Optical Transformations in Visual Navigation

Didi Sazbon, Ehud Rivlin
Faculty of Computer Science
Technion, Israel Institute of Technology
Haifa, Israel

Zeev Zalevsky, David Mendelovic
Faculty of Engineering
Tel-Aviv University
Tel-Aviv, Israel

Abstract

The navigational tasks of computing time-to-impact and controlling movements within specific range are addressed here. By using specially designed lenses various components of these procedures, consisting of mathematical transformations, can be provided at image acquisition time, and therefore, speed up execution time. This study discusses the optical implementation of different correlators based on the Fourier transform and Mellin transform. In addition, the fractional versions of these correlators are defined and analyzed here. Based on the experimental results it can be concluded that the optical implementation of transformations can indeed play a significant role in speeding up execution time in respect to the above mentioned navigational tasks.

1. Introduction

Algorithms concerning visual navigation aspects are usually known to be computationally heavy and considered as time consuming. Therefore, physical sensors that provide as much information as possible directly at acquisition time have an utmost importance. Sensors performing special operations can considerably save processing time. An example to such a sensor can be a retina-like sensor performing the log-polar mapping, which has many applications in motion estimation [1-8]. Constructing such special purpose sensors can be a very complicated task from an engineering point of view, because it usually involves a special geometrical positioning of known sensors. Moreover, the use of many sensors and the supporting algorithms enlarges significantly the overall cost of the system.

Here, a different approach is taken, which suggest the common use of cameras, but with optically implemented functions. This means that the functions we wish to obtain directly at acquisition time, will be implemented within the camera, by using special purpose lenses. 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 special lenses are formed by different combinations of regular Fourier lenses and special purpose designed filters [9-11]. This paper concentrates on two designs: (a) Fourier and Mellin based correlators, and (b) Fractional Fourier and Fractional Mellin based correlators, which concept is first introduced here. Each design answers a specific problem in visual navigation. The first design serves in the computation of time-to-impact, and the later in controlling and estimating specific ranges.

2. Optical Transformations and Visual Navigation

In what follows we present two optical transformations and their possible implementation to visual navigation. First, correlators, optical correlators, and their possible application to time-to-impact calculation are discussed. Then, the concept of fractional correlators is introduced, and its possible application to selective range estimation is explained.

2.1 Correlators and Their Application to Time-to-Impact Estimation

The 2D Mellin Transform is given by:

$$M(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\cdot \frac{\exp\left\{-2\pi i\left(u\ln\sqrt{x^2+y^2}+v\tan^{-1}\frac{y}{x}\right)\right\}}{x^2+y^2}dxdy \qquad (1)$$

Therefore, the Mellin Transform of an image is in fact its Fourier Transform in a log-polar representation. Replacing x and y with the log-polar coordinates (ρ,θ), the Fourier Transform of the image in its log-polar coordinates is obtained. The log-polar mapping is defined by:

$$\rho = \ln\sqrt{x^2+y^2} \qquad \theta = \tan^{-1}\frac{y}{x} \tag{2}$$

and therefore,

$$M(u,v) = \int_{0}^{2\pi}\int_{-\infty}^{\infty} f(\rho,\theta)\exp\{-2\pi i(u\rho+v\theta)\}d\rho d\theta = FT(f(\rho,\theta)) \tag{3}$$

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. Therefore, the resulted transform is invariant both to scaling and rotation.

An interesting application of the correlators is the calculation of time-to-impact. In this case the motion is restricted to rotation and scaling, without allowing any other form of translation. 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

$$l = F\frac{L}{R} \tag{4}$$

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

$$\ln l = \ln L + \ln F - \ln R \tag{5}$$

and by taking its derivative, we obtain

$$\frac{d\ln l}{dt} = \frac{d(\ln L + \ln F - \ln R)}{dt} = -\frac{1}{R}\frac{dR}{dt} = -\frac{1}{R}v = -\frac{v}{R} \tag{6}$$

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

$$\frac{1}{TTI} = \frac{v}{R} \tag{7}$$

where TTI denotes the time-to-impact. Combining the last two equations yields

$$TTI = \left(-\frac{d\ln l}{dt}\right)^{-1} \tag{8}$$

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

$$TTI = \left(-\frac{\ln L_1 - \ln L_2}{\Delta t}\right)^{-1} \tag{9}$$

Both L 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 straight forward. Let L_2 be a scaled version of L, meaning $L_2 = sL$, then:

$$TTI = \left(-\frac{\ln L_1 - \ln L_2}{\Delta t}\right)^{-1} = \left(-\frac{\ln L_1 - \ln sL_1}{\Delta t}\right)^{-1} = \left(-\frac{\ln\frac{L_1}{sL_1}}{\Delta t}\right)^{-1} = \left(\frac{\ln s}{\Delta t}\right)^{-1} \tag{10}$$

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

2.2 Fractional 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 an helicopter that should fly over a target in a specific range.

Here, a new type of correlators, that can recognize movements in a specific range, are 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], 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:

$$F^{(p_1,p_2)}(u,v) = C \cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \cdot \exp\left\{ \begin{array}{l} -2\pi i\left(\frac{ux}{\sin\phi_1} + \frac{vy}{\sin\phi_2}\right) \\ +\pi i\left(\frac{u^2+x^2}{\tan\phi_1} + \frac{v^2+y^2}{\tan\phi_2}\right) \end{array} \right\} dxdy \quad (11)$$

where C is a constant, and

$$\phi_1 = \frac{p_1\pi}{2} \quad \text{and} \quad \phi_2 = \frac{p_2\pi}{2} \quad (12)$$

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 of order (p_1, p_2) by:

$$M^{(p_1,p_2)}(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) \cdot x^{-\frac{2\pi i}{in\phi_1}u-1} \cdot \exp\left\{\frac{\pi i}{\tan\phi_1}\left(u^2 + \ln^2 x\right)\right\} \cdot x^{-\frac{2\pi i}{\sin\phi_2}v-1} \cdot \exp\left\{\frac{\pi i}{\tan\phi_2}\left(v^2 + \ln^2 y\right)\right\} dxdy \quad (13)$$

where ϕ_1 and ϕ are the same as in the FRT case. 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 resulted 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.

3. Experimental Results and Discussion

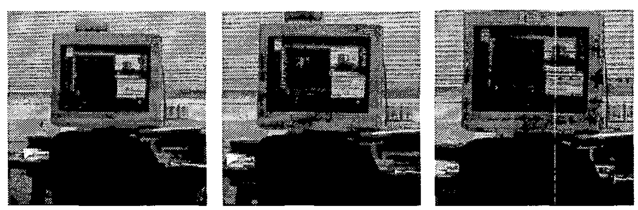
The images of the computer depicted in Figure 1a and Figure 1b are taken one time unit apart, which we will assume to be equal to one second. The first image is taken approximately 2.85 meters away from the computer and the second image is taken approximately 2.6 meters away from the computer. A rough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

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 scaling factor using the Mellin correlator on these two images resulted in a 1.1 scaling, and the resulted time-to-impact is:

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

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 straight forward by using the Mellin based correlators. The location of the correlation peak corresponds to the scaling of the object in the second image in 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.



(a) (b) (c)

Figure 1: A sequence of images with a scaling factor of 1.1.

The sequence of the three images depicted in Figure 1 can be used also in order to demonstrate the specific range estimation capability of the fractional correlators. Images 1a, 1b, and 1c, are taken in such a way that each is a scaled version of the previous with a scaling factor of 1.1. Image 1a 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 Figure 2 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 resulted 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.

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

Figure 2: Using the Fractional Mellin correlator on the images from Figure 1 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 can not be detected.

The conclusion that can be drawn out of such 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.

References

1. R. Jain, *Direct Computation of the Focus of Expansion*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 5, No. 1, pp. 58-64, 1983.
2. M. Tistarelli, and G. Sandini, *Estimation of Depth from Motion Using an Anthropomorphic Visual Sensor*, Image and Vision Computing, Vol. 8, No. 4, pp. 271-278, 1990.
3. M. Tistarelli, and G. Sandini, *On the Advantages of Polar and Log-Polar Mapping for Direct Estimation of Time-to-Impact from Optical Flow*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, No. 4, pp. 401-410, 1993.
4. H. Tunley, and D. Young, *First Order Optic Flow from Log-Polar Sampled Images*, ECCV, vol. A, pp. 132-137, 1994.
5. P. Questa, and G. Sandini, *Time to Contact Computation with a Space-Variant Retina-Like C-mos Sensor*, Proc. IROS, pp. 1622-1629, 1996.
6. F.L. Lim, G.A.W. West, and S. Venkatesh, *Use of Log Polar Space for Foveation and Feature Recognition*, IEE Proceedings of Vision, Image, and Signal Processing, Vol. 144, No. 6, pp. 323-331, 1997.
7. C. Silva, and J. Santos-Victor, *Egomotion Estimation Using Log-Polar Images*, ICCV, pp. 967-972, 1998.
8. I. Ahrns, and H. Neumann, *Real-Time Monocular Fixation Control Using the Log-Polar Transformation and Confidence-Based Similarity measure*, ICPR, pp. 310-315, 1998.
9. D. Mendlovic, and H.M. Ozaktas, *Fractional Fourier Transformations and Their Optical Implementation: Part I*, JOSA A10, pp. 1875-1881, 1993.
10. H.M. Ozaktas, and D. Mendlovic, *Fractional Fourier Transformations and Their Optical Implementation: Part II*, JOSA A10, pp. 2522-2531, 1993.
11. H.M. Ozaktas, and D. Mendlovic, *Fourier Transforms of Fractional Orders and Their Optical Interpretation*, Opt. Commun., Vol. 101, pp. 163-165, 1993.
12. A.W. Lohmann, *Image Rotation, Wigner Rotation and the Fractional Fourier Transform*, JOSA A10, pp. 2181-2186, 1993.
13. D. Mendlovic, H.M. Ozaktas, and A.W. Lohmann, *Graded Index Fibers, Wigner Distribution Functions and the Fractional Fourier Transform*, Applied Optics, Vol. 33, pp. 6188-6193, 1994.
14. J. Garcia, D. Mendlovic, Z. Zalevsky, and A.W. Lohmann, *Space Variant Simultaneous Detection of Several Objects Using Multiple Anamorphic Fractional Fourier Transform Filters*, Applied Optics, Vol. 35, pp. 3945-3952, 1996.
15. J.H. Caulfield, D. Mendlovic, and Z. Zalevsky, *Fractional Correlator With Real Time Control of the Space-Invariance Property*, Applied Optics, Vol. 36, pp. 2370-2375, 1997.