

Object Recognition Using Frequency Domain Blur Invariant Features

Ville Ojansivu and Janne Heikkilä

Machine Vision Group, Department of Electrical and Information Engineering,
University of Oulu, PO Box 4500, 90014, Finland
{vpo,jth}@ee.oulu.fi

Abstract. In this paper, we propose novel blur invariant features for the recognition of objects in images. The features are computed either using the phase-only spectrum or bispectrum of the images and are invariant to centrally symmetric blur, such as linear motion or defocus blur as well as linear illumination changes. The features based on the bispectrum are also invariant to translation, and according to our knowledge they are the only combined blur-translation invariants in the frequency domain. We have compared our features to the blur invariants based on image moments in simulated and real experiments. The results show that our features can recognize blurred images better and, in a practical situation, they are faster to compute using FFT.

1 Introduction

Recognition of objects and patterns in images is a fundamental part of computer vision with numerous applications. The task is difficult as the objects rarely look exactly similar in different conditions. In real applications, images contain various artifacts such as geometrical and convolutional degradations. Image analysis systems should be able to operate also in these nonideal conditions. There has been a vast amount of research in this field of invariant pattern and object recognition [1]. However, the invariant recognition of objects degraded by blur is a much less studied topic.

In various applications, images may contain blur, which can result, for example, from atmospheric turbulence, out-of-focus, or relative motion between the camera and the scene. This degradation process can be modeled as a linear shift-invariant system in which the relation between an ideal image $f(\mathbf{x})$ and an observed image $g(\mathbf{x})$ is given by

$$g(\mathbf{x}) = f(\mathbf{x}) * h(\mathbf{x}) + n(\mathbf{x}) , \quad (1)$$

where \mathbf{x} is a 2-D spatial coordinate vector, $h(\mathbf{x})$ the point spread function (PSF) of the system, $n(\mathbf{x})$ additive noise, and $*$ denotes 2-D convolution. The point spread function $h(\mathbf{x})$ represents blur while other degradations are captured by the noise term $n(\mathbf{x})$.

The analysis of blurred images is often carried out by first deblurring the images and then applying standard methods for further analysis. Unfortunately,

image deblurring is a difficult problem. Conventional solutions include the estimation of the PSF of the blur and deconvolution of the image using that PSF. When the PSF is known, the latter ill-posed problem can be solved using approaches which use regularization [2]. In practice, the PSF is often unknown and very hard to estimate accurately. In this case, blind image restoration algorithms are used [3].

The analysis of the blurred images can also be performed without deblurring using features which are invariant to blur. Flusser and Suk proposed the first blur invariant features in [4]. These invariants are based on geometric moments, central moments (MOMIs) or the spectrum (SPEIs) of the image. Of these, the MOMIs are also invariant to translation. The MOMIs have been applied to template matching [4], recognition of defocused objects [5] and registration of X-ray images [6]. In [7], blur, rotation and scale invariants based on complex moments were proposed. In [8,9], the theory was extended to blur and affine moment invariants. A shortcoming of these blur invariant features is their sensitivity to noise, especially in the case of SPEIs [10]. Probably for this reason, the SPEIs do not have known applications. In addition, translation invariance has not been incorporated into Fourier domain blur invariants.

The features presented in this paper are invariant to centrally symmetric blur, which is exactly the same condition as with the MOMIs and SPEIs. The invariants are computed from the phase-only spectrum, (phase blur invariants, PBIs) or using phase-only bispectrum in which case we achieve blur-translation invariants (PBTIs). The computation of the invariants can be done efficiently using FFT.

2 Frequency Domain Blur Invariant Features

2.1 Features Invariant to Blur

In this section, we show how invariance to blur is obtained in the Fourier domain by using the phase-only spectrum.

If noise $n(\mathbf{x})$ is neglected, (1) can be expressed in the Fourier domain using the convolution theorem by

$$G(\mathbf{u}) = F(\mathbf{u}) \cdot H(\mathbf{u}) , \quad (2)$$

and in the phasor form by

$$G(\mathbf{u}) = |G(\mathbf{u})| e^{-i\phi_g(\mathbf{u})} , \quad (3)$$

where \mathbf{u} is a vector in the 2-D frequency space.

If the Fourier transform $G(\mathbf{u})$ is normalized by its magnitude, only the complex exponential containing the phase remains, namely

$$\frac{G(\mathbf{u})}{|G(\mathbf{u})|} = e^{-i\phi_g(\mathbf{u})} = e^{-i[\phi_f(\mathbf{u}) + \phi_h(\mathbf{u})]} , \quad (4)$$

where $\phi_f(\mathbf{u})$ is the phase of the original image $f(\mathbf{x})$ and $\phi_h(\mathbf{u})$ the phase of the blur PSF $h(\mathbf{x})$.

Since $h(\mathbf{x})$ is assumed to be centrally symmetric, its Fourier transform $H(\mathbf{u})$ is real and its phase $\phi_h(\mathbf{u})$ has only two possible values

$$\phi_h(\mathbf{u}) = 0 \vee \phi_h(\mathbf{u}) = \pi . \tag{5}$$

It follows from this and from the periodicity of the complex argument that the equality

$$\begin{aligned} [e^{-i\phi_g(\mathbf{u})}]^{2n} &= e^{-i2n\phi_g(\mathbf{u})} \\ &= e^{-i2n\phi_f(\mathbf{u})} e^{-i2n\phi_h(\mathbf{u})} \\ &= [e^{-i\phi_f(\mathbf{u})}]^{2n} \end{aligned} \tag{6}$$

holds for any integer n .

Thus, any even power of the normalized Fourier transform, i.e. $e^{-i2n\phi(\mathbf{u})}$, of the observed image is invariant to the convolution of the original image with any centrally symmetric PSF. In other words, any even multiple of the Fourier transform phase modulo 2π is also invariant. We construct the invariants in this way using value $n=1$, namely

$$\mathcal{B}(\mathbf{u}) = 2\phi(\mathbf{u}) \bmod 2\pi , \tag{7}$$

where $\phi(\mathbf{u})$ is the phase spectrum of the image. Henceforth, $\mathcal{B}(\mathbf{u})$ is called phase blur invariant (PBI). Some similarities with the derivation of the SPEIs in [4] can be observed. In this case, the invariance is obtained by using the tangent of the phase spectrum $\phi(\mathbf{u})$.

2.2 Features Invariant to Blur and Translation

In this section, we extent the theory presented in Section 2.1 to incorporate also invariance to translation.

The phase spectrum of an image, which is used to construct the invariants in Section 2.1, is not invariant to translation. The amplitude spectrum is invariant to translation, but, on the other hand, blur invariants can not be constructed from it. This is the reason why we have to turn to the higher order spectra defined by

$$\Psi_n(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n) = F^*(\mathbf{s}) \prod_{i=1}^n F(\mathbf{u}_i) , \tag{8}$$

where \mathbf{u}_i with $i = 1, \dots, n$ are vectors in the 2-D frequency space, and $\mathbf{s} = \mathbf{u}_1 + \mathbf{u}_2 + \dots + \mathbf{u}_n$. It can be easily shown that Ψ_n is shift invariant [11].

When $n = 1$ in (8) we get the power spectrum and further with the value $n = 2$ the bispectrum, namely

$$\Psi_2(\mathbf{u}_1, \mathbf{u}_2) = F(\mathbf{u}_1)F(\mathbf{u}_2)F(\mathbf{u}_2 + \mathbf{u}_2) . \tag{9}$$

We are interested in the bispectrum as besides its invariance with respect to translation it does not lose any essential information about the original image. This means that it also retains the phase information in contrast to the power spectrum. For these reasons, it is possible to construct a blur invariant phase-only bispectrum where the exponentials containing the phase are raised to the second power, similar to (6), namely

$$\begin{aligned}
 P(\mathbf{u}_1, \mathbf{u}_2) &= e^{-i2\phi(\mathbf{u}_1)} e^{-i2\phi(\mathbf{u}_2)} e^{-i2\phi(\mathbf{u}_1+\mathbf{u}_2)} \\
 &= e^{-i2[\phi(\mathbf{u}_1)+\phi(\mathbf{u}_2)+\phi(\mathbf{u}_1+\mathbf{u}_2)]} .
 \end{aligned}
 \tag{10}$$

By looking the equation (9), we see that the bispectrum is a function of two vector arguments, containing totally four scalar variables. Assuming that $F(\mathbf{u})$ is an N -by- N discrete Fourier transform (DFT) of an image $f(\mathbf{x})$ the bispectrum becomes a four-dimensional N -by- N -by- N -by- N matrix. Fortunately, it is not necessary to evaluate the whole bispectrum. It is possible to take only 2-D slices of the original bispectrum, which contain basically the same information. There are various ways of defining the slices [12,13]. We define the slices as

$$S_k(\mathbf{u}) = \Psi_2(\mathbf{u}, k\mathbf{u}) \quad \forall k \in \mathcal{R} .
 \tag{11}$$

We then form blur invariants corresponding to (10) using only one slice (11) with value $k = 1$ of the whole bispectrum (9), namely

$$\begin{aligned}
 P'(\mathbf{u}) &= e^{-i2[\phi(\mathbf{u})+\phi(\mathbf{u})+\phi(\mathbf{u}+\mathbf{u})]} \\
 &= e^{-i2[2\phi(\mathbf{u})+\phi(2\mathbf{u})]} .
 \end{aligned}
 \tag{12}$$

Finally, we construct the invariants with respect to blur and translation used throughout this paper, similarly to (7), as

$$\mathcal{T}(\mathbf{u}) = 2[2\phi(\mathbf{u}) + \phi(2\mathbf{u})] \bmod 2\pi .
 \tag{13}$$

These invariants are called phase blur-translation invariants (PBTIs).

It would also be possible to use more slices in building up the invariants. It can be seen from (9), (10) and (11) that we need frequency samples in points \mathbf{u} but also at $k\mathbf{u}$ and $(k+1)\mathbf{u}$ to compute an arbitrary slice. If we assume that the images are discrete and that the spectrum is computed using DFT, the samples in the two latter cases can be extracted from DFT by utilizing its conjugate symmetry and periodicity. In our case, we only need frequency samples from points \mathbf{u} and $2\mathbf{u}$.

If the DFT size is N -by- N we have approximately $N/2$ non-redundant invariants available at once. This is opposite to the MOMIs, of which computation time depends on the number of invariants used. However, the invariants corresponding to the lower frequencies have higher signal-to-noise ratio (SNR), and there is some optimal number of invariants that should be used depending on the noise level. In the experiments, we have used L invariants of (7) or (13) for which

$\sqrt{u_1^2 + u_2^2} \leq r$, when $\mathbf{u} = [u_1, u_2]$, but without using the conjugate symmetric or zero frequency components. Other possibility would be to weight the invariants according to SNR.

The distance between images $f(\mathbf{x})$ and $g(\mathbf{x})$ to be classified is computed as

$$D = \sqrt{\sum_{i=1}^L e_i^2}, \quad (14)$$

where

$$e_i = \min \left\{ |I_i^{(f)} - I_i^{(g)}|, |I_i^{(f)} - I_i^{(g)}| - 2\pi \right\}. \quad (15)$$

Here, I_i , $i = 1, \dots, L$, are the invariants of (7) or (13).

It should be noted that the composition of our frequency domain invariants is quite robust to noise in contrast to the SPEIs of [4], which are based on the discontinuous tangent of the phase. Another useful property of our invariants, which results from the normalization of the amplitude information, is the invariance to uniform illumination changes.

As mentioned earlier, our invariants, as well as the moment invariants, are invariant to centrally symmetric blur. However, this is not exactly true for arbitrary images as the blur exchanges the information across the boundaries of the image causing some error. Also the translation invariance retains only if there is no information flow across the image boundaries.

3 Experiments

In the experiments, we classified blurred and noisy images using nearest neighbor classification and compared the results of our blur invariant features to the MOMIs and SPEIs, which were implemented as proposed in [4]. For all the frequency domain invariants $r = \sqrt{5}$ or $\sqrt{10}$, and the order of the MOMIs is up to 5 or 7. This results to either $L=10$ or 18 invariants.

In the first experiment, we compared the features invariant only to blur. In comparison we used the invariants based on central moments (MOMIs), which are also translation invariant, as they seemed to perform better compared to the invariants based on ordinary moments. As test images, we had 40 computer generated uniformly distributed noise images, which were filtered using a Gaussian low pass filter of size 10-by-10 with the standard deviation $\sigma=1$ to get an image, as in Figure 1(a), that resembles some natural texture. We picked one image at time, blurred and added noise to it, as in Figure 1(b), and tried to classify it as one of the original 40 image categories using the invariants. The blur was circular with a radius varying from 0 to 10 pixels with steps of 2 pixels. This kind of blur is a simple model of the out-of-focus effect found in many imaging systems [2]. The image size was cropped finally to 80-by-80 containing only the valid part of the convolutions. The experiment was repeated 20 times for each blur size and for each of the 40 images.

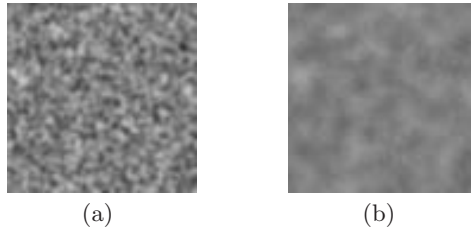


Fig. 1. (a) An example of the 40 filtered noise images used in the first experiment, and (b) a degraded version of it (blur radius is 5 and PSNR 30 dB)

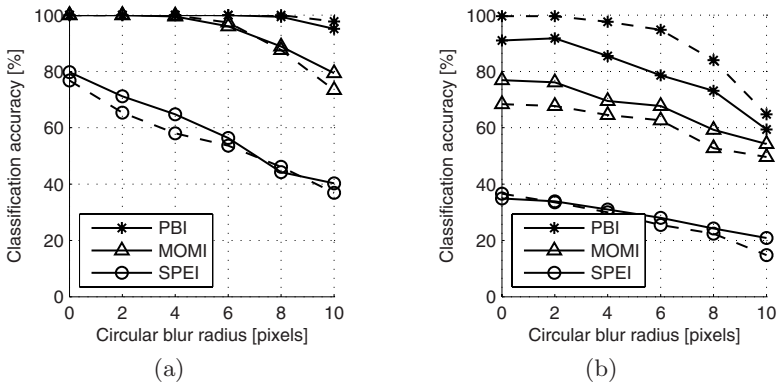


Fig. 2. Classification accuracy of nearest neighbor classification of blurred and noisy images when PSNR is (a) 40 dB (b) 20 dB. The solid lines show the cases where 10 invariants are used and the dashed lines the case of 18 invariants.

The percentage of correct classification for the three methods, our PBIs, the MOMIs and the SPEIs, is shown in Figures 2(a) and 2(b) when the peak signal-to-noise ratio (PSNR) is 40 and 20 dB, respectively. The solid lines show the case of 10 invariants and the dashed lines the case in which 18 invariants are used.

In theory, all the methods are invariant to blur, but in practice, there are differences in their robustness to noise and boundary error. The results show that the SPEIs perform worst as was expected because of their unstable construction. If compared to the PBIs, also the MOMIs seem to be more sensitive at least to noise. The use of 18 invariants seems to give significantly better results compared to the use of 10 invariants only for the PBIs. In contrast, the result of the SPEIs and the MOMIs mainly degrades when 18 invariants are used. This means that the higher order MOMIs and the high frequency SPEIs are already too sensitive to perturbations in these conditions.

The second experiment mimics the situation in which there is an object to be classified on a uniform background. Here we compared the blur-translation invariants, the PBTIs and the MOMIs. We created 20 60-by-60 images, which

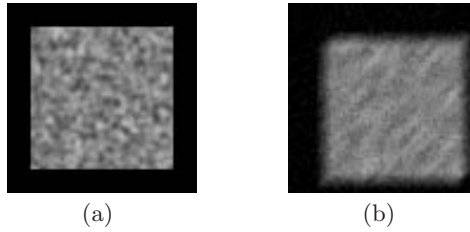


Fig. 3. (a) An Example image used in the second experiment containing an artificial object (filtered noise). (b) A motion blurred and noisy version of the same image (blur length 10 and PSNR 30 dB).

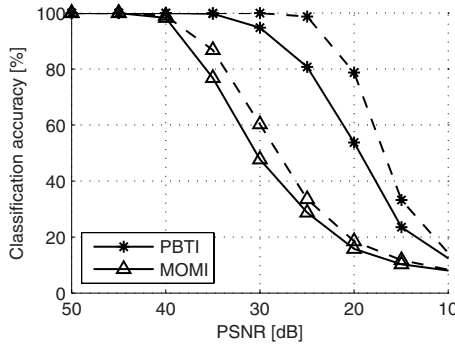


Fig. 4. The object classification accuracy of the blur-translation invariant methods. The solid lines show the cases where 10 invariants are used and the dashed lines the case of 18 invariants.

were similar to the filtered noise images of the first experiment, on a black 80-by-80 background depicting an artificial object, as shown in Figure 3(a). Distorted versions of these images were then generated, which were to be classified as one of the originals using each type of invariants and the nearest neighbor rule. The distortion included random displacement of the object in the range $[-5,5]$ pixels using linear interpolation, motion blur with the length of 10 pixels in a random direction and additive Gaussian noise (also on the background) with PSNR varying from 50 dB to 10 dB in steps of 5 dB. Figure 3(b) shows a degraded object. The border effect does not distort the invariants now when the background is uniform.

Figure 4 presents the classification accuracy of the two methods as a function of PSNR, when the experiment is repeated 20 times for each noise level and for each test image. Similar to the previous experiment, the number of the invariants used is 10 and 18 shown by the solid and dashed lines, respectively. It is clear from the results that the MOMIs are affected much more by the noise. Especially, the noise in the background seems to be harmful. On the contrary, the PBTIs can handle also this situation and perform very well compared to the MOMIs.

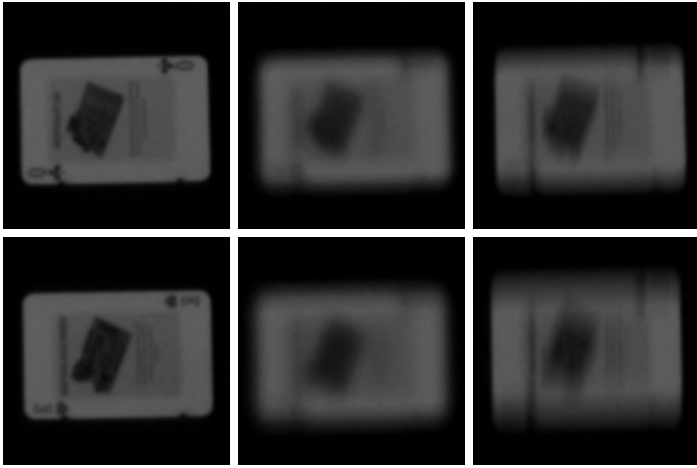


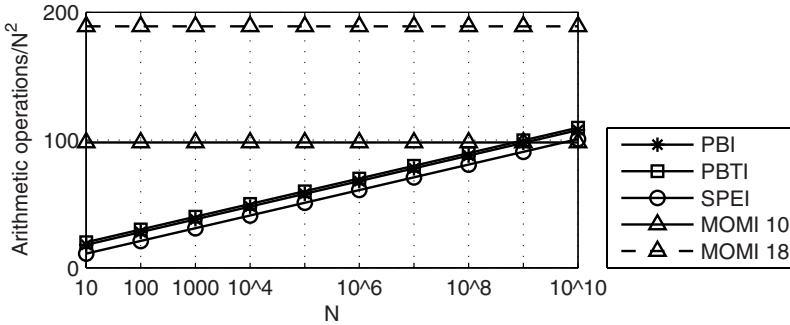
Fig. 5. Each row contains examples of one class (Queen of Clubs and King of Hearts, respectively) of the total of 10 classes of playing card images used in the third classification experiment. Cards were unusual and contained very similar figures. From left to right, there are a sharp, defocused and motion blurred image of each class.

In the final experiment, we wanted to confirm the results of the second experiment in practice. We classified real gray scale images of playing cards on the black background which were captured using a Sony DFW-X710 video camera. First, we formed 10 classes by photographing sharp images of 10 playing cards. Then, we captured three motion blurred and three out-of-focus blurred versions of these 10 cards resulting to 60 blurred card images that were to be classified as one of the 10 classes. The size of the images was 130-by-130. Motion blur was generated by panning the camera vertically in a tripod with a relatively long shutter time. Out-of-focus blur was created by defocusing the camera slightly. In Figure 5, there are sample images of two classes. In each row, from left to right, there is a sharp, defocused and motion blurred image of the same class. As can be seen, the blur is so strong that it is impossible to recognize the cards visually. Table 1 shows the classification accuracies for the two methods when 10 and 18 invariants are used. The best accuracy 95 % is achieved using 18 PBTIs. The number of correctly classified images is nearly double compared to the 18 MOMIs with the accuracy of 48 %. For example, PBTIs were able to classify all the blurred images in Figure 5, but MOMIs could only classify the blurred images on the top row where the level of blurring is lower. These results are in line with the results obtained using the artificial images.

Finally, we discuss about the computational demands of the invariants. Asymptotic complexity of the central moments, as well as the MOMIs which are build on them, is $O(N^2)$. This is less than the complexity of all types of the frequency domain invariants using radix-2 FFT, which is $O(N^2 \log_2 N)$. However, for practical image sizes the MOMIs require much more computation than the frequency domain invariants, as shown in Figure 6 where the approximate number of arithmetic

Table 1. Classification accuracies of the PBTIs and the MOMIs in the case of real degraded images of playing cards when 10 and 18 invariants are used

Method	PBTIs 18	PBTIs 10	MOMIs 18	MOMIs 10
Accuracy	95 %	73 %	48 %	45 %

**Fig. 6.** The number of arithmetic operations needed to compute different types of invariants for an N -by- N image. For MOMIs the value depends on the number of invariants, which is 10 and 18.

operations per pixel is shown as the function of N when the image size is N -by- N . All operations, including complex operations of FFT and high exponents of the moments, are counted as one for simplicity. In addition, the computation load for the MOMIs also depends on the number of invariants used, while for the frequency domain methods $N/2$ non-redundant invariants are derived simultaneously. As can be seen, 10 MOMIs becomes preferable in terms of operation count at the image size 10^9 -by- 10^9 . For 18 MOMIs this size is 10^{19} -by- 10^{19} . Both of these image sizes are unrealistic. The explanation for this property is the very large number of geometric moments that have to be computed to get the corresponding MOMI features.

4 Conclusions

In this paper, we have shown how it is possible to derive new blur invariant features, PBIs and PBTIs, based on even powers of the phase-only spectrum and bispectrum of the images, respectively. The features are invariant to centrally symmetric blur, and the PBTI features based on the bispectrum are also invariant to translation. The features can be used to recognize blurred images or objects as demonstrated in the paper. Because the features are based on the normalized phase-only spectrum or bispectrum, they are also invariant to linear brightness changes. In addition, they can be computed efficiently using FFT.

We compared the PBIs and the PBTIs with the MOMI and SPEI features proposed in [4] for classification of blurred images and objects on an uniform

background. In both cases, our features performed better. In practice, our features need also less computation, if we do not take into account the SPEIs which are very sensitive to noise.

A shortcoming of our features is that they do not carry further invariance to geometrical transformations such as rotation and scaling. On the other hand, there is no other frequency domain blur invariant method available that would carry the invariance to translation.

References

1. Wood, J.: Invariant pattern recognition: A review. *Pattern Recognition* 29, 1–17 (1996)
2. Banham, M.R., Katsaggelos, A.K.: Digital image restoration. *IEEE Signal Processing Magazine* 14(2), 24–41 (1997)
3. Kundur, D., Hatzinakos, D.: Blind image deconvolution. *IEEE Signal Processing Magazine* 13(3), 43–64 (1996)
4. Flusser, J., Suk, T.: Degraded image analysis: An invariant approach. *IEEE Trans. Pattern Anal. Machine Intell.* 20, 590–603 (1998)
5. Flusser, J., Suk, T., Saic, S.: Recognition of blurred images by the method of moments. *IEEE Transactions on Image Processing* 5(3), 533–538 (1996)
6. Bentoutou, Y., Taleb, N., Mezouar, M.C.E., Taleb, M., Jetto, L.: An invariant approach for image registration in digital subtraction angiography. *Pattern Recognition* 35, 2853–2865 (2002)
7. Flusser, J.: Combined invariants to linear filtering and rotation. *Int. J. Pattern Recognition and Artificial Intelligence* 13(8), 1123–1136 (1999)
8. Zhang, Y., Wen, C., Zhang, Y., Soh, Y.C.: Determination of blur and affine combined invariants by normalization. *Pattern recognition* 35(1), 211–221 (2002)
9. Suk, T., Flusser, J.: Combined blur and affine moment invariants and their use in pattern recognition. *Pattern Recognition* 26(12), 2895–2907 (2003)
10. Ojansivu, V.: Motion blur concealment of digital video using invariant features. In: Blanc-Talon, J., Philips, W., Popescu, D., Scheunders, P. (eds.) *ACIVS 2006*. LNCS, vol. 4179, Springer, Heidelberg (2006)
11. Chandran, V., Carswell, B., Boashash, B., Elgar, S.: Pattern recognition using invariants defined from higher order spectra: 2-d image inputs. *IEEE Transactions on Image Processing* 6(5), 703–712 (1997)
12. Dianat, S.A., Rao, R.M.: Fast algorithms for phase and magnitude reconstruction from bispectra. *Optical Engineering* 29(5), 504–512 (1990)
13. Petropulu, A.P., Pozidis, H.: Phase reconstruction from bispectrum slices. *IEEE Transactions on Signal Processing* 46(2), 527–530 (1998)