Analysis of focus measure operators in shape-from-focus

3 AUTHORS, INCLUDING:

Said Pertuz
Industrial University of Santander

11 PUBLICATIONS **61** CITATIONS

SEE PROFILE

Domenec Puig
Universitat Rovira i Virgili

113 PUBLICATIONS **970** CITATIONS

SEE PROFILE

Analysis of focus measure operators for shape-from-focus

Said Pertuz [a,*], Domenec Puig [a], Miguel Angel Garcia [b]

[a] *Intelligent Robotics and Computer Vision Group, Department of Computer Science and Mathematics, Universitat Rovira i Virgili, Av. Paisos Catalans 26, Tarragona, Spain*
[b] *Department of Electronic and Communications Technology, Autonomous University of Madrid, Francisco Tomas y Valiente 11, 28049 Madrid, Spain*

ARTICLE INFO

ABSTRACT

Shape-from-focus (SFF) has widely been studied in computer vision as a passive depth recovery and 3D reconstruction method. One of the main stages in SFF is the computation of the focus level for every pixel of an image by means of a focus measure operator. In this work, a methodology to compare the performance of different focus measure operators for shape-from-focus is presented and applied. The selected operators have been chosen from an extensive review of the state-of-the-art. The performance of the different operators has been assessed through experiments carried out under different conditions, such as image noise level, contrast, saturation and window size. Such performance is discussed in terms of the working principles of the analyzed operators.

1. Introduction

Estimating depth is an involved problem in computer vision since a dimension is lost during the projection from the 3D real world to a 2D image. If no physical interaction with the scene is allowed (e.g., it is not possible to project a light pattern), depth can be recovered by means of binocular (or trinocular) stereo systems, or by taking several images with a monocular camera whose extrinsic or intrinsic settings are changed for every frame. In the latter approach, shape-from-focus (SFF) has been proposed as a passive method for recovering 3D shapes. Its applications include: PCB inspection, robot manipulation and control [1], 3D model reconstruction [2] and manufacturing [3], among many others.

In particular, depth information is estimated in SFF from sequences of images of the same scene captured with different camera settings in order to change the degree of focus of every image. The local focus variation is then used as a depth cue. The algorithm utilized to measure the focus level for every image pixel is usually referred to as a *focus measure* (FM) operator. Many focus measure operators have been proposed in the literature for both autofocus (AF) and SFF applications.

Specifically, several comparative studies of focus measure operators for AF applications have been carried out in [4–8]. In addition, Subbarao and Tyan [9] proposed a theoretical method for assessing the uncertainty of various focus measure operators as a function of gray-level noise. In AF, the focus measure is used to determine the position of the best focused image. In that scope, the operator is applied to the whole image. In SFF, however, depth must be estimated for every pixel, with the focus measure operator being applied using a small local window. Therefore, the results of comparative studies about FM operators applied to AF can be extrapolated to SFF in a very limited way. Malik and Choi [10] studied the performance of several focus measure operators under different illumination conditions and window sizes.

The present paper presents a much more extensive analysis of up-to-date focus measure operators applied to SFF by taking into account both the operators and the reconstruction technique itself. An extensive set of experiments has been carried out using both synthetic and real image sequences in order to compare the performance of different focus measure operators for SFF under varying conditions of noise, contrast, saturation and window size. The results are analyzed and discussed according to the working principle of the tested operators. Real images have been acquired with two different monocular cameras, whereas synthetic images have been generated by a non-linear, shift-variant model of defocus. In addition, some focus measure operators previously applied to AF have been adapted to SFF.

This document is organized as follows: the next section presents the SFF technique, including both the reconstruction scheme and a summary of the analyzed FM operators. Section 3 describes the methodology that has been followed in order to compare the performance of those operators. Finally, experimental results and conclusions are presented in Sections 4 and 5, respectively.

2. Shape from focus

In order to generate the depth-map of a scene through SFF, it is necessary to estimate the distance of every point of the scene

---

* Corresponding author. Tel.: +34 634901552.
E-mail addresses: said.pertuz@urv.cat (S. Pertuz),
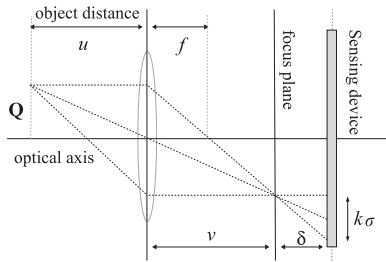domenec.puig@urv.cat (D. Puig), miguelangel.garcia@uam.es (M.A. Garcia).

Fig. 1. Defocus using the thin lens model.

from the camera by measuring its relative degree of focus in the images where that point appears. Fig. 1 illustrates the effect of defocusing in an image according to the *thin lens* model. If the sensing device is at a distance δ from the focus plane, point Q will be projected onto a circle of diameter $k\sigma$, and its radiance will spread over this circle, with Q being defocused.

The location v of the focal plane depends on the focal length of the lens, f, and the distance from the object, u. The relationship between these three variables is given by the well known Gauss lens law or thin lens equation: $1/f = 1/u + 1/v$.

Since there is a one-to-one correspondence between the object distance u and the focal plane location v, the maximum focus will only be achieved at a certain object distance. Therefore, the evaluation of the focus level is critical in applications such as autofocusing and shape-from-focus.

The SFF process consists of two main stages: focus measure and scene reconstruction. The first step aims at measuring the degree of focus of an image (or imaged point), whereas scene reconstruction aims at generating accurate 3D shapes (generally depth-maps) from the focus information obtained from the processed images. The main focus measure operators proposed in the literature are described below.

2.1. Focus measure operators

A wide variety of algorithms and operators have been proposed in the literature to measure the degree of focus of either a whole image or an image pixel for different applications, such as SFF and autofocus. In this work, focus measure operators have been grouped into six broad families according to their working principle. A brief description of each family is presented in this section. The reader is referred to Appendix A for a more detailed description of every operator. Some of the operators used in this work were originally devised for autofocus applications in order to measure the focus level of a whole image. Therefore, they have been modified and adapted in order to make them applicable to SFF. Table 1 summarizes the abbreviations used in this work to refer to the different focus measure operators. Abbreviations in boldface indicate algorithms that have been adapted from auto-focus and have not been previously used in SFF.

The six families of focus operators used in this work are:

1. *Gradient-based operators* (GRA*). This family groups focus measure operators based on the gradient or first derivative of the image. These algorithms follow the assumption that focused images present more sharp edges than blurred ones. Thus, the gradient is used to measure the degree of focus.
2. *Laplacian-based operators* (LAP*). Similarly to the previous family, the goal of these operators is to measure the amount of edges present in images, although through the second derivative or Laplacian.
3. *Wavelet-based operators* (WAV*). The focus measure operators within this family take advantage of the capability of the coefficients of the discrete wavelet transform to describe the frequency and spatial content of images. Therefore, these coefficients can be utilized to measure the focus level.
4. *Statistics-based operators* (STA*). The focus measure operators within this family take advantage of several image statistics as texture descriptors in order to compute the focus level.
5. *DCT-based operators* (DCT*). Similarly to the wavelet-based operators, this family takes advantage of the discrete cosine transform (DCT) coefficients in order to compute the focus level of an image from its frequency content. None of the operators within this family have previously been used in SFF applications to our knowledge.
6. *Miscellaneous operators* (MIS*). This family groups operators that do not belong to any of the previous five groups.

2.2. Scene reconstruction

Once a sequence of images is acquired at different focal positions and the relative degree of focus is measured for all their pixels using any of the focus measure operators described above, the second stage in SFF applies a reconstruction scheme that uses the focus information of those images in order to estimate the depth of every point of the scene.

A widely used technique for this task was proposed by Nayar [11] under the assumption of a Gaussian model of defocus (see Section 3.1). For every pixel, the image with the highest focus measure is identified and the depth corresponding to that pixel estimated by interpolating a Gaussian function around this position. In Fig. 2, the different focus measure values for a pixel at coordinates (i, j) are shown for a focus sweep of 50 images. The values F_{m-1} and F_{m+1}, as well as the maximum focus measure F_m, are used to interpolate a Gaussian function. Pixel $I(i,j)$ is assumed to be located at the depth corresponding to the maximum of the Gaussian. Alternatively, quadratic and polynomial fits have also been proposed [12].

Focus measure operators often work under the so-called *isoplanatic* assumption, by approximating the imaged surface around each pixel as a planar patch. Some researchers have further improved the obtained depth-map by means of recon-struction techniques that take into account the continuous nature of the imaged scene, without the isoplanatic restriction [12]. The proposed techniques include surface fitting and optimization by neural networks [13], and dynamic programming [14,15], among others (e.g., [16–18]). This work aims at focus measure operators. Therefore, a simple Gaussian interpolation technique has been applied in order to obtain the final depth-maps.

3. Methodology

The focus measure operators introduced in the previous section have been applied to sequences of both synthetic and real images in order to obtain the depth-maps of different scenes. The results have been evaluated by applying a quality measure to those maps. In order to assess the robustness of the focus measure operators, the evaluation procedure has been repeated under different factors: noise level, size of the evaluation window, image contrast and saturation. In this section, the image set used to test the focus measure operators is first described, emphasizing on the generation of the synthetic images. Afterwards, the evaluation procedure utilized to compare the accuracy of the obtained depth-maps is presented. Finally, the methodology to assess the robustness to noise, contrast, saturation and window size is described.

Table 1
Abbreviations of focus measure operators.

| Focus operator	Abbr.	Focus operator	Abbr.
Gradient energy	GRA2	Gray-level variance	STA3
Gaussian derivative	**GRA1**	Gray-level local variance	**STA4**
Thresholded absolute gradient	GRA3	Normalized gray-level variance	STA5
Squared gradient	GRA4	Modified gray-level variance	**STA6**
3D gradient	GRA5	Histogram entropy	STA7
Tenengrad	GRA6	Histogram range	STA8
Tenengrad variance	**GRA7**	DCT energy ratio	**DCT1**
Energy of Laplacian	LAP1	DCT reduced energy ratio	**DCT2**
Modified Laplacian	LAP2	Modified DCT	**DCT3**
Diagonal Laplacian	LAP3	Absolute central moment	**MIS1**
Variance of Laplacian	**LAP4**	Brenner's measure	**MIS2**
Laplacian in 3D window	LAP5	Image contrast	**MIS3**
Sum of wavelet coefficients	**WAV1**	Image curvature	**MIS4**
Variance of wavelet coefficients	**WAV2**	Hemli and Scherer's mean	**MIS5**
Ratio of the wavelet coefficients	WAV3	Local binary patterns-based	**MIS6**
Ratio of curvelet coefficients	WAV4	Steerable filters-based	**MIS7**
Chebyshev moments-based	**STA1**	Spatial frequency measure	**MIS8**
Eigenvalues-based	**STA2**	Vollath's autocorrelation	**MIS9**
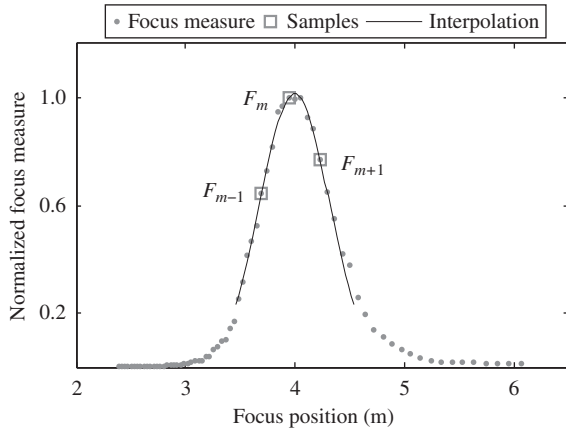
Fig. 2. Gaussian interpolation for depth estimation.

3.1. Image sequences

A set of 12 image sequences has been used to test the performance of the different focus measure operators introduced in Section 2. As described in Table 2, this test set has been divided into three groups. The first two groups correspond to image sequences captured with commercial cameras, whereas the third group contains sequences generated synthetically (Fig. 3)[1]. The procedure to generate those synthetic sequences is described below.

A defocused image is often considered in the literature to be a filtered version of a focused one. Thus, a defocused image, I_d, can be described as the convolution of the focused one, I, with a blurring function h

$$I_d = I*h. \tag{1}$$

Function h is referred to as a *Point Spread Function* (PSF), since it is the response of the camera to a unit point source [19]. In diffraction limited optics with incoherent illumination, the PSF can be simplified as a Gaussian [19–21]

$$h(\alpha,\beta) = \frac{1}{2\pi\sigma_h^2} \exp\left(-\frac{\alpha^2+\beta^2}{2\sigma_h}\right), \tag{2}$$

where σ_h is assumed to be proportional to the degree of defocus of the image.

Pentland [22] derived an expression that relates the blur parameter σ_h in (2) and the depth of the scene point u. This expression can be rewritten for every pixel $I(x,y)$ as

$$\sigma_h(x,y) = \frac{\kappa f^2 |u-u_f|}{A\,u(u_f-f)}, \tag{3}$$

where u_f is the in-focus position for a given camera setting (the distance at which the scene point should be placed in order to be focused), κ is a constant that depends on the camera, and A is the f-number of the lens.[2] However, the convolution in (1) is only valid under the assumption of a spatially invariant blurring function within the evaluation window (isoplanatism) [23,24]. Therefore, in order to avoid the isoplanatism assumption for the synthetic data used in this work, a blurred image $B_{x,y}$ is obtained for every scene point at coordinates (x, y) by convolving it with its corresponding PSF

$$B_{x,y} = I(x,y)*h_{x,y}, \tag{4}$$

where the sub-index of $h_{x,y}$ denotes the PSF corresponding to pixel $I(x,y)$ according to its depth. In turn, the defocused image for the pixel located at (x_0,y_0) given a $W \times H$ image is obtained by summing up the contributions of every defocused point

$$I_d(x_0,y_0) = \sum_{i=0}^{W-1}\sum_{j=0}^{H-1} B_{i,j}(i-x_0,j-y_0). \tag{5}$$

In the above equations, however, since every point is linearly convolved with its corresponding PSF, the overall processing is non-linear and allows the definition of a shift-variant PSF. In addition, the computation time can be reduced by taking into account that not all of the blurred images $B_{x,y}$ must be considered in (5), since the radiance of every depicted point spreads over a small image area depending on the value of its corresponding σ_h (i.e., the values of $h_{x,y}$ can be neglected for pixel coordinates beyond $2.5\sigma_h$ away from (x_0,y_0)). Therefore, the sum in (5) can be limited to those pixels that fulfill

$$(x-x_0)^2 + (y-y_0)^2 \le 6.25\sigma_h^2(x,y). \tag{6}$$

---

[1] Further details on the real scenes can be found at http://www.sayonics.com/research/focus_measure.html.

[2] The f-number is computed from the focal length, f, and the lens diameter, d, as the ratio f/d.
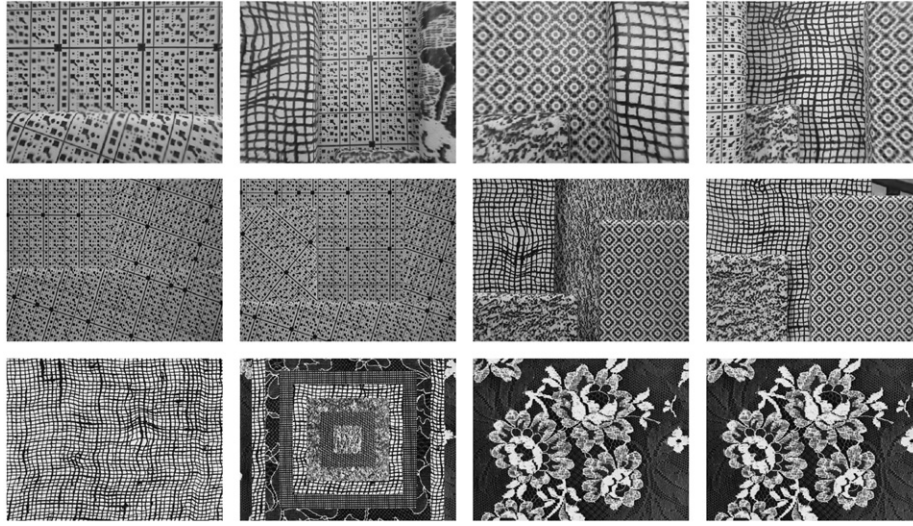
Fig. 3. All-in-focus images for each sequence of the test set. Top row: sequences from the Logitech camera; second row: sequences from the Sony camera; third row: simulated sequences.

Table 2
Image sequences.

Source	Sequences	Description
Logitech Orbit AF camera	4 sequences. 51 images of 640×480 pixels per sequence	Images acquired with an off-the-shelf webcam with controllable focus. Focus sweep between 11.9 mm and 81.0 mm away from the camera
Sony SNC-RZ50P camera	4 sequences. 50 images of 640×480 pixels per sequence	Images acquired with a controllable zoom-focus camera (PTZ). The focus sweep was performed at maximum focal length to obtain minimum depth-of-field. For all scenes, the imaged objects are within a distance range of 36.8 cm and the focus sweep was performed in a range of 1880 mm around that distance
Synthetic images	4 sequences. 25 images of 626×626 pixels per sequence	Defocus simulated for a 3.3 mm focal length camera, focusing between 50 mm and 200 mm. For all scenes, the imaged surface is between 100 mm and 150 mm away from the simulated camera

In this work, all simulated sequences consist of 25 images of 640×640 pixels. The textures mapped on the synthetic surfaces were selected in order to cope with a variety of features and intensities.

3.2. Evaluation procedure

In order to compare the performance of different focus measure operators for SFF, a measure to evaluate the quality of the obtained depth-maps must be utilized. In this paper, the inverse of the root mean square error (RMSE) has been used as a quality measure Q computed for $W \times H$ images by considering both the ground-truth G_T and the estimated depth-map Z as

$$Q = 1/RMSE, \tag{7}$$

$$RMSE = \sqrt{\frac{1}{WH} \sum_{(i,j)} (G_T(i,j) - Z(i,j))^2}. \tag{8}$$

The larger Q, the higher the quality of the considered focus measure operator. In addition, the *universal quality index* (UQI) originally proposed in [25] has also been used. This index measures the distortions between Z and G_T in the range $[-1,1]$. The best value 1 is achieved when the estimated depth-map and the ground-truth match perfectly. The UQI is evaluated locally and averaged over the whole image. This index is computed for two local regions z and g from the estimated depth-map and the ground-truth, respectively

$$UQI(z,g) = \frac{4\sigma_{zg} \overline{g} \, \overline{z}}{(\sigma_z^2 + \sigma_g^2)(\overline{z}^2 + \overline{g}^2)}, \tag{9}$$

where $\overline{z}$ and $\overline{g}$ are the means, and σ_z, σ_g and σ_{zg} the variances and covariance of z and g, respectively.

The quality measure Q is sensitive to both local and global differences between the ground-truth and the actual depth-map. However, this measure has an important drawback: since the value of the RMSE is not normalized and depends on the units of the compared variables, the measures obtained for a given image sequence cannot be directly compared to those of another image sequence (with different scene characteristics, distance range, etc.). To overcome this problem, the evaluation of every focus operator for each sequence is normalized by the maximum evaluation achieved by any of the operators in that sequence. This relative quality measure Q_r has the advantage that its value only depends on the relative performance of the compared operators, thus being independent of the scales and units of both the ground-truth and the obtained depth-maps.

For illustration purposes, Table 3(a) and (b) shows the evaluation procedure for two different image sequences. For the sake of clarity, only the last three operators in Appendix A have been shown in these tables.[3]

Table 3(a) corresponds to an image sequence obtained with the Logitech webcam, whereas Table 3(b) corresponds to the Sony SNC-RZ50P PTZ camera. It can be clearly appreciated that the RMSE values obtained by the operators differ in at least one order of magnitude between both sequences. For instance, the WAV1 operator yielded an RMSE of 1.8 mm/pixel for the first sequence, and 51 mm/pixel for the second. This is not unexpected, since the

---

[3] No special preference is given to any of these operators. They were selected due to their similarity in both definition and performance.

Table 3
Evaluation procedure.

Operator	RMSE (mm/pixel)	Q	Qr
(a) Sequence A			
WAV1	1.8	0.568	1.000
WAV2	2.1	0.471	0.830
WAV3	2.1	0.462	0.815
(b) Sequence B			
WAV1	51.1	0.019	1.000
WAV2	56.9	0.017	0.898
WAV3	54.5	0.018	0.938

first sequence corresponds to a focus sweep of 69 mm, whereas the second sequence corresponds to a sweep of 1880 mm. Therefore, such a difference in the scales and conditions of the experiments could lead to an erroneous interpretation of the results. In contrast, the values in the third column of both tables Q_r are dimensionless, although they clearly differentiate the relative performance of the different focus operators.

In the field of autofocus, the sequence of focus measure values of a focus sweep is usually referred to as *focus function*. Several features of that function have been proposed and applied in order to compare and rank different FM operators [4,7,8,26]. For instance, FM operators were ranked in [4] according to four features of the focus function: *accuracy* (deviation of the focus position from its correct value), *range* (height of the focus function), *number of false maxima* and *width* (measured at 50% of its peak value). These criteria were selected under the assumption that, for an automated focusing algorithm to work appropriately, the focus function should be monotonically decreasing as the distance from the focal plane increases, should not have any phase reversals or local maxima (or false maxima), and should come to a sharp peak at the position of best focus. Santos et al. [7] complemented this methodology by also considering the *execution time* among the key features of the focus measure operator. In [8] and [26], an extensive evaluation of several focus measure operators using an approach similar to [4] and [7] was carried out, but additionally including the noise level (energy of the local maxima) and the resolution (global distribution of the focus function) among the features of the focus function.

Although the aforementioned features of the focus function have successfully been utilized for assessing the quality of FM operators in autofocus applications, their usefulness for SFF is rather limited. In autofocus, these features are important since they directly affect the speed and accuracy of the focusing process. In contrast, the main performance measure in SFF is the reconstruction error (often measured in terms of the RMSE). In order to support this argument, the following experiment has been carried out: for each FM operator, a sample of five hundred point locations on a real focus sequence were randomly selected and the features corresponding to those positions computed, namely: the accuracy (*Acc*), range (*Ran*), width (*W*), false maxima (F_{max}), the resolution (*Res*) and noise level (N_{Lev}); as well as the proposed relative quality measure Q_r and the reconstruction error (RMSE). Table 4 shows the Spearman's rank correlation coefficients of those variables with respect to the reconstruction error with a significance level $p < 0.1$.

From Table 4, it is evident that the proposed quality measure Q_r is strongly correlated with the quality of the reconstruction in terms of the reconstruction error, with the advantage of being independent of the units and particular characteristics of each experiment.

3.3. Robustness of focus measure operators

The performance of the evaluated focus measure operators has been assessed by taking into account the effect of four different features: image contrast, image saturation, image noise and the size of the neighborhood $\Omega(i,j)$ used to evaluate each operator, as described below.

3.3.1. Sensitivity to window size

In SFF, a focus measure operator is applied to each image pixel by processing a small neighborhood or evaluation window around it. The nature and amount of image information and, hence, the size of the neighborhood used to apply the focus measure operator can strongly affect the performance of a focus measure. Malik et al. [10] addressed the problem of determining the optimum window size for the application of focus measures for shape recovery. Malik et al. observed that increasing the window size can lead to an erroneous estimation of depth due to over-smoothing effects. On the other hand, as noted in [27], small windows increase the sensitivity to noise and to the problem of image occlusion blur [27–29].

In order to evaluate the effect of the window size on the performance of the focus measure operators, five different window sizes have been considered: 21×21, 17×17, 13×13, 9×9, and 5×5.

3.3.2. Robustness to noise

This feature measures the performance of a focus measure operator when the latter is affected by the presence of image noise. A CCD camera has several primary noise sources, such as *fixed pattern noise*, *dark current noise*, *shot noise*, *amplifier noise* and *quantization noise* [30], which can be grouped into both irradiance-dependent and irradiance-independent sources. In that way, a noisy image I_n can be modeled as [31]

$$I_n = f(I + n_s + n_c) + n_q, \tag{10}$$

where I is the original image, $f(\cdot)$ is the camera response function (CRF), n_s is the irradiance-dependent noise component, n_c is the independent noise, and n_q is the additional quantization and amplification noise. According to [31], n_q is neglected, n_s and n_c are assumed to have zero mean and variances $\mathrm{Var}(n_s) = I \cdot \sigma_s^2$ and $\mathrm{Var}(n_c) = \sigma_c^2$, respectively.

All focus measure operators have been evaluated with five different noise levels, assuming an identity function for the CRF used in (10) (see Table 6).

3.3.3. Sensitivity to image contrast

Image contrast is another feature related to the image content that can affect the performance of a focus measure operator. Low contrast images usually contain smooth edges, thus increasing the difficulty to determine the relative degree of focus. Moreover, operators with high sensitivity to image contrast will exhibit a variable behavior over the image field in the presence of image aberrations such as vignetting [32]. In order to assess the robustness of the different operators to reductions of image contrast, the experiments have been repeated by pre-processing the image sequences in order to reduce their contrast. In particular, for every image sequence, contrast was reduced by compressing their histograms through the following histogram equalization transfer function:

$$I_c(x,y) = c(I(x,y) - 128) + 128, \tag{11}$$

where $I_c(x,y)$ is the new image intensity of pixel $I(x,y)$ and c is the histogram compression ratio. This equation allows for a linear compression of the image histogram around its center for gray-levels between 0 and 255. In (11), the slope c of the transfer function is reduced in order to decrease the contrast of the image. This operation must be performed in unsigned integer format to achieve a real compression of the histogram instead of a simple scaling.

Table 4
Spearman's correlation r of quality measures with respect to the RMSE.

Measure	Acc	F_{max}	N_{Lev}	W	$\mathbf{Q}_r$
r	−0.52	0.38	0.32	0.40	−**1.0**

3.3.4. Sensitivity to image saturation

Image saturation can also affect the performance of focus measure operators. In this work, image saturation has been evaluated by adding a constant offset to the original image

$$I_S(x,y) = I(x,y) + S, \tag{12}$$

where $I_S(x,y)$ is the saturated pixel at coordinates (x, y) and S is the saturation level. The values of S are between 0 and 128 in order to obtain a saturation level from 0% to 50%. Again, it is assumed that image gray levels are coded in unsigned integer format and values above 255 are set to 255.

4. Results

4.1. Overall performance

All the evaluated focus measure operators have been implemented in MATLAB and applied to the image sequences described in Section 3.1. Table 5 summarizes the mean computational time obtained for every focus operator for 640×480 images on a Pentium IV quad core at 2.5 GHz. For comparison purposes, all focus operators have been implemented and tested in equal terms. However, the performance of some operators highly depends on their particular application. For instance, the DCT-based operators were originally proposed to exploit the information inherent to some video and image formats.

Regarding the reconstruction accuracy, when the focus measure operators are applied to different image sequences within the same group (e.g., the image sequences from the Sony camera), similar rankings are obtained with respect to their quality measures, Q_r. However, sequences from different groups lead to different rankings. For instance, Fig. 4(a) and (b) compares the ground-truth with the obtained depth map for different sequences from the test set. In Fig. 4, it is evident that changes in the imaging device lead to different quality levels of the obtained reconstruction. This is reasonable if the characteristics of every image group are considered: On the one hand, the sequences from the synthetic set represent an ideal defocus, without noise or optical artifacts. On the other hand, the sequences obtained with the Logitech camera have the highest effects of noise, radial distortion, image field curvature and vignetting due to the lower quality of the camera's optics.

The difference in quality, content and nature of the acquired images can favor some operators while being detrimental to others according to their sensitivity to these factors. Therefore, the results from different image groups are presented separately. Thus, Fig. 5 ranks the focus measure operators according to the mean Q_n obtained for every group of sequences.

The results presented in Fig. 5 show that the overall ranking of focus measure operators is strongly related to the imaging device and the scene. Thus, it is difficult to determine what operator or group of operators will perform better under any imaging condition. This can be explained by considering that each combination of variables, such as the imaging device and the real scene under observation, represent a different scenario in terms of contrast, noise, saturation, etc. Therefore, a given focus measure operator will perform worse or better than others according to its sensitivity to the aforementioned factors. In addition, as shown in

Table 5
Average computation time, t, of evaluated focus measure operators for all the considered image sequences.

Method	t (ms)	Method	t (ms)	Method	t (ms)
GRA3	5.60	MIS9	11.0	WAV2	88.00
GRA4	5.90	MIS8	12.0	GRA5	111.0
MIS9	6.00	GRA1	15.0	WAV3	125.0
DCT3	6.30	LAP4	15.0	LAP5	173.0
MIS2	7.10	MIS4	17.0	STA7	388.7
LAP1	7.10	STA8	17.1	MIS6	540.0
LAP2	7.20	STA3	18.0	STA1	6490
GRA2	7.30	STA4	18.0	STA2	6770
GRA6	9.70	GRA7	18.0	DCT2	8640
STA5	10.0	MIS7	22.0	DCT1	8830
LAP3	10.0	STA4	26.0	MIS1	10100
MIS5	10.7	WAV1	55.0	MIS3	12480

the next section, the size of the evaluation window also affects the relative performance of focus measure operators. Notwithstanding, some trends can be observed, such as some operators that generally exhibit a good performance in all cases (e.g., WAV1, LAP2) or others that yield the worst performance in general (e.g., MIS1, MIS5, MIS6).

4.2. Robustness of focus measure operators

In order to assess the robustness of the evaluated focus measure operators to image noise, contrast, saturation and size of the operator's window, several operators have been pre-selected based on their overall performance. Thus, only the eleven focus measure operators that appear within the top 15 in all the rankings shown in Fig. 5 (STA2, STA3, STA8, LAP2, LAP3, LAP4, LAP5, GRA7, WAV1, WAV2 and WAV3) have been considered. The parameters of the different tests are shown in Table 6. The pre-selected operators are mainly based on four different concepts: the image Laplacian (LAP2, LAP3 and LAP5), image statistics (STA2, STA3 and STA8), image gradient (GRA7) and the discrete wavelet transform (WAV1, WAV2 and WAV3). As will be shown in the following sections, operators based on similar concepts exhibit a comparable response to changes in image conditions.

The conducted experiments show that the family of Laplacian-based operators have the best overall performance at normal imaging conditions (i.e., without addition of noise, contrast reduction or image saturation). The image Laplacian is a discrete approximation of the second derivative of the image and highlights regions with rapid changes in intensities. This makes it suitable for detecting changes in focus. These results are in agreement with [6] for autofocus applications, and [9] for SFF. In contrast, Sun et al. [8] found that statistics-based methods have a better performance for autofocus.

4.2.1. Sensitivity to window size

In general, the performance of all focus measure operators decreases for small evaluation windows. For instance, Fig. 6(a) shows the mean RMSE obtained by the different operators for all the real sequences of the test set.

From Fig. 6(a), it is evident that, as the window size is reduced, the RMSE increases. This is expected since the size of the evaluation window directly affects the amount of texture and image information that makes it possible to detect changes in focus. However, as already noted by previous researchers, increments in the window size yield a reduction of spatial resolution.[4]

---

[4] In this work, the reduction of spatial resolution did not lead to greater errors probably since the reconstructed scenes mostly consist of planar patches.
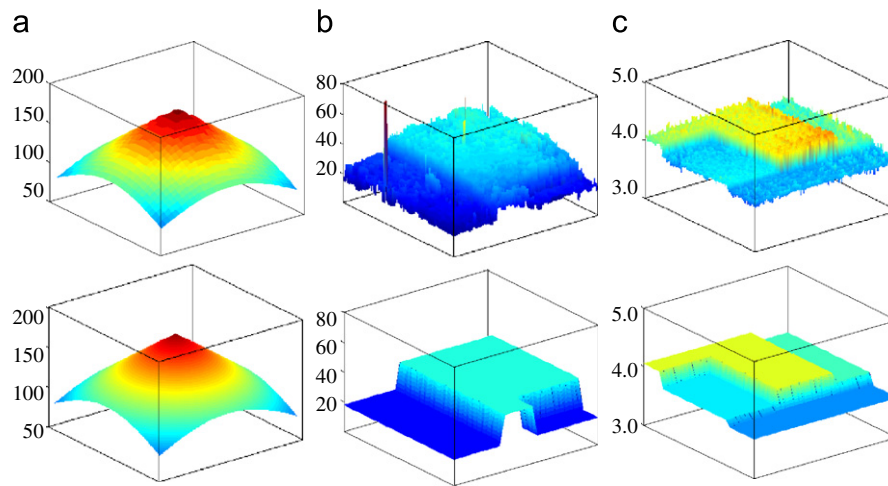
Fig. 4. Depth-maps obtained with SFF using the LAP2 operator with a 7×7 window (top row) and the corresponding ground-truths (bottom row). (a) Simulation. (b) Logitech camera. (c) Sony camera. Z-axis: pixel depth (mm).

Malik and Choi [10] tackled the problem of finding the optimum window size for SFF. In particular, they pointed out that increasing the window size yields a reduction of the quality of the reconstruction by excessively smoothing the depth-map. Therefore, the optimum window size for a particular application must be a trade-off between spatial resolution and robustness to the lack of texture. In this section, the effect of the window size on the relative performance of focus measure operators according to their working principle is assessed.

In order to compare the influence of the window size, Fig. 6(b) shows the mean of the normalized quality measure Q_n for all the analyzed focus measure operators. The results indicate that the differences between the Laplacian-based and statistics-based operators tend to decrease as the size of the evaluation window is reduced. On the other hand, the gradient-based operator, GRA7, is the operator most affected by the reduction of window size. In contrast, the performance of the wavelet-based operators shows a significant improvement. This result is significant since wavelet-based operators have not been applied to SFF in the literature.

From the tests described in this section and throughout this work, it is possible to observe that focus measure operators based on similar concepts respond similarly to variations in the imaging conditions. Therefore, it is easier and more meaningful to understand the behavior of the various families of focus measure operators instead of each operator on its own.[5] In this way, Fig. 7 shows the mean performance of each family of focus measure operators after averaging the quality measures obtained by the operators within the same family for each window size.

Fig. 7 confirms that wavelet-based operators perform better for small evaluation windows, whereas gradient-based operators are the most sensitive to this feature. The wavelet decomposition of an image can be interpreted as a simultaneous frequency and scale-space analysis where the detail sub-bands are related to the highest frequencies of the image [33]. This fact has previously been exploited in the field of autofocus for the definition of focus measure operators, in image fusion for the computation of all-in-focus images [34,35] and in image compression for the JPEG2000 standard [36]. According to the theory of defocus, changes in focus mostly affect the high frequency components of the image. This explains why wavelet-based operators improve their relative

performance as the window size decreases. Actually for small windows, the change in focus can successfully be detected at the coefficients of the low scale sub-bands of the DWT. For instance, Fig. 8 compares the depth-maps obtained using a wavelet-based operator (WAV1) against those generated with a Laplacian-based operator (LAP2) for different window sizes. It is evident that, their performance is comparable for the largest window. However, as the size decreases, the response of the Laplacian-based operator quickly deteriorates, while the wavelet-based operator responds more robustly. The sequence used to generate the depth-maps of Figs. 8 and 11 corresponds to the right-most synthetic sequence of Fig. 4.

It is also important to remark that, for a certain evaluation window size, the ranking (i.e., the quality measure) of operators may vary depending upon the image set used to perform the tests. Thus, the ranking for a given window size will differ if the sequences from the Sony camera, the Logitech camera or both are used. However, the overall behavior of the different operators is independent of the image sequences used to perform the tests. For example, Fig. 9 compares the mean quality measures obtained by the analyzed families for the sequences acquired with the Logitech camera and the Sony camera, respectively.

By comparing Fig. 9(a) and (b), it is evident that, for a given window size, the families of focus measure operators may be sorted differently depending on the acquisition device. However, they have a similar overall behavior. This reasoning is followed for the tests described in the next sections. For the sake of brevity, the results corresponding to sequences from both cameras are shown.

4.2.2. Robustness to noise

As described in Section 3.3.2, the robustness to noise of the focus measure operators has been assessed by performing 3D reconstructions under five different noise levels (see Table 6). The results are summarized in Fig. 10. In these experiments, image noise was one of the factors that most affected the performance of all operators, with all the measured RMSE increasing with the amount of noise level.

It is important to note that statistics-based operators have the highest robustness to noise, with the STA2 operator being the best. In fact, from noise levels 3–5, statistics-based operators are the ones with the best accuracy. The gradient-based operator, GRA7, also shows a good response to noise, qualifying in the second place for the highest noise level. On the other hand,

---

[5] Individualized performance of the pre-selected focus measure operators can be found at http://www.sayonics.com/research/focus_measure.html.
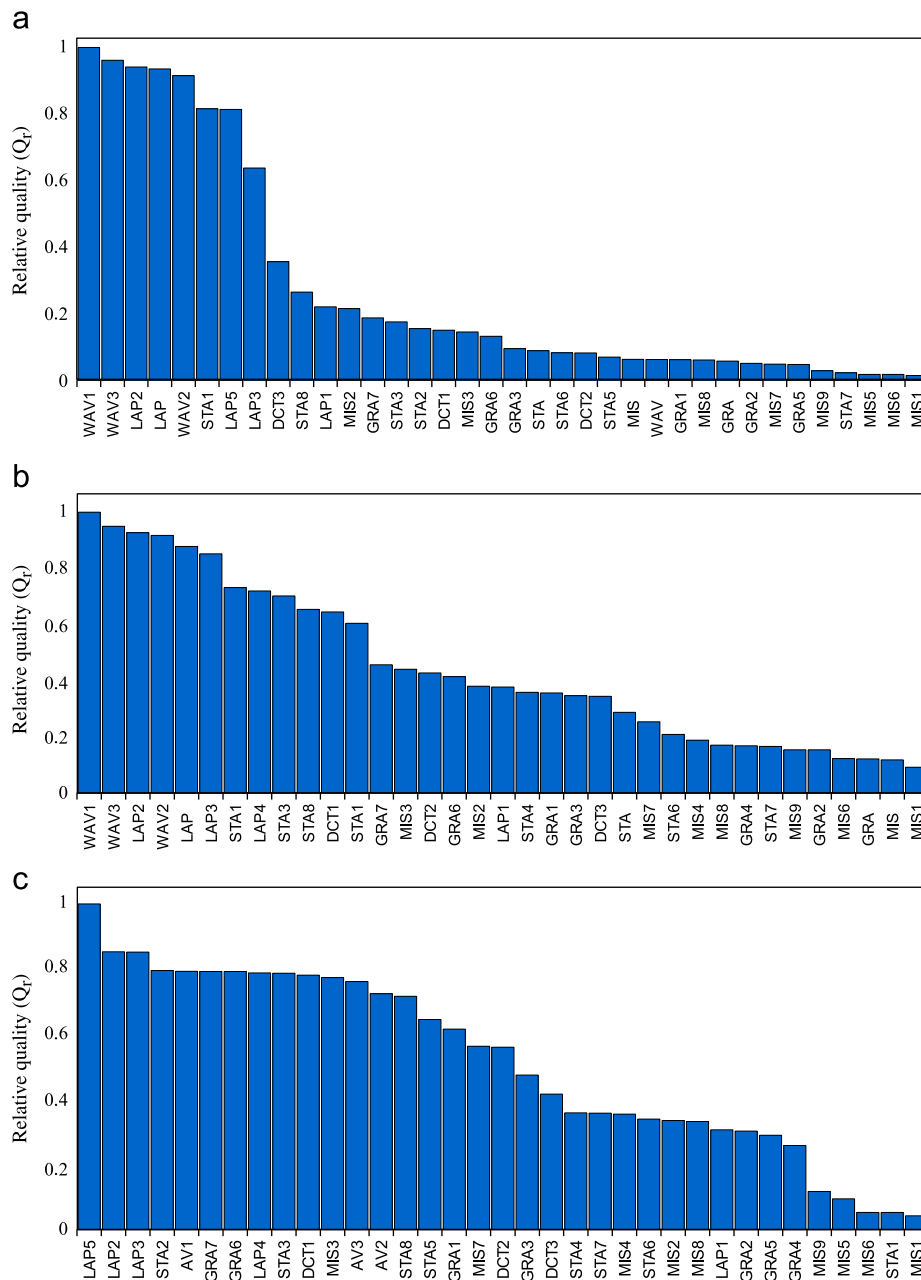
Fig. 5. Ranking of focus measure operators according to the relative quality Q_r (vertical axis) for different groups of image sequences. (a) Synthetic sequences. (b) Logitech's sequences. (c) Sony's sequences.

Laplacian-based operators, which exhibit the best performance at low noise levels (level 1 and 2), are the most sensitive to noise, showing the greatest reduction in their quality measure.

The sensitivity to image noise of Laplacian-based operators is a well known fact and the robustness of statistics-based operators is in agreement with the prior knowledge. On the one hand, low-order statistical moments are theoretically expected to have a low correlation with the high-frequency components of noise. Therefore, according to the properties of linear functions of random variables [37], variance-based operators will detect focus accurately provided that the variance of noise is below the one of the signal (intensity values). For the particular case of the eigenvalues-based operator (STA2), the optimal dimensionality reduction property of principal component analysis (PCA) will

Table 6
Robustness of focus measure operators.

Level	Window size (pixels)	Noise $\sigma_c = \sigma_s$	Contrast (%)	Saturation (%)
Lv1	21×21	0.00050	80.0	10
Lv2	17×17	0.00176	62.5	20
Lv3	13×13	0.00320	45.0	30
Lv4	9×9	0.00429	27.5	40
Lv5	5×5	0.00555	10.0	50

provide this operator with a good robustness to noise as long as the noise energy is below the total energy of the data in the neighborhood of a pixel [38]. This can explain the robustness
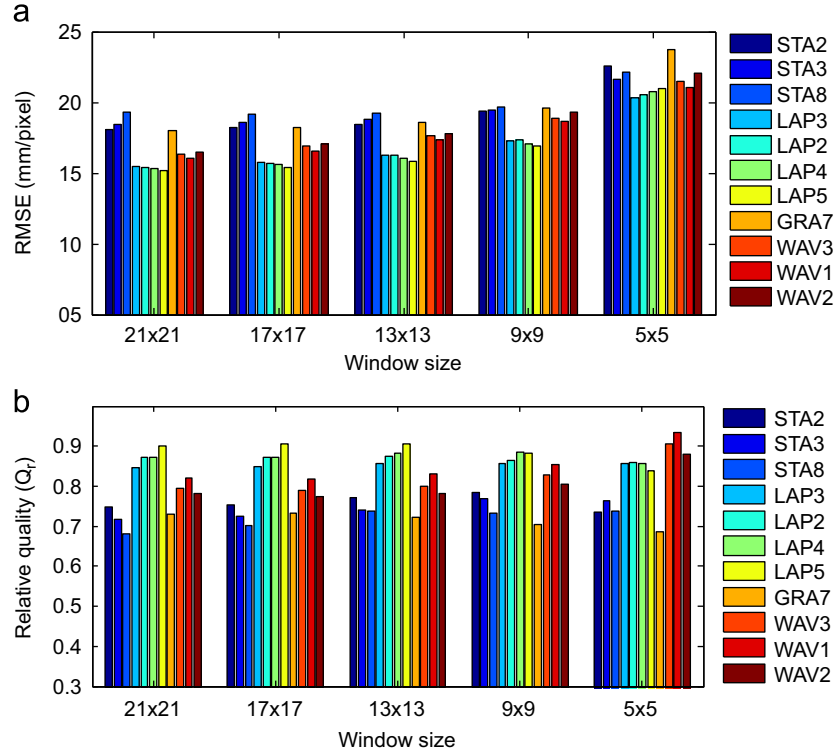
Fig. 6. Sensitivity to windows size. (a) Mean RMSE for all real sequences. (b) Relative quality for all the real sequences.
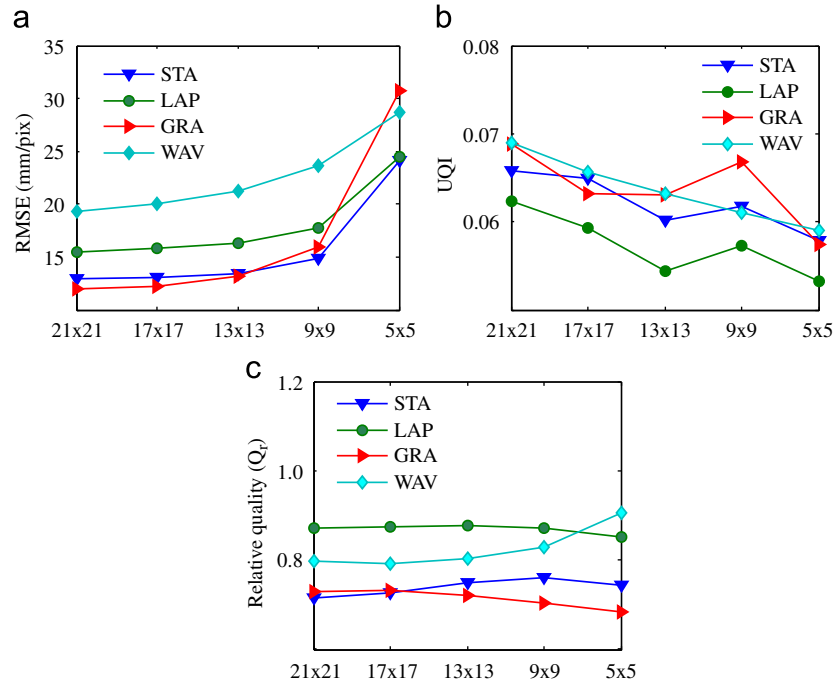

Fig. 7. Quality measures for different families of operators and different window sizes. (a) RMSE. (b) Universal quality index. (c) Relative quality. *X*-axis: window size.

of statistics-based operators to image noise. In contrast, Laplacian-based operators are the most sensitive to noise since it is well known that second derivatives are very sensitive to it [39]. For instance, Fig. 11 compares the depth-maps obtained using a statistical-based operator (STA3) against those generated with a Laplacian-based operator (LAP5). It is evident that the Laplacian-based operator has a better performance for the lowest noise level. However, as the

noise level increases, its performance quickly deteriorates with respect to the statistical-based operator.

In addition, wavelet-based operators also have a high sensitivity to noise. This can be explained by the fact that image noise mostly corresponds to high-frequency components. Therefore, it is likely to have an impact on the coefficients of the detail sub-bands of the DWT. In fact, in the literature related to image denoising, noise is often suppressed by thresholding the coefficients
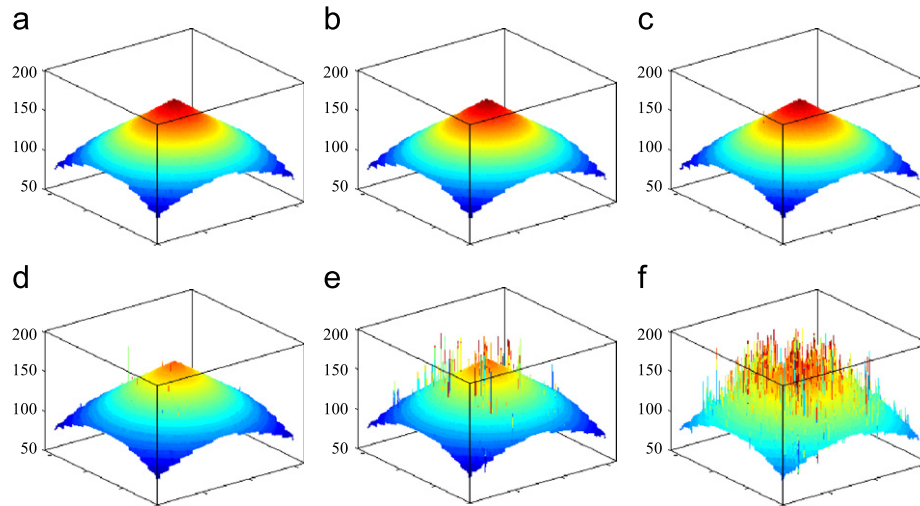
Fig. 8. Depth-maps obtained with SFF using WAV1 (top row) and LAP2 (bottom row) for different window sizes (7×7, 5×5 and 3×3, from left to right). These depth-maps correspond to the ground-truth shown in Fig. 4(a). Z-axis: pixel depth (mm). (a) $Q_r = 1.00$, (b) $Q_r = 0.99$, (c) $Q_r = 0.98$, (d) $Q_r = 0.96$, (e) $Q_r = 0.60$, and (f) $Q_r = 0.26$.
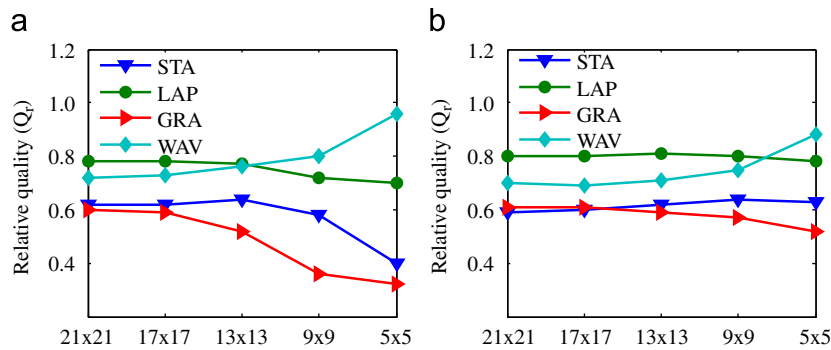

Fig. 9. Performance for different acquisition devices. (a) Logitech camera. (b) Sony camera. X-axis: window size.

of the DWT sub-bands [40]. Thus, the focus measure will deteriorate due to the effects of noise on these sub-bands.

4.2.3. Sensitivity to image contrast

The experiments in this section show that the effects of contrast on the performance of focus measure operators are marginal, only with a slight increase in the RMSE of the obtained depth-maps, even for contrast levels reduced up to 10%. Moreover, the relative performance of the focus measure operators remains almost unaltered for the different contrast levels. The results of the evaluation of focus measure operators for different contrasts are summarized in Fig. 12. From them, it can be concluded that contrast affects all the compared operators similarly, since their relative performance almost remains unchanged.

4.2.4. Sensitivity to image saturation

As can be observed in Fig. 13, the performance of focus measure operators remains unaltered for saturations below 30%. In general, all operators decrease their performance as the saturation level is high, but this behavior is more evident for saturation levels above 30% (Lv3). This can be explained by the fact that, for the imaging conditions of the captured sequences, the upper bounds of the image histogram only contain a small percentage of the total energy. Therefore, low saturation levels only affect a small fraction of image pixels. Thus, the effect of saturation is only significant above 30%. Above this threshold

both Q_r and UQI indicate that Laplacian-based operators have the highest sensitivity to saturation, as can be appreciated in Fig. 13.

In general, features such as image contrast and saturation affect all the analyzed operators similarly. A difference in the relative performance of some operators can only be observed at high levels of contrast and the increment of saturation. Both, the reduction in image contrast and the increment of saturation can be thought of as a reduction in the pixel depth or, in terms of image intensities, as a reduction in the number of gray levels. In this work, all the focus measure operators have been implemented in double-precision arithmetic. Therefore, no quantization or overflow problems arise in the computations independently of the operations performed by each focus measure operator.

5. Conclusions

Focus measure operators are a fundamental part of 3D scene reconstruction through shape-from-focus. In this work, a methodology to compare the performance of several focus measure operators has been proposed and tested. The selected operators have been chosen from an extensive review of up-to-date literature. Since some of them were originally proposed for autofocus applications, it has been necessary to adapt them in order to be applicable to SFF. Experiments have been carried out on a test set constituted by both synthetic and real image sequences. Synthetic sequences were generated based on a non-linear shift-variant computation of defocus.
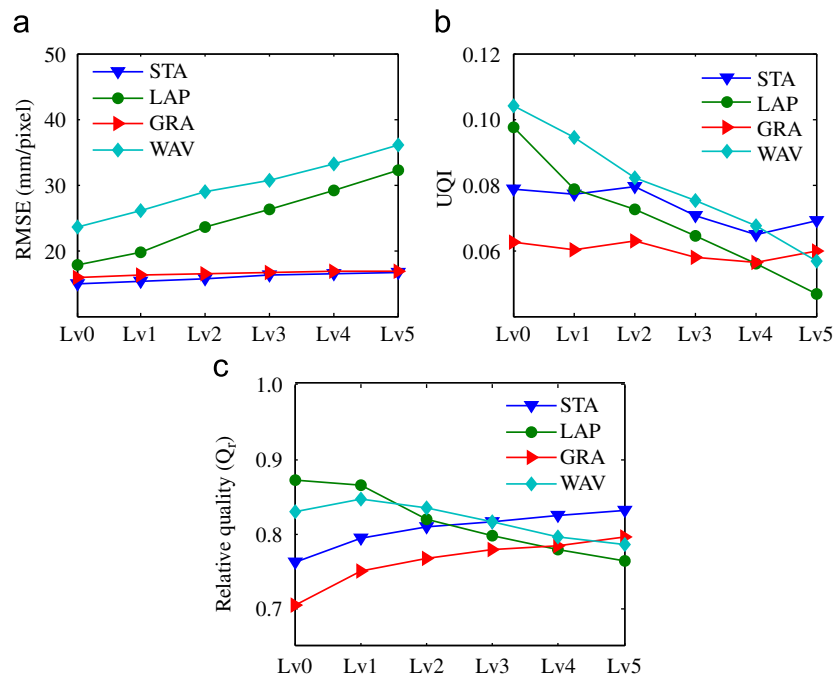
Fig. 10. Quality measures for different families of operators and noise levels. (a) RMSE. (b) Universal quality index. (c) Relative quality.
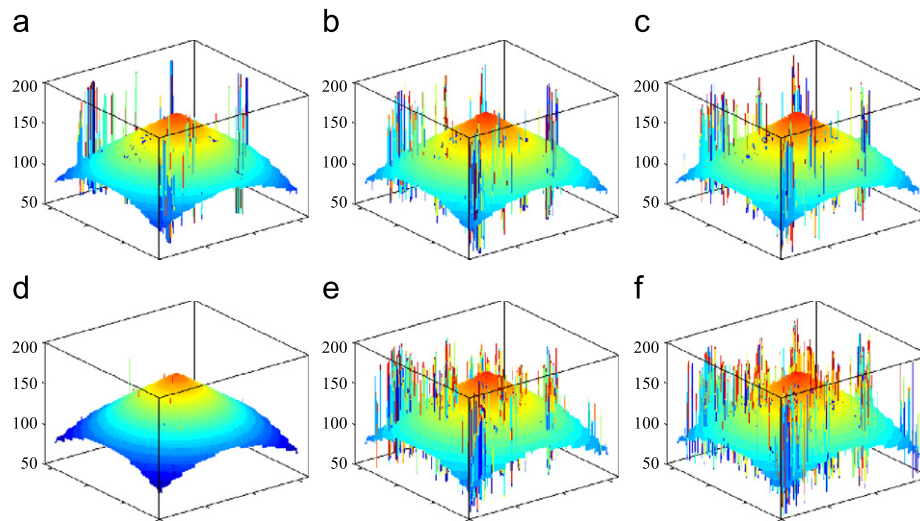

Fig. 11. Depth-maps obtained with SFF using STA3 (top row) and LAP5 (bottom row) with a 7×7 window and increasing noise levels (Lv0, Lv1 and Lv2). These depth-maps correspond to the ground-truth shown in Fig. 4(a). Z-axis: pixel depth (mm). (a) $Q_r=0.21$, (b) $Q_r=0.17$, (c) $Q_r=0.16$, (d) $Q_r=1.00$, (e) $Q_r=0.15$, and (f) $Q_r=0.11$.

The analyzed focus measure operators are based on different mathematical principles. From an initial group of 36 operators, the best 11 operators have been chosen in order to compare their performance under different imaging conditions. The selected group includes algorithms based on the image Laplacian, image statistics, the image gradient and the wavelet transform.

Experiments have shown that Laplacian-based operators have the best overall performance at normal imaging conditions (i.e., without addition of noise, contrast reduction or image saturation). However, it is difficult to determine which focus measure operators have the best performance for specific imaging conditions (i.e., a given noise level, contrast, saturation and window size), since this strongly depends on the particular capturing device with which the image sequences are acquired. Notwithstanding, the overall behavior of the different operators is independent of the capturing device and they respond similarly to changes in noise, contrast, saturation and window size even for different devices. Moreover, experiments have also shown that operators belonging to the same family, which are thus based on similar principles, have a similar response to changes in the imaging conditions.

In summary, the results presented in this work provide an insight on how different imaging conditions can affect the different families of focus measure operators. Moreover, the group of best operators for SFF has been identified, which can be useful for future development of new focus measure operators and reconstruction schemes in this particular field.

The texture content is an important factor that influences the performance of SFF [41–43]. However, the problem of identifying what texture families and what texture features are relevant for focus detection still needs to be assessed. Future work will focus
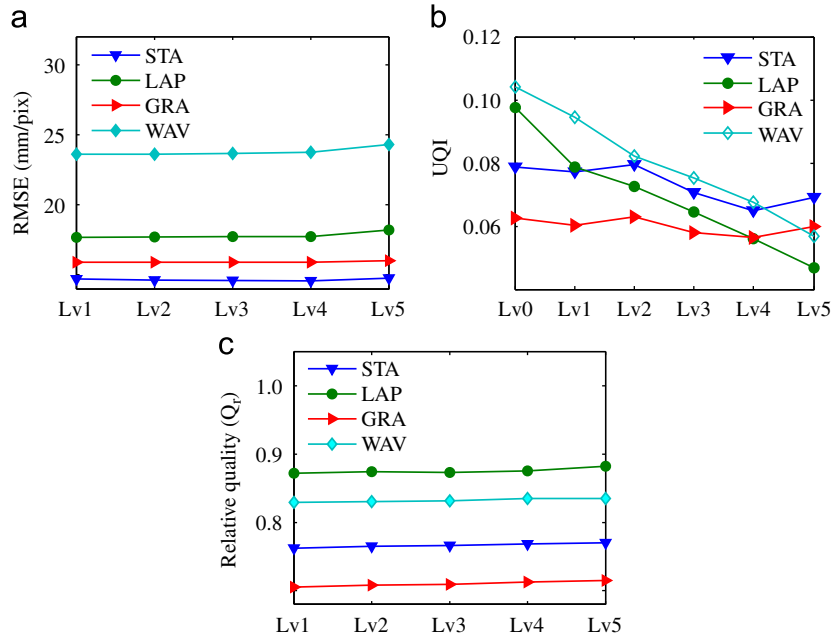
Fig. 12. Quality measures for different families of operators and contrast levels. (a) RMSE. (b) Universal quality index. (c) Relative quality.
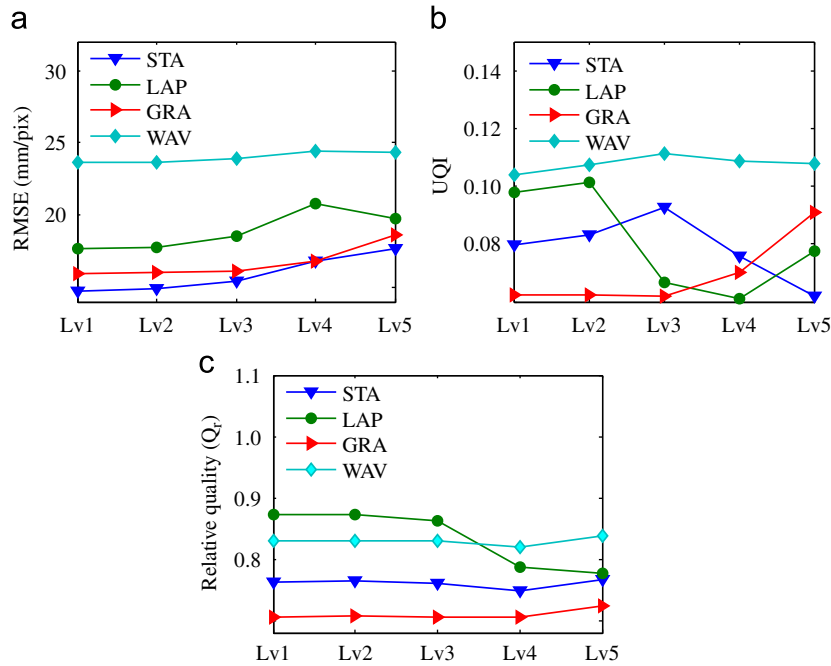

Fig. 13. Quality measures for different families of operators and saturation levels. (a) RMSE. (b) Universal quality index. (c) Relative quality.

on the application of SFF to complex scenes, where changes in characteristics such as the texture content may occur over the same scene. The aim is to study the response of the analyzed focus measure operators to different families of microtextures, such as the ones previously proposed in [44]. In addition, the present work will be expanded by incorporating new published focus measure operators, such as the ones in [45–47].

Appendix A. Focus measure operators

This appendix summarizes the focus measure operators evaluated in this work. Some operators have not been previously used

in the field of SFF and have been adapted from AF applications for this work. For homogeneity, the original notation used by some authors has been adapted.

A.1. Absolute central moment (MIS1)

Shirvaikar et al. [48] proposed a focus measure for AF, the *absolute central moment* (*ACMo*), based on statistical measures and the image histogram H

$$ACMo = \sum_{k=1}^{L} |k-\mu| P_k, \tag{A.1}$$

where μ is the mean intensity value of H, L the number of gray-levels in the image and P_k the relative frequency of the k-th gray-level. This operator has been adapted to SFF by accumulating the values of $ACMo$ computed over the neighborhood $\Omega(x,y)$ of pixel $I(x,y)$.

A.2. Brenner's focus measure (MIS2)

A focus measure based on the second difference of the image gray-levels of an image I is defined as [4,7,8]

$$\phi = \sum_{(i,j)} |I(i,j)-I(i+2,j)|^2. \tag{A.2}$$

A variation in (A.2) also allows taking into account the vertical variations of the image [7]. In addition, the values above a given threshold can only be accumulated [7,8]. This measure can be adapted to SFF if the focus measure for every pixel $I(x,y)$ is computed by limiting the sum in (A.2) to its local neighborhood $\Omega(x,y)$.

A.3. Image contrast (MIS3)

Nanda et al. [49] used the image contrast as a focus measure for autofocus

$$C(x,y) = \sum_{i=x-1}^{x+1} \sum_{j=y-1}^{y+1} |I(x,y)-I(i,j)|, \tag{A.3}$$

where $C(x,y)$ is the image contrast for pixel $I(x,y)$. This operator can be adapted to SFF if the contrast is accumulated over the pixel's neighborhood

$$\phi_{x,y} = \sum_{(i,j)\in\Omega(x,y)} C(i,j), \tag{A.4}$$

A.4. Image curvature measure (MIS4)

This operator was proposed in [50] for SFF applied to microscopy. If the image gray-levels are interpolated by means of a surface, the curvature of this surface may be used as a focus measure [50,51]

$$\phi = |c_0|+|c_1|+|c_2|+|c_3|, \tag{A.5}$$

where $C=(c_0,c_1,c_2,c_3)^T$ is the vector of coefficients used to interpolate a quadratic surface $f(x,y)=c_0x+c_1y+c_2x^2+c_3y^2$. C is computed through least squares by applying two convolution masks [50]

$$c_0 = M_1*I \quad c_2 = \frac{3}{2}M_2*I-M_2^T*I$$

$$c_1 = M_1^T*I \quad c_3 = \frac{3}{2}M_2^T*I-M_2*I revx,$$

where

$$M_1 = \frac{1}{6}\begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \quad M_2 = \frac{1}{5}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

A.5. Helmli and Scherer's mean method (MIS5)

Helmli and Scherer [50] proposed to measure the local contrast by computing the ratio, $R(x,y)$, between the intensity level of every pixel $I(x,y)$ and the mean gray level of its neighborhood $\mu(x,y)$

$$R(x,y) = \begin{cases} \dfrac{\mu(x,y)}{I(x,y)}, & \mu(x,y) \geq I(x,y) \\ \dfrac{I(x,y)}{\mu(x,y)} & \text{otherwise}. \end{cases} \tag{A.6}$$

This ratio is one if there is either a constant gray value or low contrast. An $M \times N$ neighborhood centered at (x, y) is used to compute $\mu(x,y)$. The focus measure for $I(x,y)$ is computed by summing the values of $R(x,y)$ within $\Omega(x,y)$.

A.6. Local binary patterns-based measure (MIS6)

Lorenzo et al. [52] studied the use of Local Binary Patterns (LBPs) as a focus measure for autofocus applications. In order to compute the LBP operator for a given pixel $I(x,y)$, n pixels within a radius R around (x, y) are selected [52]

$$LBP_{x,y}(n,R) = \sum_{k=1}^{n} S(I_k-I(x,y)), \tag{A.7}$$

where I_k is the intensity level of the k-th pixel around (x, y) and $S(x)$ is

$$S(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise}. \end{cases} \tag{A.8}$$

The focus measure for pixel $I(x,y)$ is computed as

$$\phi_{x,y} = \sum_{(i,j)\in\Omega(x,y)} LBP_{i,j}(n,R). \tag{A.9}$$

Values of $n=8$ and $R=2$ have been used in the experiments of the present paper.

A.7. Steerable filters-based measure (MIS7)

Minhas et al. [51] proposed a focus measure based on a filtered version of the image I_f

$$\phi_{x,y} = \sum_{(i,j)\in\Omega(x,y)} I_f(i,j), \tag{A.10}$$

$I_f(i,j)$ is defined as

$$I_f(i,j) = \max\{R_{(i,j)}^{\theta_1},R_{(i,j)}^{\theta_2},\ldots,R_{(i,j)}^{\theta_N}\}, \tag{A.11}$$

where R^{θ_n}, $n=1,2,\ldots,N$ is the image response to the n-th steerable filter defined as [53]

$$R^{\theta_n} = \cos(\theta_n)(I*\Gamma_x)+\sin(\theta_n)(I*\Gamma_y), \tag{A.12}$$

with Γ_x and Γ_y being the Gaussian derivatives (A.15).

Recently, an efficient algorithm for the computation of the focus measure based on steerable filters by means of integral images has been proposed in [54].

A.8. Spatial frequency measure (MIS8)

This operator was proposed in [55] for the fusion of multi-focal images

$$\phi_{x,y} = \sqrt{\sum_{(i,j)\in\Omega(x,y)} I_x(i,j)^2 + \sum_{(i,j)\in\Omega(x,y)} I_y(i,j)^2}, \tag{A.13}$$

where I_x and I_y denote the first derivatives of an image in the X and Y direction, respectively.

A.9. Vollath's autocorrelation (MIS9)

A focus measure based on image autocorrelation was proposed in [7,56] and [8] for autofocus. Its adaptation to SFF is straightforward

$$
\phi_{x,y} = \sum_{(i,j) \in \Omega(x,y)} (I(i,j) \cdot I(i+1,j))
$$
$$
- \sum_{(i,j) \in \Omega(x,y)} I(i,j) \cdot I(i+2,j).
\tag{A.14}
$$

A.10. Gaussian derivative (GRA1)

Based on the defocus modeling, Geusebroek et al. [57] proposed a focus measure for autofocus in microscopy based on the first order Gaussian derivative [6,57]

$$
\phi = \sum_{(x,y)} (I * \Gamma_x)^2 + (I * \Gamma_y)^2,
\tag{A.15}
$$

where Γ_x and Γ_y are the x and y partial derivatives of the Gaussian function $\Gamma(x,y,\sigma)$, respectively

$$
\Gamma(x,y,\sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right).
\tag{A.16}
$$

In order to apply this measure to small neighborhoods in SFF, the value of σ in (A.16) must be computed accordingly. For the results shown in this paper, the value of σ was selected such that, for a neighborhood of size $W \times W$, a total of five σ's are contained along W. The focus measure for a pixel $I(x,y)$ is computed by applying (A.15) within its neighborhood, $\Omega(x,y)$.

A.11. Gradient energy (GRA2)

The sum of squares of the first derivative in the x and y directions has also been proposed as a focus measure [19,55,58]

$$
\phi_{x,y} = \sum_{(i,j) \in \Omega(x,y)} (I_x(i,j)^2 + I_y(i,j)^2).
\tag{A.17}
$$

A.12. Thresholded absolute gradient (GRA3)

The first derivative of the image in the horizontal dimension is a simple measure of its degree of focus

$$
\phi_{x,y} = \sum_{(i,j) \in \Omega(x,y)} |I_x(i,j)|, \quad |I_x(i,j)| \geq T.
\tag{A.18}
$$

The performance of this measure is affected by the selection of T. For the sake of generality, no threshold has been considered in this work. An alternative definition of this method considers both vertical and horizontal image derivatives by either addition [59] or selection of the maximum value [7].

A.13. Squared gradient (GRA4)

Instead of applying (A.18), the first derivative is squared in order to increase the influence of larger gradients [7,8,55,60]. If both vertical and horizontal derivatives are considered and added, this measure is equivalent to the energy of the image gradient (GRAE).

A.14. 3D gradient (GRA5)

Ahmad proposed in [61] the use of the 3D gradient as a focus measure operator. In [61], the whole image sequence is stacked in a single image volume $V(x,y,z)$, where x and y denote the image coordinates and z the image number. The magnitude of the 3D

gradient is given by

$$
|\nabla V| = \sqrt{\nabla V_x^2 + \nabla V_y^2 + \nabla V_z^2},
\tag{A.19}
$$

where the three components of the gradient are obtained by convolving V with the $3 \times 3 \times 3$ operator oriented in the x, y and z direction, respectively. The focus measure at pixel $I(i,j)$ for the k-th image is computed as the sum of the 3D gradient in a small 2D neighborhood, provided this gradient is greater than a threshold T

$$
\phi_{x,y,k} = \sum_{(i,j) \in \Omega(x,y)} |\nabla V(i,j,k)|, \quad |\nabla V(i,j,k)| \geq T.
\tag{A.20}
$$

A.15. Tenengrad (GRA6)

A popular focus measure based on the magnitude of image gradient is defined as [7,8,19,28,50,51,55,58,59,62–69]

$$
\phi_{x,y} = \sum_{(i,j) \in \Omega(x,y)} (G_x(i,j)^2 + G_y(i,j)^2),
\tag{A.21}
$$

where G_x and G_y are the X and Y image gradients computed by convolving the given image I with the Sobel operators.

A.16. Tenengrad variance (GRA7)

This operator uses the variance of the image gradient as a focus measure. It was originally used for autofocus in [65], but can also be applied to SFF

$$
\phi_{x,y} = \sum_{(i,j) \in \Omega(x,y)} (G(i,j) - \overline{G})^2,
\tag{A.22}
$$

where $\overline{G}$ is the mean value within $\Omega(x,y)$ of the gradient magnitude, which is computed as: $G = \sqrt{G_x^2 + G_y^2}$.

A.17. Energy of Laplacian (LAP1)

The energy of the second derivative of the image has been used as a focus measure for both autofocus [6,8,19,55,59, 63,64,66,67,69,70] and SFF [61]

$$
\phi_{x,y} = \sum_{(i,j) \in \Omega(x,y)} \Delta I(i,j)^2,
\tag{A.23}
$$

where ΔI is the image Laplacian obtained by convolving I with the Laplacian mask.

A.18. Modified Laplacian (LAP2)

Nayar [11] proposed a focus measure based on an alternative definition of the Laplacian

$$
\phi(x,y) = \sum_{(i,j) \in \Omega(x,y)} \Delta_m I(i,j),
\tag{A.24}
$$

where $\Delta_m I$ is the modified Laplacian of I, computed as

$$
\Delta_m I = |I * \mathcal{L}_X| + |I * \mathcal{L}_Y|.
\tag{A.25}
$$

The convolution masks used to compute the modified Laplacian are

$$
\mathcal{L}_X = [-1 \ \ 2 \ \ -1],
$$

and $\mathcal{L}_Y = \mathcal{L}_X^T$.

A.19. Diagonal Laplacian (LAP3)

Thelen et al. [71] also included vertical variations of the image in order to compute the modified Laplacian of the image

$$
\Delta_m I = |I * \mathcal{L}_X| + |I * \mathcal{L}_Y| + |I * \mathcal{L}_{X1}| + |I * \mathcal{L}_{X2}|,
\tag{A.26}
$$

where $\mathcal{L}_X$ and $\mathcal{L}_Y$ are defined as in (A.25), and $\mathcal{L}_{X1}$ and $\mathcal{L}_{X2}$ are given by

$$\mathcal{L}_{X1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mathcal{L}_{X2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

A.20. Variance of Laplacian (LAP4)

This measure utilizes the variance of the image Laplacian as a focus measure for autofocus [65]. In SFF, this measure can be defined as

$$\phi_{i,j} = \sum_{(i,j) \in \Omega(x,y)} (\Delta I(i,j) - \overline{\Delta I})^2, \tag{A.27}$$

where $\overline{\Delta I}$ is the mean value of the image Laplacian within $\Omega(x,y)$.

A.21. Laplacian in 3D window (LAP5)

An et al. [72] proposed the use of a 3D neighborhood for accumulating the focus measure

$$\phi_{x,y,k} = \sum_{f=k-1}^{k+1} \sum_{(i,j) \in \Omega(x,y)} |\Delta_M I_f(i,j)|, \tag{A.28}$$

where $\Delta_M I_f$ is the modified Laplacian of the f-th image, computed as in (A.25).

A.22. Chebyshev moments-based (STA1)

An image focus measure for AF based on Chebyshev moments was proposed in [69] as the ratio between the energy of the high-pass band and the energy of the low-pass band extracted from the image by using the Chebyshev moments. In [69], this measure is applied to a normalized image $\tilde{I}$ and can be computed as

$$\phi = \frac{\|\mathbb{H}(\tilde{I};p)\|}{\|\mathbb{L}(\tilde{I};p)\|}, \tag{A.29}$$

where $\|\mathbb{H}(\tilde{I};p)\|$ and $\|\mathbb{L}(\tilde{I};p)\|$ respectively denote the high-order and low-order Chebyshev moments up to order p of the normalized image $\tilde{I}$, which is computed as

$$\tilde{I} = \frac{I}{\sqrt{\sum_{(i,j)}[I(i,j)]^2}}. \tag{A.30}$$

Note that (A.29) and (A.30) must be applied to the whole image in order to compute a single focus measure. Nevertheless, this measure can be used in SFF by performing a sliding-block operation within a neighborhood $\Omega(x,y)$ and assigning the obtained measure to its central pixel. However, this procedure is expected to affect the performance of the operator for small neighborhoods, since the kernels used to compute the Chebyshev moments will lose their discriminating capability as the number of points (window size) is decreased. Parameter p also determines the sensitivity to the frequency components of the image. According to [69] and [73], a value of $p=2$ has been used in this work.

A.23. Eigenvalues-based (STA2)

A sharpness measure of an image proposed in [67] is obtained from the trace of the matrix of eigenvalues, Λ, of the image covariance S. Thus, the variances of the principal components of the image are used as a focus measure [67,73,74]

$$\phi = \text{trace}[\Lambda_k], \tag{A.31}$$

where the trace of Λ_k is the sum of the first k diagonal elements of Λ. k has been set to 5 in this work.

The image covariance S is

$$S = \frac{JJ^T}{MN-1}, \tag{A.32}$$

where J is the normalized image in (A.30) after removing its mean value: $J = \tilde{I} - mean(\tilde{I})$; and $M \times N$ is the size of the neighborhood. pixel's neighborhood. The trace of Λ_k is the sum of its first k diagonal elements. This focus measure, originally proposed for a whole image, can be applied to SFF in a sliding block-like fashion. However, the computational cost is dramatically increased since the normalization procedure in (A.30) is iterated for every pixel's neighborhood $\Omega(x,y)$.

A.24. Gray-level variance (STA3)

The variance of image gray-levels is one of the most popular methods to compute the focus measure of an image. It has been applied to both autofocus [4,7,8,19,55,59,62,64,66–70,75] and SFF [50,51,58,72]

$$\phi_{x,y} = \sum_{(i,j) \in \Omega(x,y)} (I(i,j) - \mu)^2, \tag{A.33}$$

where μ is the mean gray-level of pixels within $\Omega(x,y)$.

A.25. Gray-level local variance (STA4)

Pech et al. [65] proposed the local variance of gray-levels as a focus measure for autofocus of diatoms in brigthfield microscopy. For its application to SFF, this operator is re-formulated as

$$\phi_{x,y} = \sum_{(i,j) \in \Omega(x,y)} (L_v(i,j) - \overline{L_v})^2, \tag{A.34}$$

where $L_v(i,j)$ is computed as the variance of gray-levels within a neighborhood of size $w_x \times w_y$ centered at (i,j). $\overline{L_v}$ is the mean value of L_v. In this work, w_x and w_y have been chosen to coincide with the size of $\Omega(x,y)$.

A.26. Normalized gray-level variance (STA5)

The gray-level variance can be compensated for differences in the average image brightness among different images by normalizing the value of ϕ in (A.33) by the mean gray-level value μ [7,8,63].

A.27. Modified gray-level variance (STA6)

The computation of the gray-level variance in (A.33) can be thought of as a non-linear filtering of the image. An alternative focus measure can be obtained if the mean value $\mu(x,y)$ of every pixel within its neighborhood $\Omega(x,y)$ is computed

$$\phi_{x,y} = \sum_{(i,j) \in \Omega(x,y)} (I(i,j) - \mu(i,j))^2, \tag{A.35}$$

where $\mu(x,y)$ is obtained through a linear convolution filter.

A.28. Histogram entropy (STA7)

Since a focused image is expected to have a higher information content, the entropy and range of the image histogram can be used to compute the focus measure. The histogram entropy

operator is defined as [4,7,8,59,62,70]

$$\phi = -\sum_{k=1}^{L} P_k \log(P_k), \tag{A.36}$$

where P_k is the relative frequency of the k-th gray-level.

In order to compute a focus value for a pixel at coordinates (x, y), the image histogram used in (A.36) is obtained from the gray-level values within $\Omega(x,y)$.

A.29. Histogram range (STA8)

The histogram range has been used as a focus measure for autofocus [4,7,8]

$$\phi = \max(k|H > 0) - \min(k|H > 0). \tag{A.37}$$

In this work, the histogram H is computed within every $\Omega(x,y)$.

A.30. DCT energy ratio (DCT1)

The discrete cosine transform (DCT) is now part of many image and video encoding systems. As noted in [75], the sum of the AC components of the DCT is equal to the variance of the image intensity and can be used as a focus measure. Later, Shen and Chen proposed in [66] the DC/AC ratio as a focus measure. Let $F_{u,v}$ be the DCT of an $M \times N$ sub-block of the image (typically, $M = N = 8$). The focus measure associated with this sub-block, ϕ_S, can be computed as

$$\phi_S = \frac{\sum_{\substack{u=0 \\ (u,v) \neq (0,0)}}^{M-1} \sum_{v=0}^{N-1} F(u,v)^2}{F_{0,0}^2}. \tag{A.38}$$

For SFF, the focus measure for a pixel $I(x,y)$ is computed by accumulating the values of ϕ_S within its neighborhood $\Omega(x,y)$.

A.31. DCT reduced energy ratio (DCT2)

Lee et al. [63] applied the DCT to 8×8 image sub-blocks in order to measure focus. They suggested that the computation time and robustness to noise of the energy ratio measure in (A.38) can be improved if only 5 out of the 63 AC coefficients are used to compute the AC energy. Thus, the focus measure is defined as

$$\phi = \frac{F_{0,1}^2 + F_{1,0}^2 + F_{2,0}^2 + F_{1,1}^2 + F_{0,2}^2}{F_{0,0}^2}.$$

A.32. Modified DCT (DCT3)

An efficient implementation of a focus measure based on an 8×8 modified DCT can be obtained by performing a linear convolution with a mask $\mathcal{M}$ [76]. Similarly to DCTR and DCTE, the focus measure for SFF is computed for every pixel according to its neighborhood

$$\phi_{x,y} = \sum_{(i,j) \in \Omega(x,y)} (I \ast \mathcal{M}), \tag{A.39}$$

where

$$\mathcal{M} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}.$$

A.33. Sum of wavelet coefficients (WAV1)

Wavelet-based focus measure operators are mostly based on the statistical properties of the discrete wavelet transform (DWT) coefficients. In the first level DWT, the image is decomposed into four sub-images, where W_{LH1}, W_{HL1}, W_{HH1} and W_{LL1} denote the three detail sub-bands and the coarse approximation sub-band, respectively. For a higher level DWT, the coarse approximation is successively decomposed into detail and coarse sub-bands. The information of the detail and coarse sub-bands is then used to compute the focus measure.

Yang and Nelson [68] proposed a focus operator for autofocus computed from the sub-bands

$$\phi = \sum_{(i,j) \in \Omega_D} |W_{LH1}(i,j)| + |W_{HL1}(i,j)| + |W_{HH1}(i,j)|, \tag{A.40}$$

where Ω_D is the corresponding window of Ω in the DWT sub-bands. In this work, the focus measure of all the wavelet-based operators has been computed using the coefficients of the over-complete wavelet transform, thus avoiding the need for computing the corresponding neighborhood within each sub-band. Thus, Ω_D is simply the same as Ω.

Huang et al. [77] used a focus measure similar to (A.40) with 2-level DWT and Daubechies-10 filters. In this work, 1-level DWT with Daubechies-6 filters have been used by following [68] and [70].

A.34. Variance of wavelet coefficients (WAV2)

The variance of the wavelet coefficients within Ω_D can also be used to compute the focus measure [68]

$$\begin{aligned} \phi = &\sum_{(i,j) \in \Omega_D} (W_{LH1}(i,j) - \mu_{LH1})^2 \\ &+ \sum_{(i,j) \in \Omega_D} (W_{HL1}(i,j) - \mu_{HL1})^2 \\ &+ \sum_{(i,j) \in \Omega_D} (W_{HH1}(i,j) - \mu_{LL1})^2, \end{aligned} \tag{A.41}$$

where μ_{LH}, μ_{HL} and μ_{HH} denote the mean value of the respective DWT sub-bands within Ω_D.

A.35. Ratio of wavelet coefficients (WAV3)

Xie et al. [70] proposed the use of the ratio between the high frequency coefficients M_H and the low frequency coefficients M_L of the wavelet transform as a focus measure [70]

$$\phi = \frac{M_H^2}{M_L^2}, \tag{A.42}$$

where M_H and M_L are defined as follows:

$$M_H^2 = \sum_k \sum_{(i,j) \in \Omega_D} W_{LHk}(i,j)^2 + W_{HLk}(i,j)^2 + W_{HHk}(i,j)^2, \tag{A.43}$$

$$M_L^2 = \sum_k \sum_{(i,j) \in \Omega_D} W_{LLk}(i,j)^2. \tag{A.44}$$

Sub-index k indicates that the k-th level wavelet is used to compute the coefficients. According to [70], the coefficients of the first level DWT are used in (A.43), whereas the third level coefficients are used in (A.44). The WAV1, WAV2 and WAV3 operators were originally proposed for autofocus applications. In order to adapt them to SFF, a focus measure is computed for every pixel $I(x,y)$ by restricting the sums in (A.40)–(A.42) to the corresponding $\Omega(x,y)$.

A.36. Ratio of curvelet coefficients

Minhas et al. [78] proposed a focus measure operator based on the coefficients of the discrete curvelet transform. In the k-th level, the curvelet transform decomposes an image into N bands at different orientations. Similarly to the wavelet-based focus measure operators described previously, the focus measure is computed as

$$\phi = \sum_{(i,j) \in \Omega_D} F_\theta(i,j), \tag{A.45}$$

where $F_\theta(i,j)$ is calculated as the ratio between the summed coefficients of the k-th and $(k-1)$-th level sub-bands. Let C_k denote the coefficients of the k-th sub-band, $F_\theta(i,j)$ is defined as

$$F_\theta(i,j) = \frac{\sum C_k(i,j)}{\sum C_{k-1}(i,j)}. \tag{A.46}$$

Following [78], 2-level curvelet decomposition with eight orientations has been implemented. In order to perform a fair comparison with other focus measure operators, the pre-processing steps of contrast enhancement and denoising described in [78] have been omitted.

References

[1] I. Nourbakhsh, D. Andre, C. Tomasi, M. Genesereth, Mobile robot obstacle avoidance via depth from focus, Robotics and Autonomous Systems 22 (1997) 151–158.

[2] H.Y. Lin, M. Subbarao, A vision system for fast 3D model reconstruction, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, 2001, pp. 663–668.

[3] Z. Jiang, D. Xu, M. Tan, H. Xie, An improved focus measure for MEMS assembly, in: Proceedings of the International Conference on Mechatronics and Automation, vol. 2, 2005, pp. 1118–1122.

[4] L. Firestone, K. Cook, K. Culp, N. Talsania, K. Preston Jr., Comparison of autofocus methods for automated microscopy, Cytometry 12 (1991) 195–206.

[5] F.C.A. Groen, I.T. Young, G. Ligthart, A comparison of different focus functions for use in autofocus algorithms, Cytometry 6 (1985) 81–91.

[6] M.J. Russell, T.S. Douglas, Evaluation of autofocus algorithms for tuberculosis microscopy, in: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2007, pp. 3489–3492.

[7] A. Santos, C.O. de Solorzano, J.J. Vaquero, J.M. Pena, N. Mapica, F.D. Pozo, Evaluation of autofocus functions in molecular cytogenetic analysis, Journal of Microscopy 188 (1997) 264–272.

[8] Y. Sun, S. Duthaler, B.J. Nelson, Autofocusing in computer microscopy: selecting the optimal focus algorithm, Microscopy Research and Technique 65 (2004) 139–149.

[9] M. Subbarao, J.-K. Tian, Selecting the optimal focus measure for autofocusing and depth-from-focus, IEEE Transactions on Pattern Analysis and Machine Intelligence 20 (1998) 864–870.

[10] A.S. Malik, T.S. Choi, Consideration of illumination effects and optimization of window size for accurate calculation of depth map for 3d shape recovery, Pattern Recognition 40 (2007) 154–170.

[11] S. Nayar, Y. Nakagawa, Shape from focus, IEEE Transactions on Pattern Analysis and Machine Intelligence 16 (1994) 824–831.

[12] M. Subbarao, T. Choi, Accurate recovery of three-dimensional shape from image focus, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (1995) 266–274.

[13] M. Asif, T.S. Choi, Shape from focus using multilayer feedforward neural networks, IEEE Transactions on Image Processing 10 (2001) 1670–1675.

[14] M. Ahmad, T.S. Choi, A heuristic approach for finding best focused shape, IEEE Transactions on Circuits and Systems for Video Technology 15 (2005) 566–574.

[15] M. Ahmad, T.S. Choi, Shape from focus using optimization technique, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 2, 2006, pp. II–II.

[16] S.O. Shim, T.S. Choi, A novel iterative shape from focus algorithm based on combinatorial optimization, Pattern Recognition 43 (2010) 3338–3347.

[17] M. Muhammad, T.S. Choi, Sampling for shape from focus in optical microscopy, IEEE Transactions on Pattern Analysis and Machine Intelligence 34 (2012) 564–573.

[18] M.T. Mahmood, T.S. Choi, Nonlinear approach for enhancement of image focus volume in shape from focus, IEEE Transactions on Image Processing 21 (2012) 2866–2873.

[19] M. Subbarao, T. Choi, A. Nikzad, Focusing techniques, Journal of Optical Engineering 32 (1993) 2824–2836.

[20] J. Dias, A. de Almeida, H. Araujo, Depth recovery using active focus in robotics, in: Proceedings of the IEEE/RSJ International Workshop on Intelligent Robots and Systems, vol. 1, 1991, pp. 249–255.

[21] E. Wong, A new method for creating a depth map for camera auto focus using an all in focus picture and 2d scale space matching, in: Proceedings of the International Conference on Acoustics, Speech and Signal Processing, vol. 3, 2006.

[22] A.P. Pentland, A new sense for depth of field, IEEE Transactions on Pattern Analysis and Machine Intelligence 9 (1987) 523–531.

[23] M. Subbarao, A. Nikzad, A model for image sensing and digitization in machine vision, in: SPIE Conference, 1990.

[24] P.Y. Maeda, P.B. Catrysse, B.A. Wandell, Integrating lens design with digital camera simulation, in: SPIE Conference, vol. 5678, International Society for Optical Engineering, 2005, pp. 48–58.

[25] Z. Wang, A. Bovik, A universal image quality index, IEEE Signal Processing Letters 9 (2002) 81–84.

[26] H. Xie, W. Rong, L. Sun, Construction and evaluation of a wavelet-based focus measure for microscopy imaging, Microscopy Research and Technique 70 (2007) 987–995.

[27] J.A. Marshall, D. Ariely, C.A. Burbeck, T.D. Aricly, J.P. Rolland, K.E. Martin, Occlusion edge blur: a cue to relative visual depth, International Journal of the Optical Society of America 13 (1996) 681–688.

[28] H. Nair, C. Stewart, Robust focus ranging, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1992, pp. 309–314.

[29] Y. Xiong, S. Shafer, Depth from Focusing and Defocusing, Technical Report CMU-RI-TR-93-07, Robotics Institute, Pittsburgh, PA, 1993.

[30] G. Healey, R. Kondepudy, Radiometric CCD camera calibration and noise estimation, IEEE Transactions on Pattern Analysis and Machine Intelligence 16 (1994) 267–276.

[31] C. Liu, R. Szeliski, S.B. Kang, C. Zitnick, W. Freeman, Automatic estimation and removal of noise from a single image, IEEE Transactions on Pattern Analysis and Machine Intelligence 30 (2008) 299–314.

[32] B.K.P. Horn, Robot Vision, 6th edition, The MIT Press, 1990.

[33] R. Gopinath, J. Odegard, C. Burrus, Optimal wavelet representation of signals and the wavelet sampling theorem, IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing 41 (1994) 262–277.

[34] B. Forster, D. Van De Ville, J. Berent, D. Sage, M. Unser, Complex wavelets for extended depth-of-field: a new method for the fusion of multi-channel microscopy images, Microscopy Research and Technique 65 (2004) 33–42.

[35] W.W. Wang, P.-L. Shui, G.X. Song, Multifocus image fusion in wavelet domain, in: Proceedings of the International Conference on Machine Learning and Cybernetics, vol. 5, 2003, pp. 2887–2890.

[36] D.S. Taubman, M.W. Marcellin, JPEG 2000: Image Compression Fundamentals, Standards and Practice, Kluwer Academic Publishers, 2002.

[37] D.C. Montgomery, G.C. Runger, Applied Statistics and Probability for Engineers, 5th edition, John Willey & Sons, 2010.

[38] L. Zhang, R. Lukac, X. Wu, D. Zhang, PCA-based spatially adaptive denoising of CFA images for single-sensor digital cameras, IEEE Transactions on Image Processing 18 (2009) 797–812.

[39] M. Juneja, P.S. Sandhu, Performance evaluation of edge detection techniques for images in spatial domain, International Journal of Computer Theory and Engineering 1 (2009) 614–621.

[40] S. Chang, B. Yu, M. Vetterli, Adaptive wavelet thresholding for image denoising and compression, IEEE Transactions on Image Processing 9 (2000) 1532–1546.

[41] H. Sundaram, S. Nayar, Are textureless scenes recoverable? in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997, pp. 814–820.

[42] M. Muhammad, H. Mutahira, A. Majid, T.S. Choi, Recovering 3D shape of weak textured surfaces, in: Proceedings of the International Conference on Computational Science and its Applications, 2009, pp. 191–197.

[43] V. Gaganov, A. Ignatenko, Robust shape from focus via Markov random fields, in: Proceedings of the International Conference on Computer Graphics and Vision, 2009, pp. 74–80.

[44] A.R. Rao, G.L. Lohse, Towards a texture naming system: identifying relevant dimensions of texture, Vision Research 36 (1996) 1649–1669.

[45] P. Mendapara, R. Minhas, Q. Wu, Depth map estimation using exponentially decaying focus measure based on SUSAN operator, in: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, 2009, pp. 3705–3708.

[46] M. Mahmood, T.S. Choi, Focus measure based on the energy of high-frequency components in the S transform, Optics Letters 35 (2010) 1272–1274.

[47] A. Baradarani, Q.J. Wu, M. Ahmadi, P. Mendapara, Tunable halfband-pair wavelet filter banks and application to multifocus image fusion, Pattern Recognition 45 (2012) 657–671.

[48] M. Shirvaikar, An optimal measure for camera focus and exposure, in: Proceedings of the Southeastern Symposium on System Theory, 2004, pp. 472–475.

[49] H. Nanda, R. Cutler, Practical Calibrations for a Real-Time Digital Omnidirectional Camera, Technical Report, Technical Sketches, Computer Vision and Pattern Recognition, 2001.

[50] F. Helmli, S. Scherer, Adaptive shape from focus with an error estimation in light microscopy, in: Proceedings of the International Symposium on Image and Signal Processing and Analysis, 2001, pp. 188–193.

[51] R. Minhas, A.A. Mohammed, Q.M. Wu, M.A. Sid Ahmed, 3D shape from focus and depth map computation using steerable filters, in: Proceedings of the International Conference on Image Analysis and Recognition, 2009, Springer-Verlag, Berlin, Heidelberg, 2009, pp. 573–583.

[52] J. Lorenzo, M. Castrillon, J. Mendez, O. Deniz, Exploring the use of local binary patterns as focus measure, in: International Conference on Computational Intelligence for Modelling Control & Automation, 2008, pp. 855–860.

[53] W.T. Freeman, E.H. Adelson, The design and use of steerable filters, IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (1991) 891–906.

[54] R. Minhas, A. Mohammed, Q. Wu, An efficient algorithm for focus measure computation in constant time, IEEE Transactions on Circuits and Systems for Video Technology 22 (2012) 152–156.

[55] W. Huang, Z. Jing, Evaluation of focus measures in multi-focus image fusion, Pattern Recognition Letters 28 (2007) 493–500.

[56] V. Hilsenstein, Robust autofocusing for automated microscopy imaging of fluorescently labelled bacteria, in: Proceedings of the Digital Image Computing: Techniques and Applications, 2005, pp. 15–15.

[57] J. Geusebroek, F. Cornelissen, A. Smeilders, H. Geerts, Robust autofocusing in microscopy, Cytometry 39 (2000) 1–9.

[58] A.S. Malik, T.S. Choi, A novel algorithm for estimation of depth map using image focus for 3D shape recovery in the presence of noise, Pattern Recognition 41 (2008) 2200–2225.

[59] N.K. Chern, P.A. Neow, M.H. Ang, Practical issues in pixel-based autofocusing for machine vision, in: Proceedings of the International Conference on Robotics and Automation, vol. 3, 2001, pp. 2791–2796.

[60] A.M. Eskicioglu, P.S. Fisher, Image quality measures and their performance, IEEE Transactions on Communications 43 (1995) 2959–2965.

[61] M. Ahmad, T.S. Choi, Application of three dimensional shape from image focus in LCD/TFT displays manufacturing, IEEE Transactions on Consumer Electronics 53 (2007) 1–4.

[62] E. Krotkov, J.P. Martin, Range from focus, in Proceedings of the International Conference on Robotics and Automation, vol. 3, 1986, pp. 1093–1098.

[63] S.Y. Lee, J.T. Yoo, Y. Kumar, S.W. Kim, Reduced energy-ratio measure for robust autofocusing in digital camera, IEEE Signal Processing Letters 16 (2009) 133–136.

[64] J. Lee, K. Kim, B. Nam, J. Lee, Y. Kwon, H. Kim, Implementation of a passive automatic focusing algorithm for digital still camera, IEEE Transactions on Consumer Electronics 41 (1995) 449–454.

[65] J. Pech Pacheco, G. Cristobal, J. Chamorro Martinez, J. Fernandez Valdivia, Diatom autofocusing in brightfield microscopy: a comparative study, in:

[66] C.H. Shen, H.H. Chen, Robust focus measure for low-contrast images, Digest of Technical Papers of International Conference on Consumer Electronics, 2006, pp. 69–70.

[67] C. Wee, R. Paramesran, Measure of image sharpness using eigenvalues, Information Sciences 177 (2007) 2533–2552.

[68] G. Yang, B. Nelson, Wavelet-based autofocusing and unsupervised segmentation of microscopic images, in: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 3, 2003, pp. 2143–2148.

[69] P. Yap, P. Raveendran, Image focus measure based on Chebyshev moments, IEEE Proceedings—Vision, Image and Signal Processing 151 (2004) 128–136.

[70] H. Xie, W. Rong, L. Sun, Wavelet-based focus measure and 3-d surface reconstruction method for microscopy images, in: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, 2006, pp. 229–234.

[71] A. Thelen, S. Frey, S. Hirsch, P. Hering, Improvements in shape-from-focus for holographic reconstructions with regard to focus operators, neighborhood-size, and height value interpolation, IEEE Transactions on Image Processing 18 (2009) 151–157.

[72] Y. An, G. Kang, I.J. Kim, H.S. Chung, J. Park, Shape from focus through Laplacian using 3D window, in: International Conference on Future Generation Communication and Networking, vol. 2, 2008, pp. 46–50.

[73] C.Y. Wee, R. Paramesran, Comparative analysis of eigenvalues-based and tchebichef moments-based focus measures, in: International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, vol. 1, 2008, pp. 517–520.

[74] C. Wee, R. Paramesran, Image sharpness measure using eigenvalues, in: Proceedings of the International Conference on Signal Processing, 2008, pp. 840–843.

[75] J. Baina, J. Dublet, Automatic focus and iris control for video cameras, in: International Conference on Image Processing and its Application, 1995, pp. 232–235.

[76] S.Y. Lee, Y. Kumar, J.M. Cho, S.W. Lee, S.W. Kim, Enhanced autofocus algorithm using robust focus measure and fuzzy reasoning, IEEE Transactions on Circuits and Systems for Video Technology 18 (2008) 1237–1246.

[77] J. Huang, C. Shen, S. Phoong, H. Chen, Robust measure of image focus in the wavelet domain, in: Proceedings of the International Symposium on Intelligent Signal Processing and Communication Systems, 2005, pp. 157–160.

[78] R. Minhas, A.A. Mohammed, Q.J. Wu, Shape from focus using fast discrete curvelet transform, Pattern Recognition 44 (2011) 839–853.

Said Pertuz was born is Sincelejo, Colombia. He received the B.S. degree in Electronics Engineering at the Industrial University of Santander, Colombia (2007) and the M.S. degree in Computer Science from the Rovira i Virgily University, Spain (2009). In 2008, he joined the Intelligent Robotics and Computer Vision group of the Rovira Virgili University as Ph.D. student and his current research interest include robotics vision and 3D scene reconstruction.

Domenec Puig received the M.S. and Ph.D. degrees in computer science from Polytechnic University of Catalonia, Barcelona, Spain, in 1992 and 2004 respectively. In 1992, he joined the Department of Computer Science and Mathematics at Rovira i Virgili University, Tarragona, Spain, where he is currently Associate Professor. Since July 2006, he is the Head of the Intelligent Robotics and Computer Vision group at the same university. His research interests include image processing, texture analysis, perceptual models for image analysis, scene analysis, and mobile robotics.

Miguel Angel Garcia received the B.S., M.S., and Ph.D. degrees in computer science from Polytechnic University of Catalonia, Barcelona, Spain, in 1989, 1991, and 1996 respectively. He joined the Department of Software at Polytechnic University of Catalonia in 1996 as an Assistant Professor. From 1997 to 2006, he was with the Department of Computer Science and Mathematics at Rovira i Virgili University, Tarragona, Spain, where he was the Head of Intelligent Robotics and Computer Vision group. In 2006, he joined the Autonomous University of Madrid, Spain, where he is currently Associate Professor. His research interests include mobile robotics, image processing, and 3D modeling.