PERGAMON

Pattern Recognition 35 (2002) 2993–2999

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Using Fourier/Mellin-based correlators and their fractional versions in navigational tasks

Didi Sazbon[a,*], Zeev Zalevsky[b], Ehud Rivlin[a], David Mendlovic[b]

[a] *Faculty of Computer Science, Technion—Israel Institute of Technology, 32000 Haifa, Israel*
[b] *Department of Electrical Engineering, Tel-Aviv University, Tel-Aviv, Israel*

Received 13 October 2000; accepted 30 October 2001

Abstract

The navigational tasks of computing time to impact and controlling movements in specific range are addressed here. We show how time to impact can be obtained via the Mellin-based correlator, and we introduce the concept of fractional based correlators. The fractional Fourier/Mellin based correlators can be used in detecting or controlling specific range of movements. Also, both the conventional and the fractional-based correlators can be easily implemented optically in lenses, thus providing correlation images directly at image acquisition time. Lenses are considered cost-effective and therefore, the optical correlators are optimal in both senses, speed and cost. © 2002 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Correlators; Fourier transform; Mellin transform; Fractional Fourier transform; Fractional Mellin transform; Time to impact; Range control

1. Introduction

Correlators are well known for their role in object recognition. Here, their possible applications to navigational tasks were investigated. We will show that when the motion is assumed to consist of rotation and scaling, the Mellin-based correlator can be used to compute time to impact. In addition, we will introduce the concept of fractional correlators. The fractional Fourier-based correlator can be used to control the range of translation, while the Fractional Mellin-based correlator, can be used to control the range of rotation and scaling.

Algorithms concerning the above mentioned navigational tasks are usually known to be computationally heavy and considered as time consuming [1–7]. Therefore, physical sensors that provide as much information as possible directly at acquisition time have an utmost importance. The major advantage of the correlators is the fact that they can be easily implemented optically. Here, we suggest the common use of cameras, but with optically implemented functions. This means that the correlators can be implemented within the camera, by using special purpose lenses, and can be obtained at image acquisition time. The special lenses can be formed by different combinations of regular Fourier lenses and special purpose designed filters [8–10]. In addition to the high speed in which relevant data can be supplied, lenses are usually cost effective. Thus, this solution is optimal in both senses, speed and cost.

The usage of optical correlators in real time vehicle navigation was previously demonstrated by Psaltis [11]. There, an opto-electronic information processing system based on holographic memory database (realized using the DuPont HRF-150 photopolymer) is applied to perform the navigational tasks. In contrast to the database technique, this paper uses transformations that are optically implemented, and thus, shows how translation, rotation, and scaling can be computed simply, without the requirement for a digital processor of high capacity.

This paper is organized as follows: Section 2 discusses the Mellin correlator and its role in time to impact estimation. Section 3 presents the theoretical concept of the fractional based correlators and their role in range control. Section

* Corresponding author. Tel.: +972-4-8294894; fax: +972-4-8221128.
E-mail address: didi@cs.technion.ac.il (D. Sazbon).

PII: S0031-3203(02)00018-3

4 shows and discusses experiments that demonstrate these concepts, and Section 5 is a summary.

2. Time to impact estimation using the Mellin-based correlator

The correlation image provides information concerning the translation of a tested pattern in relation to a centered reference pattern. The correlation of two functions, h and g, is given by:

$$h \otimes g = FT^{-1}\{H^* G\} = \int_{-\infty}^{\infty} h(x')g^*(x'-x)\,\mathrm{d}x' \qquad (1)$$

where H and G are the Fourier transform of h and g, respectively, H^* is the complex conjugate of H, and FT denotes the Fourier transform. The correlation of the pattern with itself will produce a maximum (called 'correlation-peak') in the center of the correlation image. If the tested pattern is a translated version of the reference pattern, the value of the correlation-peak will be translated accordingly.

The major drawback of the correlation function is that it is sensitive to rotation and scaling. In order to overcome this problem, the log-polar mapping of the functions can be used. The coordinate transformation given by

$$\rho = \ln\sqrt{x^2+y^2}, \quad \theta = \tan^{-1}\frac{y}{x}, \qquad (2)$$

defines a log-polar mapping, $f(\rho,\theta)$, of a function $f(x,y)$.

The Fourier transform of an image in its log-polar representation is known as the Mellin transform. The 2D Mellin transform is given by

$$M(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \times \frac{\exp\{-2\pi\mathrm{i}(u\ln\sqrt{x^2+y^2}+v\tan^{-1}y/x)\}}{x^2+y^2}\,\mathrm{d}x\,\mathrm{d}y. \qquad (3)$$

Replacing x and y with the log-polar coordinates described by Eq. (2), the Fourier transform of the image in its log-polar coordinates is obtained:

$$M(u,v) = \int_{0}^{2\pi}\int_{-\infty}^{\infty} f(\rho,\theta)\exp\{-2\pi\mathrm{i}(u\rho+v\theta)\}\,\mathrm{d}\rho\,\mathrm{d}\theta = FT(f(\rho,\theta)), \qquad (4)$$

where FT denotes the Fourier transform.

Thus, if the tested pattern is a rotated and scaled version of the reference pattern, the correlation image using the Mellin transform instead of the Fourier transform will produce a translated correlation-peak in the ρ and θ plane. Clearly, a translation in the ρ coordinate is actually a scaling in the x–y plane, and a translation in the θ coordinate is actually a rotation in the x–y plane.

An important observation should be made that by gaining invariance to scale and rotation, the invariance to translation is lost. Thus, there are two distinct types of correlators, one is translation invariant and the second is scale and rotation invariant. Moreover, the Mellin transform described above is commonly known as the Mellin transform of type 2, where there exists an additional transformation known as Mellin transform of type 1. In the Mellin transform of type 1, the x and y coordinates are replaced with $\ln x$ and $\ln y$, correspondingly. The analysis of a correlator based on Mellin of type 1 will be similar to that of type 2, and it can be used in cases where the scaling along the two axes is different.

We will show an interesting and straightforward way to compute time to impact using the Mellin-based correlator. In this case the motion is restricted to rotation and scaling. It is known that under perspective projection of a camera with focal length F, the length, l, of an object in the image plane is given by

$$l = F\frac{L}{R}, \qquad (5)$$

where L is its physical length, and R is its distance from the center of projection of the camera. If the natural logarithm of l is taken, then

$$\ln l = \ln L + \ln F - \ln R \qquad (6)$$

and by taking the derivative of Eq. (6), we obtain

$$\frac{\mathrm{d}\ln l}{\mathrm{d}t} = \frac{\mathrm{d}(\ln L + \ln F - \ln R)}{\mathrm{d}t} = -\frac{1}{R}\frac{\mathrm{d}R}{\mathrm{d}t} = -\frac{1}{R}v = -\frac{v}{R}, \qquad (7)$$

where v is the velocity of the imaged object. Note that L and F are constants, and thus the derivatives of their components are zero. On the other hand, the distance of the camera from the object is given by

$$R = v\,TTI \qquad (8)$$

or

$$\frac{1}{TTI} = \frac{v}{R}, \qquad (9)$$

where TTI denotes the time to impact. Combining Eq. (7) with Eq. (9), yields

$$TTI = \left(-\frac{\mathrm{d}\ln l}{\mathrm{d}t}\right)^{-1} \qquad (10)$$

meaning, that the time to impact can be extracted from the change in the temporal variations of the logarithm of the object in the image plane.

Therefore, if two images of the object are taken Δt time apart, the derivative of $\ln l$ can be calculated from these images, and the time to impact can be obtained directly. If the object has length L_1 in the first image and length L_2 in the second image then

$$TTI = \left(-\frac{\ln L_1 - \ln L_2}{\Delta t}\right)^{-1}. \qquad (11)$$

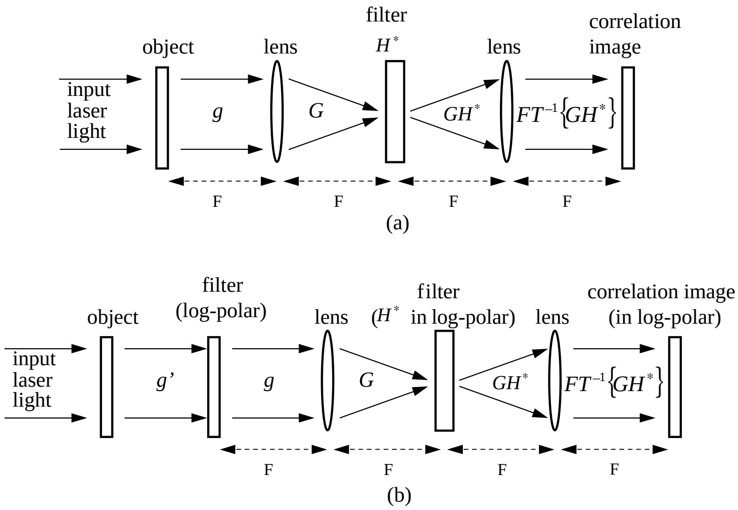

Fig. 1. The schematic implementation of the conventional correlators in optics: (a) The correlator based on the Fourier transform (used in detecting translations). (b) The correlator based on the Mellin transform (used in detecting rotation and scaling). *FT* denotes the Fourier transform, G and H are the Fourier transforms of the tested image and the reference image accordingly, and F is the focal length of the lenses.

Both L_1 and L_2 can be given directly in pixels, because any other representation is equivalent to summation and subtraction of the same constant.

Note that by using the Mellin-based correlators, the time to impact computation of the imaged object is straightforward. Let L_2 be a scaled version of L_1, meaning $L_2 = sL_1$ then:

$$TTI = \left(-\frac{\ln L_1 - \ln L_2}{\Delta t}\right)^{-1} = \left(-\frac{\ln L_1 - \ln sL_1}{\Delta t}\right)^{-1}$$
$$= \left(-\frac{\ln L_1/sL_1}{\Delta t}\right)^{-1} = \left(\frac{\ln s}{\Delta t}\right)^{-1}. \quad (12)$$

Thus, one needs to know only the scaling factor, s, which is calculated directly from the shift of the correlation peak.

3. Fractional Fourier-based correlators and their application to range control

The common correlators provide information regarding the translation of a certain object or about its scaling and rotation. In some visual systems there is no need to know every type of motion, and the only interesting movements are limited to a specific range. For example, if a robot is moving around in a room consisting of known obstacles, it should not come near an obstacle if it is at a certain distance away. We would like to be able to detect only the nearby obstacles and ignore the others. A variant of this example would be a stationary robot sorting objects from the same type, that should ignore objects bigger or smaller than a desired range of sizes. In this case, we would like to recognize only the objects that are in the desired range and ignore all others. Another example would be a helicopter that should fly over a target in a specific range.

Here, a new type of correlators, that can detect movements in a specific range, is introduced. The correlators are based on the concepts of fractional transformations, i.e. the fractional Fourier transform (FRT) and the fractional Mellin transform. The optical implementation of the FRT in optics is well documented [9–15] (see a schematic description in Fig. 1), and is done by changing the distances and the focal length of the lenses performing the Fourier transform and its inverse.

The 2D fractional Fourier transform of order (p_1, p_2) is given by

$$F^{(p_1,p_2)}(u,v) = C \cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \times \exp\left\{-2\pi i\left(\frac{ux}{\sin\phi_1} + \frac{vy}{\sin\phi_2}\right) + \pi i\left(\frac{u^2+x^2}{\tan\phi_1} + \frac{v^2+y^2}{\tan\phi_2}\right)\right\} dx\,dy \quad (13)$$

where C is a constant, and

$$\phi_1 = \frac{p_1\pi}{2} \quad \text{and} \quad \phi_2 = \frac{p_2\pi}{2}. \quad (14)$$

Note that $F^{(0,0)}(u,v) = f(x,y)$, which is the image itself, and $F^{(1,1)}(u,v) = F(u,v)$, which is the Fourier transform of the image.

We define the 2D fractional Mellin transform (type 2) of order (p_1, p_2) by

$$M^{(p_1,p_2)}(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)x^{-2\pi i/in\phi_1(u-1)} \times \exp\left\{\frac{\pi i}{\tan\phi_1}(u^2+\ln^2 x)\right\} \times x^{-2\pi i/\sin\phi_2(v-1)}\exp\left\{\frac{\pi i}{\tan\phi_2}(v^2+\ln^2 y)\right\}\,dx\,dy, \quad (15)$$

where ϕ_1 and ϕ_2 are the same as in Eq. (14). Note that $M^{(0,0)}(u,v) = f(\rho,\theta)$, which is the image itself in log-polar representation, and $M^{(1,1)}(u,v) = M(u,v)$, which is the Mellin transform of the image. In the general case $M^{(p_1,p_2)}(u,v) = FRT^{(p_1,p_2)}(f(\rho,\theta))$, which is the fractional Fourier transform of order (p_1, p_2) of the image in its log-polar representation.

How does the correlation pattern change when we apply the fractional correlators instead of the conventional correlators? The fractional transformations of order (1, 1) are actually the 2D Fourier or Mellin transforms of the image, and the correlation pattern has a peak at the positions corresponding to those of the conventional correlators. As the order decreases, enlarging the movement results in a lower magnitude of the principal peak and increase of the side lobes (additional peaks). If the order is only slightly decreased, then the location of the principal peak still corresponds to the position of the peak in the regular correlators, otherwise it may not correspond. Therefore, in order to limit the range of the desired movement, we should pre-determine the order of the fractional transformation used and the threshold level, in such a way, that the resulting peak would still be in a position corresponding to the correct movement. An object moving outside the range that was pre-determined will produce a correlation-peak which falls below the threshold level.

4. Experimental results and discussion

4.1. Time-to-impact computation

Fig. 1a depicts the schematic implementation of the Fourier-based correlators in optics, used in detecting translations, and Fig. 1b depicts the schematic implementation of the Mellin-based correlator in optics, used in detecting scaling and rotation. Here, we concentrate on the Mellin based correlator and its usage in estimating time to impact. Although we use it in a computational fashion, it can be incorporated into a single lens commercially, and therefore, can be obtained at image acquisition time.

The images of the computer depicted in Fig. 2(a) and (b) are taken one time unit apart, which we will assume to be equal to 1 s. The first image is taken ~ 2.85 m away from the computer and the second image is taken approximately 2.6 m away from the computer. A rough approximation of the time to impact can be logically deduced by assuming a constant velocity of 0.25 m per second, which turns out to be ~ 10.4 s.

Now, we will assume that we do not know the actual location of the computer in the room, and we would like to estimate the time to impact using only the two images. The way to do so is by using the Mellin-based correlators, as was explained in Section 2.1, to measure the scaling factor, and then to use Eq. (12). The scaling factor using the Mellin correlator on these two images resulted in a 1.1 scaling, and the resulting time to impact is:

$$TTI = \left(\frac{\ln 1.1}{\Delta t}\right)^{-1} = \frac{1}{\ln 1.1} = 10.492 \text{ s}. \quad (17)$$

This result agrees with the rough approximation done by calculating the real distances and velocities.

Hence, the computation of time to impact can be deduced straightforwardly by using the Mellin-based correlators. The location of the correlation peak corresponds to the scaling of the object in the second image with respect to the object in the first image. This technique can be implemented optically to give a correlation image at acquisition time, and therefore, can be used as a real-time technique.

4.2. Specific range estimation

The sequence of the three images depicted in Fig. 2 can be used also in order to demonstrate the specific range estimation capability of the fractional correlators. Images 2a–c, are taken in such a way that each image is a scaled version of the previous one with a scaling factor of 1.1. Image 2a was correlated using the fractional Mellin correlator with itself and with the other two images, and the correlation peaks were measured. A variety of fractional orders, p, were used, and the table presented in Fig. 3 indicates the locations of the correlation peaks measured for each fractional order. All the fractional orders used had values in the range [0.95,1], which implies that regular Mellin properties are strongly dominant. In other words, where shifts in specific ranges are concerned, the locations of the correlation peaks in these ranges should be around the locations of the correlation peaks resulting from using the Mellin transform.

The table shows that a fractional order of 0.995 allows scaling of 1.1 but does not allow scaling of 1.21. This means that when using $p = 0.995$, an object that is located away from the camera, a distance corresponding to a scale of 1.1 will be detected. On the other hand, positioning the object in a distance corresponding to a scale of 1.21 will not produce a detection indication. Note that the auto-correlation peaks of the image with itself are located at the center of the image for any choice of p, as is expected. Generally speaking, in

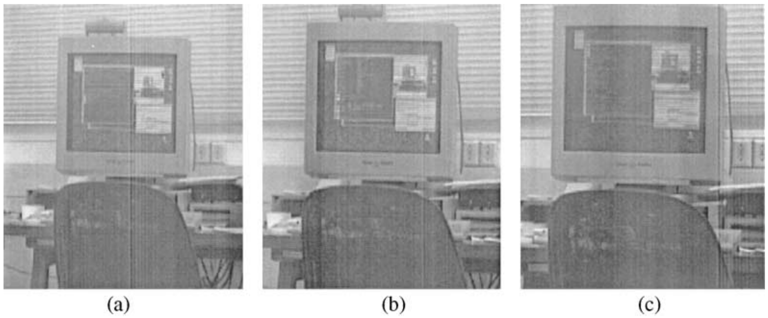

Fig. 2. A sequence of three images with a scaling factor of 1.1.

p	correlation peak image 5a (scaling=1)	correlation peak image 5b (scaling=1.1)	correlation peak image 5c (scaling=1.21)
1 peak location: peak value:	121,121 31.9345	127,121 28.9544	132,121 26.9186
0.999 peak location: peak value:	121,121 31.9326	127,121 28.9530	132,121 26.9071
0.998 peak location: peak value:	121,121 31.9269	127,121 28.9491	132,121 26.8752
0.997 peak location: peak value:	121,121 31.9176	127,121 28.9426	132,121 26.8556
0.996 peak location: peak value:	121,121 31.9052	127,121 28.9359	125,121 26.8332
0.995 peak location: peak value:	121,121 31.8900	126,121 28.9327	122,121 26.8193
0.994 peak location: peak value:	121,121 31.8727	126,121 28.9312	122,121 26.8121
0.99 peak location: peak value:	121,121 31.7952	122,121 28.9273	122,121 26.7971
0.95 peak location: peak value:	121,121 31.5852	121,121 28.8348	121,121 26.7478

Fig. 3. Using the Fractional Mellin correlator on the images from Fig. 5 produced the correlation peaks (locations in pixels and values) given in the table. It can be seen that a choice of fractional order $p = 0.995$ allows scaling of 1.1 but not of 1.21. This implies that ranges up to a scaling of 1.1 can be detected, and ranges larger that 1.21 cannot be detected.

fractional correlators the shift invariance range, R, can be controlled according to Ref. [15]:

$$|R| < \frac{\lambda F tg(\pi p/2)}{\Delta}, \tag{18}$$

where, λ is the wavelength, F is the focal length of the system performing optically the FRT, and Δ is the spatial region in which the object is defined. For digital computation of FRT one may choose $\lambda F = 1$.

An additional example is illustrated in Fig. 4, where a sequence of six images depicting a robot is used. Image 4d was correlated with itself and with the other images using the fractional Mellin correlator. The scaling factors of the images with respect to image 7d are 0.83, 0.87, 0.96, 1, 1.13, 1.21. Fractional orders in the range $[0.96, 1]$ were tested, and an analysis similar to the analysis performed in Fig. 3 was conducted. The results showed that $p = 0.999$ did not allow a scaling smaller than 0.83, $p = 0.996$ did not allow a

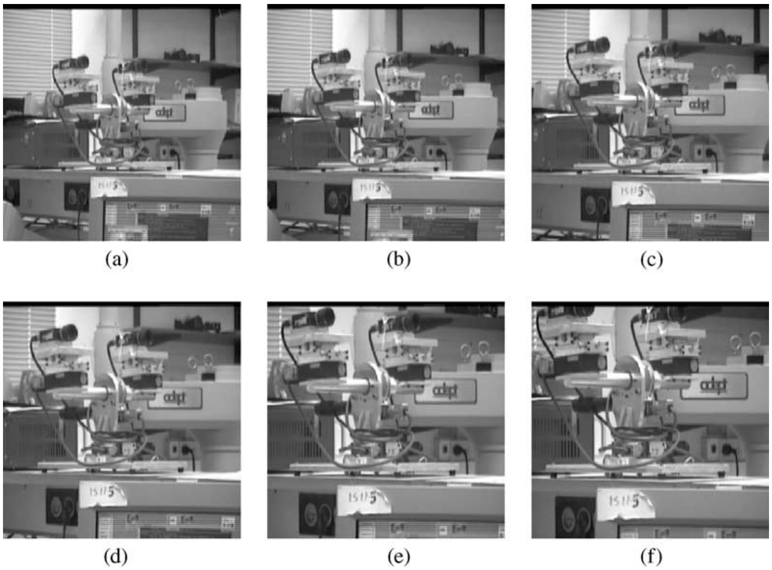

Fig. 4. A sequence of a laboratory images with a scaling factor of 1.1.

scaling bigger than 1.21 or smaller than 0.87, and $p = 0.994$ did not allow a scaling bigger than 1.13 or smaller than 0.96.

The conclusion that can be drawn out of these experiments is that movements with specific ranges can be controlled via the fractional correlators, by a proper choice of the fractional order that fits the desired ranges. This technique can be implemented optically to give a correlation image at acquisition time, and therefore, can be a real-time technique.

5. Summary

This paper introduced different types of correlators that can find applications in common tasks of visual navigation. The major advantage of the correlators is the fact that they can be implemented optically, and therefore can be obtained at image acquisition time and can be obtained at real time. The correlators discussed here are: (a) The Mellin based correlator that can be used in computation of time to impact. (b) The Fractional Mellin and fractional Fourier-based correlators (first defined in this study) that can be used to control movements in specific ranges.

The experiments proved that both applications can be executed using the correlators. The time to impact experiment showed that time to impact can be estimated using only two images and the Mellin-based correlator (optical or computational), with no additional a-priory knowledge. The specific ranges estimation experiments showed how the specific desired ranges of movement can be controlled, while all others ignored.

References

[1] R. Jain, Direct computation of the focus of expansion, IEEE Trans. Pattern Anal. Mach. Intell. 5 (1) (1983) 58–64.
[2] M. Tistarelli, G. Sandini, Estimation of depth from motion using an anthropomorphic visual sensor, Image Vision Comput. 8 (4) (1990) 271–278.
[3] M. Tistarelli, G. Sandini, On the advantages of polar and log-polar mapping for direct estimation of time to impact from optical flow, IEEE Trans. Pattern Anal. Mach. Intell. 15 (4) (1993) 401–410.
[4] H. Tunley, D. Young, First order optic flow from log-polar sampled images, ECCV A (1994) 132–137.
[5] F.L. Lim, G.A.W. West, S. Venkatesh, Use of log polar space for foveation and feature recognition, IEE Proc. Vision Image Signal Process. 144 (6) (1997) 323–331.
[6] C. Silva, J. Santos-Victor, Egomotion estimation using log-polar images, ICCV (1998) 967–972.
[7] I. Ahrns, H. Neumann, Real-time monocular fixation control using the log-polar transformation and confidence-based similarity measure, ICPR (1998) 310–315.
[8] D. Mendlovic, H.M. Ozaktas, Fractional Fourier transformations and their optical implementation: Part I, JOSA A 10 (1993) 1875–1881.
[9] H.M. Ozaktas, D. Mendlovic, Fractional Fourier transformations and their optical implementation: Part II, JOSA A 10 (1993) 2522–2531.
[10] H.M. Ozaktas, D. Mendlovic, Fourier transforms of fractional orders and their optical interpretation, Opt. Commun. 101 (1993) 163–165.
[11] A. Pu, R. Denkewalter, D. Psaltis, Real time vehicle navigation using a holographic memory, Opt. Eng. 36 (10) (1997) 2737–2746.
[12] A.W. Lohmann, Image rotation, Wigner rotation and the fractional Fourier transform, JOSA A10 (1993) 2181–2186.

[13] D. Mendlovic, H.M. Ozaktas, A.W. Lohmann, Graded index fibers, Wigner distribution functions and the fractional Fourier transform, Appl. Opt. 33 (1994) 6188–6193.
[14] J. Garcia, D. Mendlovic, Z. Zalevsky, A.W. Lohmann, Space variant simultaneous detection of several objects using multiple anamorphic fractional Fourier transform filters, Appl. Opt. 35 (1996) 3945–3952.
[15] J.H. Caulfield, D. Mendlovic, Z. Zalevsky, Fractional correlator with real time control of the space-invariance property, Appl. Opt. 36 (1997) 2370–2375.

About the Author—DIDI SAZBON received her B.Sc. degree in Computer Science and her M.Sc. degree in Biomedical Engineering from the Technion, Israel Institute of Technology. Currently she is a Ph.D. student in the Computer Science Department at the Technion, Israel Institute of Technology. Her current research interests are in computer vision and robot navigation.

About the Author—ZEEV ZALEVSKY received his B.Sc. degree in Electrical Engineering from the Tel-Aviv University, Israel, in 1993. He completed his Ph.D. in the Department of Physical Electronics at the Tel-Aviv University in 1996. His major research interests are optical signal processing, optical pattern recognition, super resolution, temporal optics, computer vision, and optic communication.

About the Author—EHUD RIVLIN received his B.Sc. and his M.Sc. degrees in Computer Science and his M.B.A. degree from the Hebrew University in Jerusalem, Israel, and his Ph.D. degree from the University of Maryland. Currently he is an Associate Professor in the Computer Science Department at the Technion, Israel Institute of Technology. His current research interests are in machine vision and robot navigation.

About the Author—DAVID MENDLOVIC received his B.Sc. and his Ph.D. degrees in Electrical Engineering from the Tel-Aviv University, Israel, in 1987 and 1990, respectively. At present, he is a Professor in the Department of Physical Electronics. His research interests include optics for computers, optical signal processing, diffractive optics, three-dimensional surface sensing and temporal optics.