Fundamental structure of Fresnel diffraction: natural sampling grid and the fractional Fourier transform

Haldun M. Ozaktas,* Sercan Ö. Arık, and Türker Coşkun

Bilkent University, Department of Electrical Engineering, TR-06800 Bilkent, Ankara, Turkey

**Corresponding author: haldun@ee.bilkent.edu.tr*

Received February 15, 2011; revised May 6, 2011; accepted May 20, 2011;
posted May 23, 2011 (Doc. ID 142599); published June 29, 2011

Fresnel integrals corresponding to different distances can be interpreted as scaled fractional Fourier transformations observed on spherical reference surfaces. We show that by judiciously choosing sample points on these curved reference surfaces, it is possible to represent the diffracted signals in a nonredundant manner. The change in sample spacing with distance reflects the structure of Fresnel diffraction. This sampling grid also provides a simple and robust basis for accurate and efficient computation, which naturally handles the challenges of sampling chirplike kernels. © 2011 Optical Society of America

OCIS codes: 070.2575, 070.2580, 070.2025, 050.1940, 050.5082, 070.0070.

In this Letter we show that, by appropriately choosing sampling points on spherical rather than planar surfaces, it is possible to represent diffracted signals with the minimum possible number of samples. The grid of sample points reflects the structure of Fresnel diffraction and also facilitates fast and accurate computation.

Appropriate choice of sampling points depends not only on propagation parameters, but also on the space and frequency extents of the signals, and involves proper scaling of the input. Unlike some other sampling approaches, this allows representation of the signal nonredundantly and without information loss, using the same number of samples that are required to represent the input field.

For simplicity in presentation, we will work with dimensionless space and frequency coordinates [1] of a single variable. Let $\hat{f}(x)$ and $\hat{F}(\sigma_x)$ denote the space and frequency representation of a signal. We will use $f(u)$ and $F(\mu)$ to denote corresponding functions with dimensionless arguments u and μ:

$$\hat{f}(x) \equiv \frac{1}{\sqrt{s}} f\left(\frac{x}{s}\right), \qquad \hat{F}(\sigma_x) \equiv \sqrt{s} F(s\sigma_x), \tag{1}$$

where s is a scaling parameter with dimensions of length.

The fractional Fourier transform (FRT) is a generalization of the ordinary Fourier transform (FT) with an order parameter a. It is well known that the FRT operation corresponds to a rotation in the space–frequency plane by an angle $a\pi/2$. The FRT $f_a(u)$ of a function $f(u)$ may be defined as [1]

$$f_a(u) = A_a \int_{-\infty}^{\infty} \exp\left[i\pi\left(u^2 \cot\frac{a\pi}{2} - 2uu' \csc\frac{a\pi}{2} + u'^2 \cot\frac{a\pi}{2}\right)\right] f(u')\mathrm{d}u', \tag{2}$$

where $A_a = \sqrt{1 - i\cot(a\pi/2)}$.

The Fresnel integral describes the propagation of light from one transverse plane along the optical axis to another. For the one-dimensional case, the output field $\hat{g}(x)$ is related to the input field $\hat{f}(x)$ by [2]

$$\hat{g}(x) = e^{i2\pi d/\lambda} e^{-i\pi/4} \sqrt{\frac{1}{\lambda d}} \int_{-\infty}^{\infty} \exp\left[\frac{i\pi(x-x')^2}{\lambda d}\right] \hat{f}(x')\mathrm{d}x', \tag{3}$$

where d is the distance of propagation and λ is the wavelength.

It is known that the Fresnel integral can be decomposed into a FRT, followed by magnification, followed by chirp multiplication [1,3–6]:

$$\hat{g}(x) = e^{i2\pi d/\lambda} e^{-ia\pi/4} \sqrt{\frac{1}{sM}} \exp\left(\frac{i\pi x^2}{\lambda R}\right) f_a\left(\frac{x}{sM}\right), \tag{4}$$

where

$$a = \frac{2}{\pi} \arctan\frac{\lambda d}{s^2}, \tag{5}$$

$$M = \sqrt{1 + \frac{\lambda^2 d^2}{s^4}} = \sec\frac{a\pi}{2}, \tag{6}$$

$$R = \frac{s^4 + \lambda^2 d^2}{\lambda^2 d} = d \csc^2\frac{a\pi}{2}. \tag{7}$$

If we choose to observe the diffracted light on a spherical reference surface of radius R, the chirp multiplication can be dispensed with and we simply observe the FRT of the input, magnified by M. (The constant phase terms $e^{i2\pi d/\lambda} e^{-ia\pi/4}$ are not of significance.) Equation. (4) holds true regardless of the choice of s.

We assume that the energy of the signal at the $z = 0$ plane is confined to an ellipse with diameters Δx and $\Delta\sigma_x$ in the space–frequency plane (phase space), in the sense that most of the energy lies within this ellipse. For concreteness, one may employ the Wigner distribution as a space–frequency representation [1], although this is not essential for our development. Δx and $\Delta\sigma_x$ also correspond to the space and frequency extents of the signal. Since a frequency extent of $\Delta\sigma_x$ implies a sampling

 © 2011 Optical Society of America

interval of $1/\Delta\sigma_x$, we would need $N = \Delta x/(1/\Delta\sigma_x) = \Delta x\Delta\sigma_x$ samples to characterize the signal in terms of its samples, a quantity also referred to as the space-bandwidth product. In dimensionless coordinates the diameters of the ellipse become $\Delta x/s$ and $s\Delta\sigma_x$.

We now determine the spatial extent of the diffracted signal observed on the Fresnel output plane by using the fact that [1,7,8] Fresnel propagation shears the Wigner distribution $\hat{W}_f(x,\sigma_x)$ into the form $\hat{W}_f(x+\lambda d\sigma_x,\sigma_x)$. If the original Wigner distribution occupied an ellipse with diameters Δx and $\Delta\sigma_x$ in the space–frequency plane, the sheared Wigner distribution will exhibit a spatial extent of $\Delta x'' = \sqrt{(\lambda d\Delta\sigma_x)^2 + \Delta x^2}$ and a frequency extent of $\Delta\sigma_x'' = \Delta\sigma_x$. The output spatial extent $\Delta x''$ will begin to be significantly larger than the input spatial extent Δx beyond the distance $d = \Delta x/(\lambda\Delta\sigma_x)$. We refer to this distance as the "knee-of-the-curve" point along the z axis. This distance is easy to interpret if we note that the angular divergence of the input signal is $\Delta\theta \approx \lambda\Delta\sigma_x$, so that the spatial spreading of the signal after propagating a distance d will be $d\Delta\theta = \lambda d\Delta\sigma_x$. This begins to exceed Δx at $d = \Delta x/(\lambda\Delta\sigma_x)$.

We will now rederive the spatial extent of the diffracted signal by working our way through Eq. (4). In dimensionless coordinates, the diameters of the original ellipse will be $\Delta x/s$ and $s\Delta\sigma_x$. Fractional Fourier transformation of order a will rotate this ellipse by an angle $\alpha = a\pi/2$, producing an ellipse with dimensionless spatial extent $\sqrt{(s\Delta\sigma_x\sin\alpha)^2 + (\Delta x\cos\alpha/s)^2}$. Going back to dimensional coordinates and multiplying this with the parameter M gives us the spatial extent of the diffracted signal on the spherical reference surface:

$$\begin{aligned}\Delta x' &= M\sqrt{(s^2\Delta\sigma_x\sin\alpha)^2 + (\Delta x\cos\alpha)^2}\\ &= \sqrt{(\lambda d\Delta\sigma_x)^2 + \Delta x^2},\end{aligned} \tag{8}$$

where we have inserted the expressions for α and M from Eqs. (5) and (6). We observe that the final expression obtained for $\Delta x'$ does not depend on s and is exactly the same as $\Delta x''$ derived in the previous paragraph using the Wigner distribution. The spatial extent on the Fresnel output plane is equal to that on the spherical reference surface since there is only a multiplicative factor between these surfaces.

Now, we turn our attention to the spatial frequency extent $\Delta\sigma_x'$ of the diffracted signal on the spherical reference surface. Again from the geometry of the rotated ellipse, we find that the dimensionless frequency extent is $\sqrt{(s\Delta\sigma_x\cos\alpha)^2 + (\Delta x\sin\alpha/s)^2}$, which after going back to dimensional coordinates and dividing by M, gives us the spatial frequency extent of the diffracted signal on the spherical reference surface:

$$\Delta\sigma_x' = \cos\alpha\sqrt{(\Delta\sigma_x\cos\alpha)^2 + (\Delta x\sin\alpha/s^2)^2}. \tag{9}$$

Notice that the spatial frequency extent $\Delta\sigma_x'' = \Delta\sigma_x$ on the Fresnel output plane would be different from this, due to the final chirp multiplication.

The space-bandwidth product $N' = \Delta x'\Delta\sigma_x'$ on the spherical reference surface can be calculated by using Eqs. (8) and (9). N' is the minimum number of Nyquist samples required to characterize the diffracted signal on the spherical reference surface. It can be shown that N' is always greater than or equal to $N = \Delta x\Delta\sigma_x$, with equality if and only if $s = \sqrt{\Delta x/\Delta\sigma_x}$ (except in the special cases $d = 0$ and $d = \infty$). With this choice of s, on the reference surface we have $\Delta x' = M\Delta x$ and $\Delta\sigma_x' = \Delta\sigma_x/M$ so that their product is equal to the original space-bandwidth product $N = \Delta x\Delta\sigma_x$.

We can also write the space-bandwidth product on the Fresnel output plane as $N'' = \Delta x''\Delta\sigma_x'' = \Delta x'\Delta\sigma_x$. We always have $N' \geq N$ and $N'' \geq N$. Whether $N'' > N'$ holds depends on whether $\Delta\sigma_x > \Delta\sigma_x'$ holds, which in turn can be shown to depend on whether $2s^4 > (\Delta x/\Delta\sigma_x)^2 - \lambda^2 d^2$. In particular, this condition holds when $s = \sqrt{\Delta x/\Delta\sigma_x}$ so that for this value of s we have $N'' > N' = N$.

Fresnel propagation results in horizontal shearing in the space–frequency plane, which increases the spatial extent, but does not decrease the bandwidth. This increases the space-bandwidth product and number of samples at the output plane, despite the fact that Fresnel transformation is a unitary and information-preserving operation. Moreover, we have seen that this remains the case, even if we take our output reference surface to be the spherical surface with radius R, unless s is chosen equal to $s = \sqrt{\Delta x/\Delta\sigma_x}$. This choice of s equates the spatial extent Δx and the frequency extent $\Delta\sigma_x$ in the dimensionless space–frequency plane, where they both become equal to $\sqrt{\Delta x\Delta\sigma_x} = \sqrt{N}$. In this case, the original ellipse becomes a circle with this diameter. Since FRT corresponds to rotation in the space–frequency plane [1,9], this circular region will not change shape or size after an FRT operation. Therefore, when we go back to dimensional coordinates, the space extent will merely be M times the space extent of the original signal, and the frequency extent will merely be the frequency extent of the original signal divided by M. Consequently, the samples on the spherical reference surface will be spaced $M/\Delta\sigma_x$ apart covering an extent of $M\Delta x$. The number of samples N' needed to characterize the signal will be $M\Delta x/(M/\Delta\sigma_x) = \Delta x\Delta\sigma_x = N$.

Recall that we have a magnified FRT relationship between the input plane and the output spherical reference surface. With the choice of s above, the only effect of the magnification on sampling is to magnify the spacing and extent of the samples by M. Therefore, the N samples, spaced $M/\Delta\sigma_x$ apart on the spherical reference surface, constitute a "natural sampling grid" for diffraction calculations. These samples are sufficient to reconstruct the continuous diffracted field in the Nyquist–Shannon sense. The same number of samples would not have been sufficient on other reference surfaces or with other values of s. Thus $s = \sqrt{\Delta x/\Delta\sigma_x}$ may be referred to as the "natural scale parameter."

It is known that the discrete FRT approximately maps the samples of a function to the samples of its FRT in the same sense that the ordinary discrete FT does for the ordinary FT [10–13]. Therefore, the values of the diffracted field at the natural sampling grid points can be well approximated by the discrete FRT of the samples

of the input field. With this observation, *diffraction computation is seen to be reduced to discrete FRT*.

We observe that all three of Eqs. (5)–(7) exhibit significant change of behavior around $d \sim s^2/\lambda$. (a changes from linear increasing to saturation, M changes from saturation to linear increasing, and R changes from decreasing to increasing.) Thus, all three parameters share a common knee-of-the-curve. We now compare this knee-of-the-curve with that encountered during our calculation of $\Delta x''$ with the Wigner distribution. Equating the two knee-of-the-curves as $\Delta x/(\lambda \Delta \sigma_x) = s^2/\lambda$, we find that both approaches yield the same knee-of-the-curve location when $s = \sqrt{\Delta x/\Delta \sigma_x}$, which is the same special value of s found before.

We note that the parameters M, a, and R are independent of the input signal. Their dependence on wavelength λ, distance d, and scale s define the spherical surfaces that reflect the physical structure of diffraction. They are parameters which characterize the system we are investigating. On the other hand, the parameters Δx, $\Delta \sigma_x$, and, hence, s characterize the signal and are not related to the physics of diffraction. They are parameters characterizing the signal passing through our system. The natural sampling grid has two ingredients: the spherical surfaces on which the samples are taken and the sample spacings on the spherical surfaces. The general structure of the spherical surfaces is determined by the nature of diffraction as a system, through the parameters M, a, and R. The set of signals determine the parameter s, which in turn matches the spherical surfaces to the signals, and determines the sample spacings. Since the spherical surfaces and the sample spacings together completely define the grid, we may also say that choice of s matches the grid to the set of signals. The fact that the same value of s matches the common knee-of-the-curve of the structure-defining parameters M, a, and R, with the knee-of-the-curve of the diffracting signal, reinforces this observation.

Sampling for purposes of digital computation of the Fresnel integral can be approached in a number of ways [14–21]. We have seen that, despite the spreading of light and the space–frequency shearing behavior of the Fresnel integral, accurate representation and efficient computation does not require an increase in the number of samples. Our method employs the minimal number of samples at the input and output and fast $\sim N \log N$ computation between these samples is possible [10]; furthermore, the accuracy of computation is the same as that in computing the FT with the fast Fourier transform (FFT) [6]. In this sense, representation and computation based on the natural sampling grid is optimal. We also note that our formulation does not depend on the algorithm used for computing the FRT. Therefore, improved algorithms or enhancements (such as in [22]) can be used without any modification of our formulation.

In conclusion, by defining the output on spherical reference surfaces, and with appropriate scaling, we can recover the output field from the same number of samples as the input. This minimum number of samples would not have been sufficient with other reference surfaces or with other values of s, despite the fact that the Fresnel integral is unitary and preserves space–frequency area and number of degrees of freedom.

H. M. Ozaktas acknowledges partial support of the Turkish Academy of Sciences.

References

1. H. M. Ozaktas, Z. Zalevsky, and M. A. Kutay, *The Fractional Fourier Transform with Applications in Optics and Signal Processing* (Wiley, 2001).
2. J. W. Goodman, *Introduction to Fourier Optics*, 2nd ed. (McGraw-Hill, 1996).
3. H. M. Ozaktas and D. Mendlovic, J. Opt. Soc. Am. A **12**, 743, 1995.
4. H. M. Ozaktas and M. F. Erden, Opt. Commun. **143**, 75 (1997).
5. H. M. Ozaktas, A. Koc, I. Sari, and M. A. Kutay, Opt. Lett. **31**, 35 (2006).
6. A. Koc, H. M. Ozaktas, C. Candan, and M. A. Kutay, IEEE Trans. Signal Process. **56**, 2383 (2008).
7. M. J. Bastiaans, J. Opt. Soc. Am. **69**, 1710 (1979).
8. M. J. Bastiaans, *The Wigner Distribution: Theory and Applications in Signal Processing* (Elsevier Science, 1997), 375–426.
9. H. M. Ozaktas, B. Barshan, D. Mendlovic, and L. Onural, J. Opt. Soc. Am. A **11**, 547 (1994).
10. H. M. Ozaktas, O. Arikan, M. A. Kutay, and G. Bozdagi, IEEE Trans. Signal Process. **44**, 2141 (1996).
11. C. Candan, M. A. Kutay, and H. M. Ozaktas, IEEE Trans. Signal Process. **48**, 1329 (2000).
12. F. S. Oktem and H. M. Ozaktas, IEEE Signal Process. Lett. **16**, 727 (2009).
13. F. S. Oktem and H. M. Ozaktas, J. Opt. Soc. Am. A **27**, 1885 (2010).
14. L. Onural, Appl. Opt. **39**, 5929 (2000).
15. D. Mendlovic, Z. Zalevsky, and N. Konforti, J. Mod. Opt. **44**, 407 (1997).
16. D. Mas, J. Garcia, C. Ferreira, L. M. Bernardo, and F. Marinho, Opt. Commun. **164**, 233 (1999).
17. B. M. Hennelly and J. T. Sheridan, J. Opt. Soc. Am. A **22**, 917 (2005).
18. B. M. Hennelly and J. T. Sheridan, J. Opt. Soc. Am. A **22**, 928 (2005).
19. J. J. Healy, B. M. Hennelly, and J. T. Sheridan, Opt. Lett. **33**, 2599 (2008).
20. J. Healy and J. T. Sheridan, J. Opt. Soc. Am. A **27**, 21 (2010).
21. J. Healy and J. Sheridan, Opt. Lett. **35** 947 (2010).
22. X. Yang, Q. Tan, X. Wei, Y. Xiang, Y. Yan, and G. Jin, J. Opt. Soc. Am. A **21**, 1677 (2004).