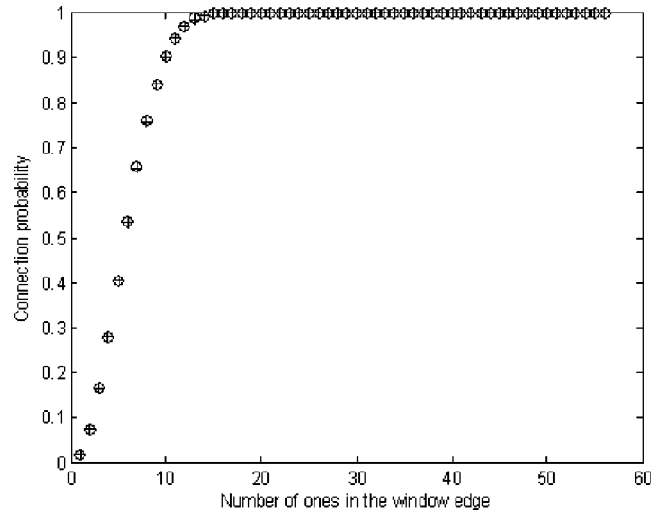


Fig. 6 Comparative results for exact and approximate connectivity in a window of 15×15 pixels. (+) Plot from Eq. (12) with $p = p_0$ and $q = p_0 + 4p_0 / (Q-4)$. (o) Plot from Eq. (14) with $p = q = p_0 + 4p_0 / (Q-4)$.

$$\binom{Q-p-k}{q} = \begin{cases} \binom{Q-p-k}{q} & \text{if } Q-p-k \geq q \\ 0 & \text{otherwise} \end{cases}.$$

Equation (12) entails a high computational cost, hence we have derived a simpler approximated expression by considering that areas I and E have the same number of pixels Q . Such approximation is based on the assumption that corner pixels in the inner region have a double connectivity valence. Moreover, we can assign the same number of 1's (p_0) to E and I , since they are neighboring areas with the same number of pixels, and therefore their expected values are the same. Since the probability α of finding a 1 filling a corner is the same of filling any other location, it follows that $\alpha = p_0 / (Q-4)$. Therefore, the probability to have k corners filled by 1's in region I is

$$P(k) = \binom{4}{k} \alpha^k (1-\alpha)^{4-k} \quad 0 \leq k \leq 4. \quad (13)$$

The expectation value of 1's in the corners is $P_{\text{cor}} = 4\alpha$, and the effective number of 1's to be considered in each region is given by $p = p_0 + 4p_0 / (Q-4)$. This yields the following probability

$$P = 1 - \frac{\binom{Q-p}{p}}{\binom{Q}{p}}, \quad (14)$$

where

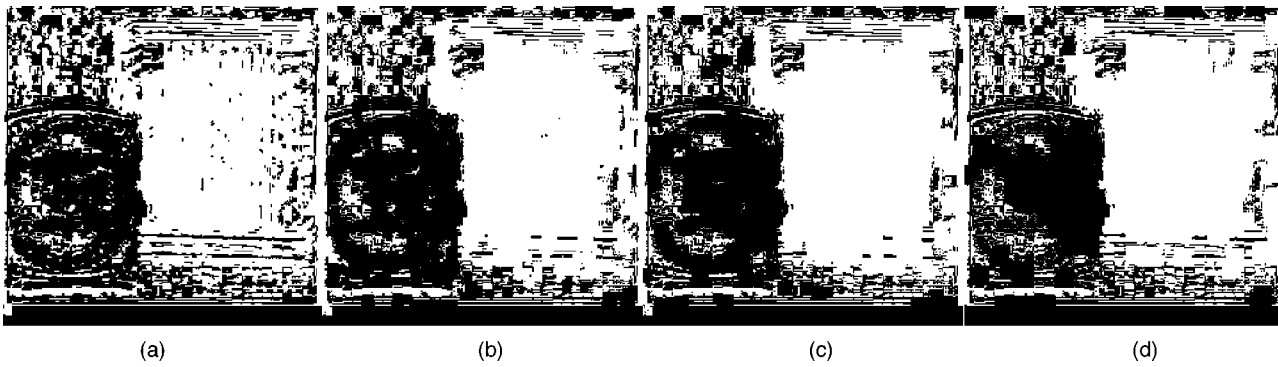


Fig. 7 Activity map from Fig. 1(d), smoothed by using a threshold value of $P_0=0.5$ and different window sizes: (a) 3×3 , (b) 7×7 , (c) 11×11 , and (d) 15×15 . Note the robustness of the scheme in relation with the smoothing window size. The main structures are preserved without introducing smoothing and the isolated points are removed. In the examples described in Sec. 4, a window size of 11×11 pixels was chosen.

$$\binom{Q-p}{p} = \begin{cases} \binom{Q-p}{p} & \text{if } Q-p \geq p \\ 0 & \text{otherwise} \end{cases}$$

Figure 6 shows the connection probability plots for a window size of 15×15 pixels. Results of using both exact [Eq. (12)] and approximate [Eq. (14)] equations are practically identical when considering appropriate p and q values. However, the computational cost of using Eq. (14) is much lower than of Eq. (12). Figure 7 shows an example of the behavior of the algorithm for different window sizes. Note that the holes in the object can be suppressed by interchanging the role of 0's and 1's in the previously described algorithm.

7 Appendix B

The quality measures selected to compare two given images X and Y are: 1. the root mean square error

$$\text{RMSE} = \left\{ \frac{\sum_{i=1}^N \sum_{j=1}^N [X(i,j) - Y(i,j)]^2}{N^2} \right\}^{1/2},$$

2. peak signal-to-noise ratio

$$\text{PSNR} = 10 \log_{10} \left(\frac{255^2}{\text{MSE}} \right),$$

where MSE stands for the mean square error between the images (note that the RMSE is the square root of the MSE), and 3. the percentage of correct decisions

$$P_c = \frac{N_c}{N_t} \times 100,$$

where N_c and N_t are the number of correct decisions and the total number of decisions, respectively.

Acknowledgments

This work has been partially supported by the following grants: TIC2001-3697-C03-02; III PRICIT of the Comunidad Autonoma de Madrid; the IM3 medical imaging thematic network from the Instituto de Salud Carlos III, and the following bilateral projects: HA2001-0087 CSIC-DAAD and 2003CZ0009 CSIC—Academy of Sciences of the Czech Republic.

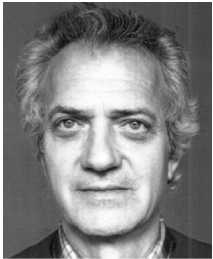
References

1. R. L. Lagendijk and J. Biemond, "Basic methods for image restoration and identification," in *Handbook of Image and Video Processing*, A. Bovik, Ed., pp. 125–139, Academic Press, San Diego, CA (2000).
2. D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.* **13**(3), 43–64 (1996).
3. A. N. Rajagopalan and S. Chaudhuri, "Space-variant approaches to recovery of depth from defocused images," *Comput. Vis. Image Underst.* **68**(3), 309–329 (1997).
4. D. Kundur, D. Hatzinakos, and H. Leung, "A novel approach to multispectral blind image fusion," *Proc. IEEE* **3067**, 83–93 (1997).
5. Z. Zhang and R. S. Blum, "A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application," *Proc. SPIE* **87**, 1315–1328 (1999).
6. L. Cohen, "Generalized phase-space distribution functions," *J. Math. Phys.* **7**, 781–786 (1966).
7. L. Cohen, *Time-Frequency Analysis*, Prentice Hall, Englewood Cliffs, NJ.
8. L. D. Jacobson and H. Wechsler, "Joint spatial/spatial-frequency representation," *Signal Process.* **14**, 37–68 (1988).
9. H. O. Bartelt, K. H. Brenner, and A. W. Lohmann, "The Wigner distribution function and its optical production," *Opt. Commun.* **1**(2), 32–38 (1980).
10. E. Wigner, "On the quantum correction for thermodynamic equilibrium," *Phys. Rev.* **40**, 749–759 (1932).
11. T. A. C. M. Claassen and W. F. G. Mecklenbräuker, "The Wigner distribution—a tool for time-frequency analysis, Parts I-III," *Philips J. Res.* **35**, 217–250, 276–300, 372–389 (1980).
12. J. C. O'Neill, P. Flandrin, and W. J. Williams, "On the existence of discrete Wigner distributions," EDICS Number: SPL. SP. 2.3 Time-Frequency Signal Analysis (Mar. 1998).
13. K. H. Brenner, "A discrete version of the Wigner distribution function," *Proc. EURASIP, Sig. Process. II: Theories Applications*, pp. 307–309 (1983).
14. C. Gonzalo, "Filtrado optico y digital de imagenes de forma espacialmente variante mediante la funcion de distribucion de Wigner," PhD Thesis, Instituto de Optica, CSIC, Madrid (1989).
15. R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, Vol. 1, Addison-Wesley, Reading, MA, 1992.
16. C. Gonzalo, J. Bescós, L. R. Berriel-Valdós, and J. Santamaría, "Space-variant filtering through the Wigner distribution function," *Appl. Opt.* **28**, 730–736 (1989).
17. N. S. Subotic and B. E. A. Saleh, "Time variant filtering of signals in

the mixed time-frequency domain," *IEEE Trans. Acoust., Speech, Signal Process.* **33**, 1479–1485 (1985).

18. H. R. Tizhoosh, "Fuzzy image processing: potentials and state of the art," *5th Intl. Conf. Soft Computing*, **1**, 321–324 (1998).

19. H. G. Senel, R. A. Peters II, and B. Dawant, "Topological median filters," *IEEE Trans. Image Process.* **10**(12), 1–16 (2001).



Salvador Gabarda received his BS degree in physics from the Universidad Nacional de Educación a Distancia, Madrid, in 1979. From 1980 to 1998 he developed a professional career at Telefónica, working in areas such as telephone traffic, failure analysis, quality and reliability normative and systems engineering, including submarine systems and stations. During this period he held diverse positions of responsibility, such as Engineering Projects Supervisor in Tarragona and later as maintenance and operations manager of all submarine cable stations of Telefónica. Now he is a freelance collaborator of the Instituto de Optica Daza de Valdés in Madrid. His major researching interest is in the area of digital image

processing. He is working on a thesis in this field to apply for his PhD degree.



Gabriel Cristóbal received his MSc and PhD degrees both in telecommunication engineering from the Universidad Politécnica de Madrid, Spain, in 1979 and 1986, respectively. He is currently a research scientist at the Instituto de Optica, Spanish Research Council (CSIC). He was a post-doctoral fellow at the international Computer Science Institute and the Electronic Research Laboratory (University of California, Berkeley) from 1989 to 1992. His current research interests are in joint representations, vision modeling, and image compression. He has published more than 40 papers in technical journals and conferences. He has been responsible for several national and EU research and development projects. He is a Senior Member of the IEEE and a member (and head of the Spanish delegation) of the ISO/IEC/SC29/WG01 JPEG 2000 Committee since 2000.

His current research interests are in joint representations, vision modeling, and image compression. He has published more than 40 papers in technical journals and conferences. He has been responsible for several national and EU research and development projects. He is a Senior Member of the IEEE and a member (and head of the Spanish delegation) of the ISO/IEC/SC29/WG01 JPEG 2000 Committee since 2000.