COMPARISON OF FOCUS MEASURES IN FACE DETECTION ENVIRONMENTS

J. Lorenzo, O. Déniz, M. Castrillón and C. Guerra

Inst. of Intelligent Systems and Numerical Appl. in Engineering (IUSIANI)
University of Las Palmas de Gran Canaria, Campus Univ. de Tafira, 35017 LAs Palmas, Spain
Campus Univ. de Tafira, 35017 LAs Palmas, Spain
[jlorenzo,odeniz,mcastrillon,cguerra]@iusiani.ulpgc.es

Abstract: This work presents a comparison among different focus measures used in the literature for autofocusing in a non previously explored application of face detection. This application has different characteristics to those where traditionally autofocus methods have been applied like microscopy or depth from focus. The aim of the work is to find if the best focus measures in traditional applications of autofocus have the same performance in face detection applications. To do that six focus measures has been studied in four different settings from the oldest to more recent ones.

1 Introduction

In face detection and face recognition methods (Pentland et al., 1994; Rowley et al., 1998; Gross et al., 2001; Hjelmas and Low, 2001; Yang et al., 2002; Zhao et al., 2003), borders play an important role because they define the facial features that appear in the face such as eyes, mouth and nose, which are needed to carry out the task. In a blurred image these facial features are not well defined, and so the detection or identification can not be done. Blurred images can be obtained in a defocused camera because defocusing can be modelled as a low-pass filtering process, opposite to focused images which have a higher frequency content (Nayar, 1994). Thus, it is desirable that the image acquisition system has an autofocus mechanism.

Automatic focusing methods fall into two main categories: active and passive systems. Active methods are based on emiting a wave in order to estimate the distance to the object of interest and consequently adjust the lens position. Passive methods estimate the position of the lens by means of finding the position where image sharpness is maximum. Passive focusing methods are normally found in low cost consumer digital cameras, while active focus is widely used in professional digital cameras and computer vision applications such as microscopy or inspection.

As mentioned above, passive autofocusing methods adjust the focus lens to maximize the high frequency components in the image. Therefore a focus value is needed to measure the amount of high frequencies in the image. As autofocusing is a long-standing topic, in the literature a wide variety of focus algorithms have been proposed each one with an associated focus measure (Krotkov, 1987; Nayar, 1994; Lee et al., 1995; Subbarao and Tyan, 1998; Choi and Ki, 1999; Lee et al., 2001; Nathaniel et al., 2001; Kehtarnavaz and Oh, 2003; Kristan and Pernus, 2004; Shirvaikar, 2004; Park and Kim, 2005; Kristan et al., 2006).

Most of the published autofocusing algorithms solve the problem of planar objects like in microscopy applications (Sun et al., 2004) or single object of interest like depth from focus applications (Nayar, 1994). In these applications, focus measures exhibit an ideal curve with a peak with step slopes at the lens position where the object is focused because there is only one object in the image or because it is a planar image. However, in human computer interaction, people do not always hold the same position in the image and exists more objects in the scene so the focus measure does not exhibit a clear maximum. In digital photograhy this drawback is eluded because the photographer selects the object of interest and centers it.

In this work we analyze the performance of some focus measures in different scenarios including typical human computer interaction. These measures are explained in Section 2. Section 3 describes the experiments and finally in Section 4 the results achieved are commented.

2 Focus Algorithms

As explained in the previous section, many focus measures have been proposed in the last years to solve the autofocus problem. All of them rely on the fact that a focused image has high contents of higher frequencies so any measure which computes these frequencies can be used. In this work, six of these measures have been chosen to make the comparison. We have compared well known focus measures with more recent ones. Below, we briefly describe each one.

The Tenenbaum Gradient (Tenengrad) (Krotkov, 1987) was one of the first proposed focus measures. This measure convolves the image with vertical (S_x) and horizontal (S_y) Sobel operators. To get a global measure over the whole image, the square of the gradient vector components are summed.

$$F_{Tenengrad} = \sum \sum S_x(x,y)^2 + S_y(x,y)^2 \quad (1)$$

The entropy measure proposed by Firestone et al. (Firestone et al., 1991) is based on the idea that the histogram of a focused image contains more information than the histogram of a defocused one. In this measure the histogram is normalized to get the probability $p(i)$ for each gray level i.

$$F_{Entropy} = - \sum_{intensities} p(i) \log p(i) \quad (2)$$

The Sum of Modified Laplace (SML) (Nayar, 1994) is based on the linear differential operator Laplacian which has the same properties in all directions and is therefore invariant to rotation. Thus, the SML measure sums the absolute values of the convolution of the image with the Laplacian operators.

$$F_{SML} = \sum \sum |L_x(x,y)| + |L_y(x,y)| \quad (3)$$

Energy Laplace (Subbarao and Tyan, 1998) is based on the same idea of the SML mesasure but the image is convolved with the following mask,

$$L = \begin{bmatrix} -1 & -4 & -1 \\ -4 & 20 & -4 \\ -1 & -4 & -1 \end{bmatrix}$$

which computes the second derivate $D(x,y)$ of the image. The value of the focus measure is the sum of the squares of the convolution results.

$$F_{EnergyLaplace} = \sum \sum D(x,y)^2 \quad (4)$$

Figure 1: Printed circuit board image

Figure 2: Box picture

Nanda and Cutler (Nanda and Cutler, 2001) proposed a focus measure from the contrast of a image as the absolute difference of a pixel with its eight neighbors, summed over all the pixels of the image.

$$F_{Contrast} = \sum \sum C(x,y) \quad (5)$$

where the contrast $C(x,y)$ for each pixel in the gray image $I(x,y)$ is computed as

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

Kristan et al. (Kristan et al., 2006) described M_{Be}, which is one of the most recent focus measures. It is based on the coefficients of the discrete cosine transform obtained after dividing the image into 8x8 non overlapped windows and then averaging over all the 8x8 windows. It must be noticed that in our implementation we have no filtered components corresponding to high order frequencies as Kristan proposes.

$$F_{MBe} = \frac{\sum M'_{Be}}{\text{num. of 8x8 windows}} \quad (6)$$

where M'_{Be} is computed from the DCT coefficients $F(\omega, \nu)$ as

$$M'_{Be} = 1 - \frac{\sum |F(\omega,\nu)|^2}{(\sum |F(\omega,\nu)|)^2}$$

Figure 3: First human computer interaction scenario

Figure 4: Second human computer interaction scenario

Figure 5: Normalized curves for the Tenengrad focus measure
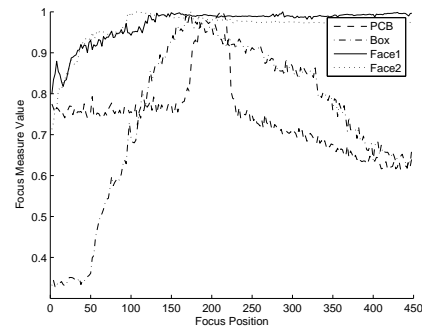
Figure 6: Normalized curves for the entropy focus measure

3 Experimental Study

The images were acquired with a Sony DFW-VL500 firewire color camera with a 16x integrated zoom in an indoor environment and four different settings were analyzed. The first setting corresponds to a printed circuit board which yields a planar image with all the scene elements to the same distance of the camera; we refer to this setting as PCB (Fig. 1). The second setting is very common in depth from focus applications where an isolated object appears in the image, this setting will be refered as Box (Fig. 2). The third and fourth settings are the ones that we typically found in a human computer interaction application where a person appears either in front of the camera or in an office enviroment. They will be refered as Face1 (Fig. 3) and Face2 (Fig. 4).

As the camera has 450 focus positions, 224 images for each of the previosly described settings were acquired with a 2 focus position step. For each acquired image the six focus measures were computed and the criterium to assess the quality of each measure was the similarity of the resulting curve with an "ideal" focus curve which exhibits only a sharp peak.

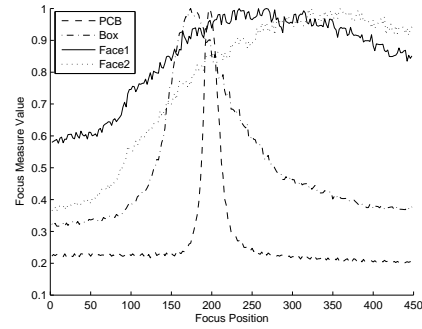Figure 5 shows the normalized curves of the Tenengrad focus measure for the four examples. This measure gives good results for PCB and Box examples because in both cases there is a well defined maximum at positions 198 and 176 respectively. On the other hand in the examples Face1 and Face2, the obtained curve does not show a sharp peak so the maximum search is more difficult.

Figure 6 shows the normalized curves of the entropy focus measure for the four examples. As it is shown in the graphics, the behaviour of this measure is not so good as Tenengrad measure. Entropy based measure only gives a good focus curve for the PCB example with the maximum located at position 216. For Face1 and Face2 examples the curve increases its value until it reaches a plateau where a maximum is really difficult to find.

The results of the SML measure in the four examples are shown in Figure 7. This measure exhibits a well defined peak in PCB and Box with maxima at positions 200 and 180 respectively. In relation to Face1 example, we get a curve with with a maximum at 246 although the peak is not so sharp as in PCB and Box examples. In example Face2 the resulting curve for the SML measure has a flattened peak with the maximum located at 368.

Table 1: Maximum of each focus measure

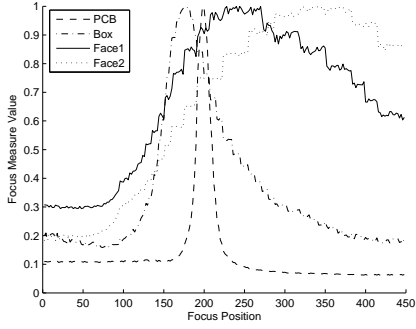| Example	$F_{Tenengrad}$	$F_{Entropy}$	F_{SML}	$F_{EnergyLaplace}$	$F_{Constrast}$	$F_{M_{Be}}$	Expert
PCB	200	216	200	200	200	200	202
Box	176	176	180	176	280	180	178
Face1	270	290	246	240	272	264	240
Face2	362	108	368	362	334	284	330-370

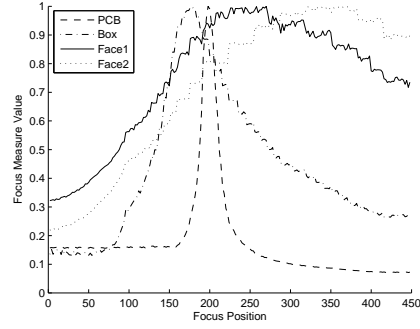Figure 7: Normalized curves for the Sum of Modified Laplacian focus measure

Figure 9: Normalized curves for the focus measure proposed by Nanda and Cutler
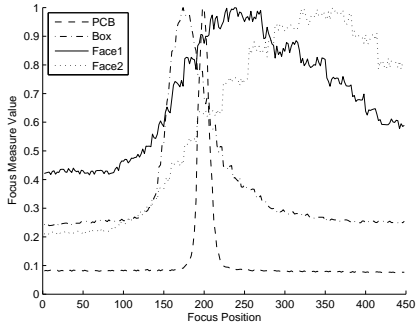

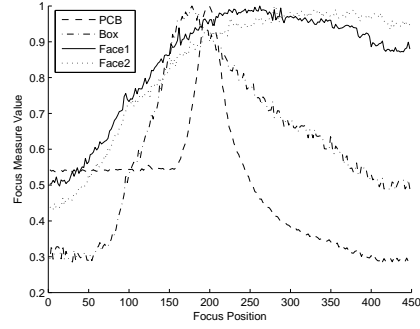Figure 8: Normalized curves for the Energy Laplace focus measure

Figure 10: Normalized curves for the M_{Be} focus measure

Energy Laplace and SML measures are based on the Laplacian of the image because it gives a high response to the higher frequencies in the image. Thus the results we get with Energy Laplace are very similar to those obtained with SML measure as it is shown in Figure 8. For PCB example the position of the maximum is the same than in the SML measure. In Box example the maximum is at position 176 while for the SML is in position 180, which are very close. The difference in focus position of this measure and SML in examples Face1 and Face2 is 6 steps in both cases that confirms the similar behavior of both measures.

As shown in Figure 9, the measure proposed by Nanda and Cutler gives similar results to those ob-

tained with the previously analyzed measures (except entropy) in examples PCB and Box. For Face1 and Face2, the results are similar too because in neither cases a sharpened peak appears at position of maximum which is located at 272 and 334 respectively.

Finally, the results obtained with the M_{Be} measure are shown in Figure 10. As one of the most recent focus measures it is expected to achieve the best results in the four settings. However, for the PCB and Box examples, the obtained results are very similar to the other measures even worse than more classical measures as SML or Tenengrad. In Face1 and Face2, the results are not so good as other measures because the curves does not exhibit a sharp peak in the maximum value so more elaborated methods for

the maximum location are needed.

The previous analisys only takes into account the shape of the focus curve but to assess the quality of the lens position which corresponds to the maximum it need to be tested by an expert because there is no possibility to compare with a reference focus measure. Sun et al. (Sun et al., 2004) propose as reference to test the accuracy of the compared methods the difference between the lens position given by the method and the lens position selected by an expert.

In this work we have followed a similar approach and for each scenario an expert was asked for getting the most focused image. The range of focus value are shown in the column labelled as `Expert` of table 1. To notice that for the `Face2` scenario the range of focus values in which the image is focused is wider because the person is further and so the depth of field is larger. Comparing in Table 1, the best focus value given by each measure and the one selected by the expert, it is observed that the most accurate measures are those based on Laplacian as SML and Energy Laplace. The others exhibit a similar accuracy except Entropy that as in the previous analysis about the shape of the curve exhibits the worst accuracy.

4 Conclusions

In this work a comparison of six focus measures have been carried out to investigate the performance of the measures in a face detection application. In face detection applications the person, which is the object of interest, normally is in an office environment so the obtained curves do not exhibit a sharp peak at one defined focus position. Instead, flattened peaks are obtained which make more difficult to get the best focus position. From the six compared focus measures, all of them, except the entropy measure, give very similar results in non face detection applications. In the two face detection settings the best results were obtained with SML and Energy Laplace measures and surprisingly the most recently published measure does not give as good results as previous ones. So we have concluded that for face detection applications the best performance is obtained with Laplacian based measures but it is necessary to use more elaborated maximum finding methods because there does not exist very sharp peaks in the focus curves. Also, a test about the accuracy of the focus position was carried out, using as reference the focus position given by an expert for each example. The results are very similar to the previous given as the two most accurate measures those based on Laplacian and the worst accuracy the Entropy measure.

REFERENCES

Choi, K.-S. and Ki, S.-J. (1999). New autofocusing technique using the frequency selective weighthed median filter for video cameras. *IEEE Transactions on Consumer Electronics*, 45(3):820–827.

Firestone, L., Cook, K., Culp, K., Talsania, N., and Preston, K. (1991). Comparison of autofocus methods for automated microscopy. *Cytometry*, 12:195–206.

Gross, R., Shi, J., and Cohn, J. (2001). Quo vadis face recognition? - the current state of the art in face recognition. Technical Report CMU-RI-TR-01-17, Robotics Institute, Carnegie Mellon University.

Hjelmas, E. and Low, B. K. (2001). Face detection: A survey. *Computer Vision and Image Understanding*, 83(3):236–274.

Kehtarnavaz, N. and Oh, H.-J. (2003). Development and real-time implementation of a rule-based auto-focus algorithm. *Real-Time Imaging*, 9:197–203.

Kristan, M. and Pernus, F. (2004). Entropy based measure of camera focus. In *Proceedings of the thirteenth Electrotechnical and Computer Science Conference ERK*, pages 179–182.

Kristan, M., Pers, J., Perse, M., and Kovacic, S. (2006). A bayes-spectral-entropy-based measure of camera focus using a discrete cosine transform. *Pattern Recognition Letters*, 27(13):1419–1580.

Krotkov, E. (1987). Focusing. *International Journal of Computer Vision*, 1:223–237.

Lee, J.-H., Kim, K.-S., and Nam, B.-D. (1995). Implementation of a passive automatic focusing algorithm for digital still camera. *IEEE Transactions on Consumer Electronics*, 41(3):449–454.

Lee, J.-S., Jung, Y.-Y., Kim, B.-S., and Sung-Jea, K. (2001). An advanced video camera system with robust AF,AE and AWB control. *IEEE Transactions on Consumer Electronics*, 47(3):694–699.

Nanda, H. and Cutler, R. (2001). Practical calibrations for a real-time digital onmidirectional camera. In *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR 2001)*.

Nathaniel, N., Neow, P., and Ang, M. (2001). Practical issues in pixel-based autofocusing for machine vision. In *Proc. of the 2001 IEEE International Conference on Robotics and Automation*, Seoul, Korea.

Nayar, Shree K.and Nakagawa, Y. (1994). Shape from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(8):824–831.

Park, R. K. and Kim, J. (2005). A real-time focusing algorithms for iris camera recognition. *IEEE Transactions on Systems, Man and Cybernetics*, 35(3):441–444.

Pentland, A., Moghaddam, B., and Starner, T. (1994). View-based and modular eigenspaces for face recognition. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'94)*, Seattle, WA.

Rowley, H. A., Baluja, S., and Kanade, T. (1998). Neural network-based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1):23–38.

Shirvaikar, M. (2004). An optimal measure for camera focus and exposure. In *Proceedings of the Thirty-Sixth Southeastern Symposium on System Theory*, pages 472–475.

Subbarao, M. and Tyan, J.-K. (1998). Selecting the optimal focus measure for autofocusing and depth-from-focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8):864–870.

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

Yang, M., Kriegman, D., and Ahuja, N. (2002). Detecting faces in images: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1):34–58.

Zhao, W., Chellappa, R., Rosenfeld, A., and Phillips, P. (2003). Face recognition: A literature survey. *ACM Computing Surveys*, pages 399–458.