

OPTIMAL FILTERING IN FRACTIONAL FOURIER DOMAINS

M. Alper Kutay, Haldun M. Ozaktas, Levent Onural, and Orhan Arikan

Bilkent University, Department of Electrical Engineering
TR-06533 Bilkent, Ankara, Turkey

ABSTRACT

The ordinary Fourier transform is suited best for analysis and processing of time-invariant signals and systems. When we are dealing with time-varying signals and systems, filtering in fractional Fourier domains might allow us to estimate signals with smaller minimum-mean-square error (MSE). We derive the optimal fractional Fourier domain filter that minimizes the MSE for given non-stationary signal and noise statistics, and time-varying distortion kernel. We present an example for which the MSE is reduced by a factor of 50 as a result of filtering in the fractional Fourier domain, as compared to filtering in the conventional Fourier or time domains. We also discuss how the fractional Fourier transformation can be computed in $O(N \log N)$ time, so that the improvement in performance is achieved with little or no increase in computational complexity.

1. INTRODUCTION

The fractional Fourier transform was defined mathematically by Namias [2], and McBride and Kerr [3]. Its analog optical implementation was discussed in [4, 5] and recently a fast digital $O(N \log N)$ time algorithm has also been developed. (The outline of this algorithm will be given in the appendix.) Other work in this area includes [1] and [8] in which many properties are derived and applications are suggested.

Several applications of the fractional Fourier transform have been suggested or explored to varying degrees. These include optical diffraction and beam propagation, optical signal processing, quantum optics, phase retrieval, signal detection, pattern recognition, noise representation, time-variant filtering and multiplexing, data compression, study of space/time-frequency distributions [1, 2, 3, 4, 5, 7, 8, 9].

The a th order fractional Fourier transform of a function, denoted by $\mathcal{F}^a[f](x)$, is defined as [1, 3]:

$$(\mathcal{F}^a[f])(x) = \int_{-\infty}^{\infty} B_a(x, x') f(x') dx',$$

$$B_a(x, x') = A_\phi e^{i\pi(x^2 \cot \phi - 2xx' \csc \phi + x'^2 \cot \phi)}, \quad (1)$$

where $A_\phi = (|\sin \phi|)^{-1/2} \exp\left(i\left(\frac{\pi \operatorname{sgn}(\sin \phi)}{4} - \frac{\phi}{2}\right)\right)$, and $\phi \equiv a\pi/2$. The kernel $B_a(x, x')$ approaches $\delta(x - x')$ or $\delta(x + x')$ when a approaches 0 or ± 2 , respectively. The definition is easily extended outside the interval $[-2, 2]$ since \mathcal{F}^4 is the identity operation [3].

Some essential properties of the fractional Fourier transform are: i) It is linear. ii) The first order transform ($a = 1$)

corresponds to the common Fourier transform. iii) It is additive in index, $\mathcal{F}^{a_1} \mathcal{F}^{a_2} f = \mathcal{F}^{a_1+a_2} f$.

An important property of the fractional Fourier transform is its relation to the Wigner distribution. Let $W(x, \nu)$ denote the Wigner distribution of the signal f . It is well known that the projection of $W(x, \nu)$ onto the x axis gives the magnitude squared of the frequency-domain representation, and the projection onto the ν axis gives the magnitude squared of the time-domain representation of the signal. The property in question states that the projection onto an axis making angle ϕ with the x -axis gives the magnitude squared of the $a = \frac{2\phi}{\pi}$ th order fractional Fourier transform of the function. This property can be formulated as iv) $\mathcal{R}_\phi[W_f(x, \nu)] = |\mathcal{F}^a[f]|^2$, where the operator \mathcal{R}_ϕ is the Radon transform evaluated at the angle ϕ . Other properties may be found in [1, 2, 3, 4, 5].

The ordinary Fourier transform is suited best for analysis and processing of time-invariant signals. When we are dealing with time-varying signals and systems, filtering in fractional Fourier domains might allow us to estimate signals with smaller minimum-mean-square error. In a similar spirit, multiplexing in fractional Fourier domains allows signals whose time-frequency distribution is irregular to be packed more efficiently in a given channel [1]. In this work we concentrate on filtering in fractional domains.

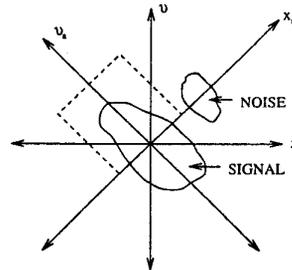


Figure 1: Noise separation in a=0.5th domain

There are a variety of time-varying digital signal processing algorithms based on mixed time-frequency signal representations. (see for example [10] and the references given there.) Many signal processing operations have been implemented by modifying these representations in some time-varying manner. However the majority of these suffer from especially two problems: i.) Most of the representations, such as the Wigner distribution, are not linear so that there exists interference terms for multicomponent signals ii.) The modified time-frequency representation may not be

the time-frequency representation of any signal, so that ad hoc approximations must be introduced.

Recently we have discussed how various space-variant operations can be performed by multiplying with a filter function in a fractional Fourier domain [1]. Signals with significant overlap in both the time and frequency domains may have little or no overlap in a fractional Fourier domain. To understand the basic idea in the simplest possible terms, consider the simple example shown in figure 1, where the Wigner distributions of a desired signal and undesired distortion term are shown. By virtue of property iv), we observe that they overlap in both the $a=0$ th and $a=1$ st domains (consider the projections on the x and ν axes), but they do not overlap in the $a = 0.5$ th domain. Thus we can eliminate undesired signal components by using a simple unit amplitude mask in the $a = 0.5$ th domain.

2. PROBLEM STATEMENT AND SOLUTION

We consider a commonly used signal observation model,

$$y = \mathcal{H}(x) + n, \quad (2)$$

where $\mathcal{H}(\cdot)$ is a linear system that degrades the input (desired) signal x , and n is an additive noise term. Our aim is to filter the observed signal y to minimize the effects of degradation and noise. Our criteria for judging the effectiveness of the restoration is the mean square error (MSE). We will assume that as a prior knowledge we know the correlation functions $R_{xx}(t, t') = E[x(t)x(t')]$, $R_{nn}(t, t') = E[n(t)n(t')]$ of the input process and noise. We will further assume that the noise is independent of the input and it is zero mean for all time, i.e. $E[n(t)] = 0 \forall t$. Under these assumptions we can also compute the cross correlation function $R_{xy}(t, t') = E[x(t)y(t')]$ of the input process x and the output process y , and the correlation function $R_{yy}(t, t') = E[y(t)y(t')]$ of the output process by virtue of equation 2. Now, consider a linear estimate of the form $\hat{x} = \mathcal{G}(y)$. The corresponding MSE is

$$\sigma_e^2 = E[\|x - \hat{x}\|^2],$$

where $\|\cdot\|^2$ denotes the L_2 norm of the signal. For the above definition of the MSE to be meaningful, we require the input processes and noise to be square-integrable. This in turn necessitates the input processes and noise to be non-stationary. For stationary processes the MSE is simply defined as the expected value of the absolute value of the difference term.

For the time-invariant degradation model \mathcal{H} with stationary processes x and n , the linear operator \mathcal{G}_{opt} that minimizes the MSE corresponds to the optimal Wiener filter. In this case, the required operation turns out to be of the convolution type, so that one can effectively implement it with a filter in the conventional Fourier domain. For an arbitrary degradation model or for non-stationary processes, finding the general linear filter that minimizes the MSE is difficult. More importantly, in general the resulting linear filter will be time-variant (not expressible as a convolution), so that it cannot be implemented in $N \log N$ time using FFT based techniques.

We consider filters in the fractional Fourier domain, that is, the estimated signal is related to the observation as,

$$\hat{x}(l) = \mathcal{F}^{-a}(\mathcal{G} \mathcal{F}^a(y))(t)$$

$$= \int_{-\infty}^{\infty} B_{-a}(t, t') g(t') \int_{-\infty}^{\infty} B_a(t', t'') y(t'') dt'' dt'$$

This filter corresponds to multiplication with a function $g(\cdot)$ in the a th domain. If $a = 1$, this corresponds to Fourier domain filtering as in the case of Wiener filtering. With this form of estimation filter, the minimization problem considered in this paper can be formally written as:

$$g_{opt}(t) = \arg \min_g \sigma_e^2 \quad (3)$$

The class of fractional Fourier domain filters is still a subclass of the class of all linear filters, so the linear filter we find will not be optimal among all linear filters. However, it is a much broader class than time-invariant Fourier domain filters, and in some cases it is possible to reduce the MSE for time-varying degradation models or non stationary processes, as compared to filtering in the conventional Fourier domain. The resulting filter can be implemented efficiently since efficient optical and digital implementations of the fractional Fourier transform are known.

We will attempt to solve the minimization problem defined by equation (3) using calculus of variations method. We define the cost function $J = \sigma_e^2$. J varies with the choice of $g(\cdot)$ since $\hat{x}(t)$ varies. This functional J is to be minimized with respect to $g(\cdot)$. We substitute $g(\cdot) = g_o(\cdot) + \alpha \delta g_o(\cdot)$ where $\alpha = \alpha_{re} + i \alpha_{im}$ is a complex scalar parameter, $g_o(\cdot)$ is the optimum filter and $\delta g_o(\cdot)$ is an arbitrary perturbation term and we fix $g_o(\cdot)$ and $\delta g_o(\cdot)$. With this substitution $\hat{x}(t)$ and also J vary with α for each fixed $\delta g_o(\cdot)$,

$$\begin{aligned} \hat{x}(t, \alpha) &= \int_{-\infty}^{\infty} B_{-a}(t, t') (g_o(t') + (\alpha_{re} + i \alpha_{im}) \delta g_o(t')) \\ &\quad \times \int_{-\infty}^{\infty} B_a(t', t'') y(t'') dt'' dt' \\ J(\alpha) &= E \left[\int_{-\infty}^{\infty} (x(t) - \hat{x}(t, \alpha)) (x(t) - \hat{x}(t, \alpha))^* dt \right] \end{aligned}$$

The necessary conditions for the optimum value of J are [11]:

$$\left. \frac{\partial J(\alpha)}{\partial \alpha_{re}} \right|_{\alpha=0} = 0, \quad \left. \frac{\partial J(\alpha)}{\partial \alpha_{im}} \right|_{\alpha=0} = 0. \quad (4)$$

It can be shown that after some algebraic manipulations, the necessary conditions for optimality of J imply,

$$\begin{aligned} E \left[\int_{-\infty}^{\infty} \delta g_o(t') \left(\int_{-\infty}^{\infty} (x(t) - \hat{x}(t, 0))^* \right. \right. \\ \left. \left. B_{-a}(t, t') \int_{-\infty}^{\infty} B_a(t', t'') y(t'') dt'' dt \right) dt' \right] = 0 \end{aligned} \quad (5)$$

Since $\delta g_o(t')$ is an arbitrary term and equation (5) is true for all $\delta g_o(t')$, the inner integral should be equal to zero. Using this condition one can find the optimal filter function:

$$g_o(t') = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B_a(t, t') B_{-a}(t', t'') R_{xy}(t, t'') dt'' dt}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B_{-a}(t', t'') B_a(t', t'') R_{yy}(t', t'') dx'' dt''} \quad (6)$$

where the correlation functions are defined as $R_{xy}(t, t'') = E[x(t)y^*(t'')]$ and $R_{yy}(t, t'') = E[y(t)y^*(t'')]$.

In the above formulation, it was assumed that the value of a is given. To find the optimal choice of a , we can substitute the optimum filter function (cf. (6)) into the definition of MSE and apply a minimization algorithm with respect to a . In practice, it might be easier to find the optimum value of a by computing the MSE as a function of a , and choosing that value of a which results in the smallest MSE.

Discrete time formulation. We now present the discrete time formulation, beginning directly from the discrete version of Eq. 2. In this case \mathcal{H} is a matrix and x , n , and y are vectors. They will be denoted by underlined bold-face letters from now on. The discrete fractional Fourier transform takes on the form of a matrix multiplication (see appendix), just as the ordinary discrete Fourier transform. The error for the discrete case is given by:

$$\sigma_e^2 = \frac{1}{N} E [(\underline{x} - \hat{\underline{x}})^H (\underline{x} - \hat{\underline{x}})] \quad (7)$$

where N is the size of the input vector \underline{x} and $\hat{\underline{x}} = \underline{F}^{-a} \underline{\Lambda}_g \underline{F}^a \underline{y}$. Here $\underline{\Lambda}_g$ is a diagonal matrix whose diagonal consists of elements of the vector \underline{g} , and \underline{F}^a is the fractional Fourier transformation matrix. The inverse transformation matrix is $\underline{F}^{-a} = (\underline{F}^a)^H$, since the fractional Fourier transformation is unitary. Here $(\cdot)^H$ denotes conjugate transpose.

This estimate corresponds to a multiplicative filter in the a th fractional Fourier domain. As in the continuous time case, if $a = 1$, \underline{F}^a is simply the DFT matrix and we obtain the common Fourier domain filter.

In order to find the optimal filter that minimizes the MSE, we define the cost function $J_d = \sigma_e^2$ and note that it varies with the choice of the vector $\underline{g} = [g_1 \ g_2 \ \dots \ g_N]$. Since in general the filter vector \underline{g} can be complex, we can write its components as $g_i = g_{i, re} + i g_{i, im}$. The optimum value of the cost function J_d is obtained from the conditions,

$$\frac{\partial J_d}{\partial g_{j, re}} = 0, \quad \text{and} \quad \frac{\partial J_d}{\partial g_{j, im}} = 0 \quad j = 1, \dots, N \quad (8)$$

It can be shown that the elements of the optimum filter function are given by,

$$g_{opt, j} = \frac{\underline{f}_j^{aH} \underline{R}_{xy} \underline{f}_j^a}{\underline{f}_j^{aH} \underline{R}_{yy} \underline{f}_j^a} \quad j = 1, \dots, N \quad (9)$$

where $\underline{R}_{xy} = E [\underline{x} \underline{y}^H]$, $\underline{R}_{yy} = E [\underline{y} \underline{y}^H]$ are the correlation matrices and \underline{f}_j^{aH} is the j th row of the matrix \underline{F}^a .

This last equation provides the solution of our minimization problem in the discrete time setting. We note that this result is fully analogous to the solution obtained in equation (6) for the continuous time case.

3. EXAMPLE

In this section, computer simulations that illustrate the applications and performance of our filtering scheme will be given. We consider a degradation model that corresponds to a time-varying low-pass filter whose bandwidth increases linearly in time. The input process is a Gaussian function which deterministic except for a random amplitude. The

noise process is finite duration bandpass noise which is modulated with an exponential quadratic chirp function ($e^{-i\pi t^2}$) so that its center frequency increases linearly with time. In Fig. 2 (a), we have shown the normalized MSE plot for different values of a . From the plot we find that the minimum MSE is achieved for $a = 0.5$. We have plotted the real part of the optimum filter function for $a = 0.5$ in Fig. 2 (b) and the imaginary part in (c). We note that the imaginary part can be set to zero without effecting the final results. Fig. 3 (a) and (b) shows realizations of the input and output processes (only real part is plotted) respectively. Resulting estimates for these realizations are plotted in Fig. 3 (c). The MSE is 0.001 for $a = 0.5$ while it is 0.063 for $a = 1$.

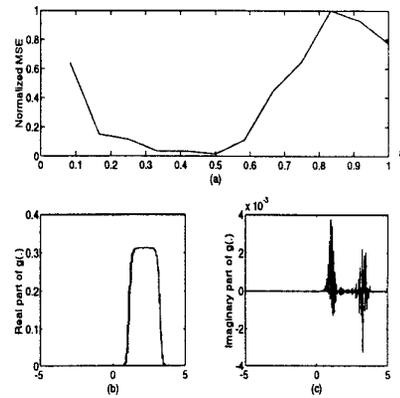


Figure 2: (a) MSE vs a , (b) Real part of optimal filter in $a=0.5$ th domain (c) Imaginary part

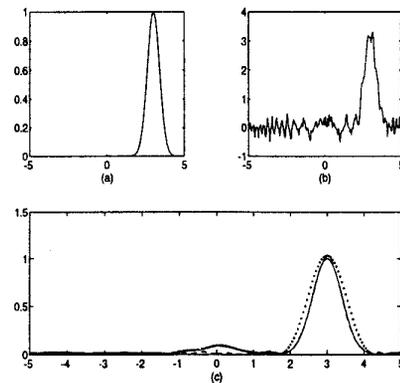


Figure 3: (a) Realization of input process x , (b) Realization of process y (c) Comparison of the estimate \hat{x} in the 0.5th domain (dashed), the estimate \hat{x} in the 1st domain (dotted) and the x (solid)

4. CONCLUSION

Filtering in fractional Fourier domains may enable significant reduction of the MSE compared to ordinary Fourier domain filtering. This reduction comes at essentially no additional cost, since the fractional Fourier transform has an

$O(N \log N)$ algorithm. We have presented a mathematical formulation and solution of this problem analogous to the formulation and solution of the the optimal Wiener filtering problem.

Filtering in fractional domains will work better for certain kinds of distortions and signal and noise statistics in comparison to others. The presence of time-varying distortion and non-stationary statistics suggest that the fractional transform may be of use, but do not guarantee significant improvements in every case.

The method can be greatly improved by filtering in not one, but several consecutive fractional Fourier domains. This not only allows one to handle a much wider variety of signals, but may also create the potential to approximate the most general optimal linear filter

5. APPENDIX

In this appendix, we present a fast algorithm for the implementation of the fractional Fourier transformation. The definition of the fractional Fourier transformation can be cast in the form,

$$(\mathcal{F}^\alpha[f])(x) = A_\phi e^{i\pi\alpha x^2} \int_{-\infty}^{\infty} e^{-i2\pi\beta x x'} \left[e^{i\pi\alpha x'^2} f(x') \right] dx', \quad (10)$$

where $\alpha = \cot \phi$ and $\beta = \csc \phi$. The last term inside the integral corresponds to a chirp modulation of the signal $f(\cdot)$. This modulation results in a vertical shear in the Wigner domain since the Wigner distribution of a chirp signal is a line delta [6].

It will be assumed that the Wigner distribution of $f(\cdot)$ is negligibly small outside a circle of diameter Δx centered around the origin. (Δx must be chosen large enough to satisfy this assumption. Normalizing the axes so that the spread of the Wigner distribution is approximately identical along both axes will allow smaller choices of Δx , which will reduce computational complexity. These are essentially the same considerations that go onto approximating a continuous Fourier transform with the DFT.) Under this assumption, and by limiting ϕ to the interval $\frac{\pi}{4} \leq \phi \leq \frac{\pi}{2}$, the amount of vertical shear in the Wigner domain resulting from chirp modulation is bounded by $\frac{\Delta x}{2}$. This ensures that the modulated function $e^{i\pi\alpha x'^2} f(x')$ is bandlimited to Δx in the frequency domain. This allows $f(\cdot)$ to be represented by Shannon's interpolation formula as follows:

$$e^{i\pi\alpha x'^2} f(x') = \sum_{n=-N}^N e^{i\pi\alpha(\frac{n}{2\Delta x})^2} f\left(\frac{n}{2\Delta x}\right) \text{sinc}\left(2\Delta x\left(x' - \frac{n}{2\Delta x}\right)\right), \quad (11)$$

where N is the largest integer smaller than $(\Delta x)^2$. The summation ranges from $-N$ to N since $f(x')$ is assumed to be negligible outside that interval. By using Eq. 11 and Eq. 10 and changing the order of integration and summation we obtain:

$$(\mathcal{F}^\alpha[f])(x) = \frac{A_\phi}{2\Delta x} \sum_{n=-N}^N e^{i\pi\alpha x^2} e^{-i2\pi\beta x \frac{n}{2\Delta x}} e^{i\pi\alpha(\frac{n}{2\Delta x})^2} \times f\left(\frac{n}{2\Delta x}\right) \text{rect}\left(\frac{\beta x}{2\Delta x}\right) \quad (12)$$

where we made use of $\int_{-\infty}^{\infty} e^{-i2\pi\beta x x'} \text{sinc}\left(2\Delta x\left(x' - \frac{n}{2\Delta x}\right)\right) dx' = e^{-i2\pi\beta x \frac{n}{2\Delta x}} \frac{1}{2\Delta x} \text{rect}\left(\frac{\beta x}{2\Delta x}\right)$; Then the samples of the transformed function are obtained by:

$$(\mathcal{F}^\alpha[f])\left(\frac{m}{2\Delta x}\right) = \frac{A_\phi}{2\Delta x} \sum_{n=-N}^N e^{i\pi\frac{(\alpha m^2 - 2\beta m n + \alpha n^2)}{(2\Delta x)^2}} \times f\left(\frac{n}{2\Delta x}\right) \text{rect}\left(\frac{\beta m}{4N}\right) \quad (13)$$

Direct computation of the above matrix-vector product would require $O(N^2)$ multiplications. An $O(N \log N)$ algorithm for the computation of this form can be readily obtained. For this purpose, we put Eq. 13 into the following form after some algebraic manipulations:

$$(\mathcal{F}^\alpha[f])\left(\frac{m}{2\Delta x}\right) = \frac{A_\phi}{2\Delta x} e^{i\pi(\alpha-\beta)(\frac{m}{2\Delta x})^2} \text{rect}\left(\frac{\beta m}{4N}\right) \times \sum_{n=-N}^N e^{i\pi\beta(\frac{m-n}{2\Delta x})^2} e^{i\pi(\alpha-\beta)(\frac{n}{2\Delta x})^2} f\left(\frac{n}{2\Delta x}\right)$$

It can be recognized that the summation corresponds to a convolution of $e^{i\pi\beta(\frac{n}{2\Delta x})^2}$ with the chirp modulated function $f(\cdot)$. The standard *FFT* can be used to compute this convolution in $O(N \log N)$ time. The output samples are then obtained by another chirp modulation. Hence the overall time complexity is $O(N \log N)$. The factor $\text{rect}\left(\frac{\beta m}{4N}\right)$ only limits the length of the output signal.

This algorithm is valid only for $\frac{\pi}{4} \leq \phi \leq \frac{\pi}{2}$ and $-\frac{\pi}{2} \leq \phi \leq -\frac{\pi}{4}$ ($0.5 \leq a \leq 1$ and $-1 \leq a \leq -0.5$). However using a basic property of the fractional Fourier transform we can extend this range to all values of a . For example: $\mathcal{F}^a = \mathcal{F}^{a-1+1} = \mathcal{F}^{a-1} \mathcal{F}^1$, $0 \leq a \leq 0.5$ etc.

The above presented method is not the only one which results in $O(N \log N)$ computation. However, it has many desirable features compared to other methods of decomposing the chirp integral and leads to good numerical behavior.

6. REFERENCES

- [1] H. M. Ozaktas, B. Barshan, D. Mendlovic, and L. Onural. *J. Opt. Soc. Am. A*, 11:547-559, 1994.
- [2] V. Namias. *J. Inst. Maths Applics.*, 25:241-245, 1980.
- [3] A. C. McBride and F. H. Kerr. *IMA J. Appl. Math.*, 39:159-175, 1987.
- [4] H. M. Ozaktas and D. Mendlovic. *Opt. Commun.*, 101:163-169, 1993.
- [5] A. W. Lohmann. *J. Opt. Soc. Am. A*, 10:2181-2186, 1993.
- [6] T. A. C. M. Claasen and W. F. G. Mecklenbraucker. *Philips Journal of Research*, 35:217-250, 1980.
- [7] H. M. Ozaktas, D. Mendlovic, "Fractional Fourier Optics," to appear in *JOSA A*.
- [8] L. M. Almeida. *Proceedings of ICASSP 1993*, III-257.
- [9] J. R. Fonollosa and C. L. Nikias. *Proceedings of ICASSP 1994*, IV-301-IV-304.
- [10] B. E. A. Saleh and N. S. Subotic. *IEEE Trans. ASSP.*, vol.33, no.6, pp.1479-1485, Dec. 1985.
- [11] G.M. Ewing, *Calculus of Variations with Applications*. Dover publications, 1985.