

ORTHONORMAL FINITE RIDGELET TRANSFORM FOR IMAGE COMPRESSION

Minh N. Do[†] and Martin Vetterli^{†§}

[†]Laboratory for Audio-Visual Communications

Swiss Federal Institute of Technology Lausanne (EPFL), CH-1015 Lausanne, Switzerland

[§]Department of EECS, University of California at Berkeley, Berkeley CA 94720, USA

Email: {Minh.Do, Martin.Vetterli}@epfl.ch

ABSTRACT

A finite implementation of the ridgelet transform is presented. The transform is invertible, non-redundant and achieved via fast algorithms. Furthermore we show that this transform is orthogonal hence it allows one to use non-linear approximations for the representation of images. Numerical results on different test images are shown. Those results conform with the theory of the ridgelet transform in the continuous domain – the obtained representation can represent efficiently images with linear singularities. Thus it indicates the potential of the proposed system as a new transform for coding of images.

1. INTRODUCTION

Many image processing tasks take advantage of *sparse* representations of image data where maximum information is packed into a small number of samples. Typically, these representations are achieved via non-redundant and invertible transforms. Currently, the most popular choice for this purpose is wavelet transforms [1, 2].

The main power of wavelets comes from the fact that they are well adapted to changes or singularities that are commonly found in real-life signals. In multidimensional cases, most often *tensor product* wavelets or *separable* schemes are employed. As a generalization from the case, wavelets in higher dimensions are well adapted for *pointlike* phenomena. But this is the *only* type of singularities that wavelets can efficiently represent. This problem was raised recently by Candes and Donoho [3, 4]. Those authors argued that in higher dimensions, there are many other kinds of intermittency such as singularities along lines and curves which wavelets do not deal with efficiently. In order to overcome this weakness, they have proposed new systems of representations namely *ridgelets* which can effectively deal with *linelike* phenomena in 2-D.

Much of the work in ridgelets has been concentrated in the *continuous* \mathbb{R}^2 space. However for practical applications, *discrete-time* space implementations of the ridgelet transform and its reconstruction on a finite \mathbb{Z}^2 plane remain an open issue. Due to the radial nature of ridgelets, straightforward implementations based on discretization of continuous formulae would require interpolation in the polar coordinate and thus the resulting transforms would be

either redundant or can not be perfectly reconstructed. In this paper, we propose one implementation of the ridgelet transform for digital images that achieves both nonredundant and invertible requirements. Numerical results on non-linear approximation comparison with other transforms on various test images are presented.

2. CONTINUOUS RIDGELET TRANSFORM

We start by briefly reviewing the continuous ridgelet transform (CRT) and draw its connection with the continuous wavelet transform. Given an integrable bivariate function $f(\mathbf{x})$, the CRT in \mathbb{R}^2 can be defined as follows [3, 4]

$$RI_f(a, b, \theta) = \int_{\mathbb{R}^2} \psi_{a,b,\theta}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}, \quad (1)$$

where the ridgelets $\psi_{a,b,\theta}(\mathbf{x})$ in 2-D are defined from a wavelet-typed function in 1-D $\psi(x)$ as:

$$\psi_{a,b,\theta}(\mathbf{x}) = a^{-1/2} \psi((x_1 \cos \theta + x_2 \sin \theta - b)/a). \quad (2)$$

In comparison, the (separable) continuous wavelet transform (CWT) in \mathbb{R}^2 can be written as:

$$W_f^2(a_1, a_2, b_1, b_2) = \int_{\mathbb{R}^2} \psi_{a_1,a_2,b_1,b_2}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}, \quad (3)$$

where the wavelets in 2-D are tensor products,

$$\psi_{a_1,a_2,b_1,b_2}(\mathbf{x}) = \psi_{a_1,b_1}(x_1) \psi_{a_2,b_2}(x_2), \quad (4)$$

of 1-D wavelets, $\psi_{a,b}(t) = a^{-1/2} \psi((t-b)/a)$ ¹.

The CRT appears similar to the 2-D CWT except that the *point* parameters (b_1, b_2) are replaced by the *line* parameters (b, θ) . In brief, those 2-D transforms are related by:

$$\begin{array}{ll} \mathbf{Wavelets:} & \rightarrow \psi_{scale, point-position} \\ \mathbf{Ridgelets:} & \rightarrow \psi_{scale, line-position} \end{array}$$

The consequence of this is: as wavelet analysis is very effective at representing objects with isolated point singularities, ridgelet analysis can be very effective at representing objects with singularities along lines. In fact, one can loosely view ridgelets as a way of concatenating 1-D wavelets along lines. Hence the motivation for using ridgelets in image processing tasks is very

This work was supported by a Department of Communication Systems, EPFL PhD Fellowship and the Swiss National Science Foundation grant number 21-52439.97.

¹In practice, however one typically enforces the same dilation scale on both directions thus leading to three wavelets corresponding to horizontal, vertical and diagonal directions.

appealing as singularities are often joined together along edges or contours in images.

Since in 2-D, points and lines are related via the Radon transform, thus the wavelet and ridgelet transforms are linked via the Radon transform. More precisely, denote the Radon transform as

$$RA_f(\theta, t) = \int_{\mathbb{R}^2} f(\mathbf{x}) \delta(x_1 \cos \theta + x_2 \sin \theta - t) d\mathbf{x}, \quad (5)$$

then the ridgelet transform is precisely the application of a 1-D wavelet transform to the slices of the Radon transform,

$$RI_f(a, b, \theta) = \int_{\mathbb{R}} \psi_{a,b}(t) RA_f(\theta, t) dt. \quad (6)$$

It is interesting to note that if in (6) instead of taking 1-D wavelet transform, the application of 1-D Fourier transform on t would result in the 2-D Fourier transform, or:

$$F_f^2(\xi \cos \theta, \xi \sin \theta) = \int_{\mathbb{R}} e^{-j\xi t} RA_f(\theta, t) dt. \quad (7)$$

This is the famous *projection-slice theorem* and is used often in image reconstruction from projection methods [5].

3. FINITE RADON TRANSFORM

As suggested from the previous section, a discrete-time ridgelet transform can be obtained using a discrete-time Radon transform. Numerous discretization of the Radon transforms have been devised to approximate the continuous formula [5]. However to our knowledge, none of them were designed to be as invertible transforms for digital images. On the other hand, the *finite* Radon transform theory [6, 7, 8] in combinatorics provides an interesting alternative.

The finite Radon transform (FRAT) is defined in a finite geometry in the same way as the continuous Radon transform is defined in the Euclidean geometry. Denote $Z_p = \{0, 1, \dots, p-1\}$, where p is a prime number. The FRAT of a real function f on the finite lattice Z_p^2 is defined as:

$$r(k, l) = FRAT_f(k, l) = \frac{1}{\sqrt{p}} \sum_{(i,j) \in L_{k,l}} f(i, j). \quad (8)$$

Here $L_{k,l}$ denotes the collection of points that make up a line on the lattice Z_p^2 or more specifically,

$$\begin{aligned} L_{k,l} &= \{(i, j) : j = ki + l \pmod{p}, i \in Z_p\}, k \in Z_p, \\ L_{p,l} &= \{(l, j) : j \in Z_p\}. \end{aligned} \quad (9)$$

In the FRAT domain, the energy is best compacted if the mean is subtracted from the image $f(i, j)$ previous to taking the transform given in (8) and we assume that in the sequel. The factor $\frac{1}{\sqrt{p}}$ is introduced to normalize the l_2 -norm.

So as in the Euclidean geometry, a line $L_{k,l}$ on the affine plane Z_p^2 is uniquely represented by its slope or direction k ($k = p$ corresponds to infinite slope or vertical lines) and its intercept l . It can be seen that there are $p^2 + p$ lines defined in this way and every line contains p points. Moreover, any two distinct points on Z_p^2 are in just one line. Also, two lines of different slopes intersect at

exactly one point. For any given slope, there are p parallel lines that provide a complete cover of the plane Z_p^2 . That means,

$$\sum_{l=0}^{p-1} r(k, l) = \frac{1}{\sqrt{p}} \sum_{(i,j) \in Z_p^2} f(i, j) = 0 \quad \forall k, 0 \leq k \leq p. \quad (10)$$

Eq. (10) explicitly reveal the redundancy of the FRAT. In each direction, there is only $(p-1)$ independent FRAT coefficients. Those coefficients at $p+1$ directions together with the mean value make up totally of $(p+1)(p-1) + 1 = p^2$ independent coefficients in the finite Radon domain (as expected!).

In analogy with the continuous case, the *finite back-projection* (FBP) operator is defined as the sum of Radon coefficients of all the lines that go through a given point, i.e.

$$FBP_r(i, j) = \frac{1}{\sqrt{p}} \sum_{(k,l) \in P_{i,j}} r(k, l), \quad (i, j) \in Z_p^2, \quad (11)$$

where $P_{i,j}$ denotes the set of indexes for lines that go through the point (i, j) . From (9) we have,

$$P_{i,j} = \{(k, l) : l = j - ki \pmod{p}, k \in Z_p\} + \{(p, i)\}$$

From the properties of the finite geometry that every two points lie in exactly one line and there are $(p+1)$ lines that go through each point, substitute (8) into (11) leads to:

$$FBP_r(i, j) = \frac{1}{p} \left(\sum_{(i',j') \in Z_p^2} f(i', j') + pf(i, j) \right) = f(i, j).$$

So the back-projection operator defined in (11) indeed computes the inverse FRAT. It is worth to note that the transform matrices for the operators *FRAT* and *FBP* are transposed of each other.

For any set A , write δ_A for the characteristic function of A . Then we can write the basis functions for the FRAT as $p^{-1/2} \delta_{L_{k,l}}$, $0 \leq k \leq p$, $0 \leq l < p$. The previously mentioned properties of lines in finite geometry leads to:

$$\langle \delta_{L_{k,l}}, \delta_{L_{k',l'}} \rangle = \begin{cases} p & \text{if } k = k', l = l' \\ 0 & \text{if } k = k', l \neq l' \\ 1 & \text{if } k \neq k' \end{cases} \quad (12)$$

Hence the minimum angle between any basis functions of the FRAT is: $\min_{k,l,k',l'} \angle(\delta_{L_{k,l}}, \delta_{L_{k',l'}}) = \cos^{-1}(1/p)$, which approaches the right angle for large p . So we can say that the finite Radon basis is “almost” orthogonal.

4. ORTHONORMAL FINITE RIDGELET TRANSFORM

Now with an invertible FRAT, applying (6) we can obtain an invertible discrete ridgelet transform by taking the discrete wavelet transform (DWT) on each vector, also called a *projection*, $\{r(k, l), l \in Z_p\}$ of Radon coefficients where the direction k is fixed². The result can be called as finite ridgelet transform (FRIT). Due to periodicity of the FRAT coefficients for each direction, the periodic wavelet transforms are chosen and assumed in the sequel.

²If p is not dyadic, a special border handling is required.

Note that the FRAT is redundant and not orthogonal. Next we will show that by taking the 1-D DWT on the projections of the FRAT in a special way, we can remove this redundancy and have an orthonormal FRIT.

Assume that the DWT is implemented by an orthogonal tree-structured filterbank with J states where G_0 and G_1 are low and high pass synthesis filters, respectively. Then the family of functions $\{g_1^{(j)}(n - 2^j m), g_0^{(j)}(n - 2^j m)\}$, $j = 1, \dots, J$, and $n, m \in \mathbb{Z}$ is the orthogonal basis of the discrete-time wavelet series [1]. Here $G^{(j)}$ are the equivalent filters at level j . The basis functions from $g_0^{(j)}$ are called the scaling functions where the others are called wavelet functions. Normally the filter G_1 is designed to satisfy the high pass condition, $G_1(z)|_{z=1} = 0$, so each wavelet basis function has zero sum.

For a general setting, let's assume that we have a collection of 1-D orthonormal transforms (which can be the same) for each projection k of FRAT, that have bases as:

$$\{w_m^{(k)}(\cdot), m \in Z_p\}, \quad 0 \leq k \leq p. \quad (13)$$

The only condition we require for each of these bases that it has a mean function, i.e. $w_0^{(k)}(l) = 1/\sqrt{p}$, $\forall l \in Z_p$. By the orthogonality condition, this means all other basis functions must have zero sum. As shown before, this requirement is satisfied for the wavelet bases where the DWT is carried to the maximum number of stages (i.e. when only one scaling coefficient is left).

The FRIT is can now be written as:

$$\begin{aligned} FRIT_f(k, m) &= \langle FRAT_f(k, \cdot), w_m^{(k)}(\cdot) \rangle \\ &= \sum_{l \in Z_p} w_m^{(k)}(l) \langle f, p^{-1/2} \delta_{L_k(l)} \rangle. \end{aligned} \quad (14)$$

Hence we can write the basis functions for the FRIT as:

$$\rho_{k,m} = \frac{1}{\sqrt{p}} \sum_{l \in Z_p} w_m^{(k)}(l) \delta_{L_k(l)}. \quad (15)$$

Consider the inner products between any two FRIT basis functions:

$$\langle \rho_{k,m}, \rho_{k',m'} \rangle = \frac{1}{p} \sum_{l, l' \in Z_p} w_m^{(k)}(l) w_{m'}^{(k')}(l') \langle \delta_{L_k(l)}, \delta_{L_{k'}(l')} \rangle.$$

Using (12), when the two FRIT basis functions have the same direction, $k = k'$, then

$$\langle \rho_{k,m}, \rho_{k,m'} \rangle = \sum_{l \in Z_p} w_m^{(k)}(l) w_{m'}^{(k)}(l) = \delta[m - m'].$$

So the orthogonality of these FRIT basis functions comes from the orthogonality of the basis $\{w_m^{(k)}(\cdot), m \in Z_p\}$. Next for the case when the two FRIT basis have different directions, $k \neq k'$, again using (12) leads to

$$\begin{aligned} \langle \rho_{k,m}, \rho_{k',m'} \rangle &= \frac{1}{p} \sum_{l, l' \in Z_p} w_m^{(k)}(l) w_{m'}^{(k')}(l') \\ &= \frac{1}{p} \left(\sum_{l \in Z_p} w_m^{(k)}(l) \right) \left(\sum_{l' \in Z_p} w_{m'}^{(k')}(l') \right). \end{aligned}$$

In this case, if either m or m' is non-zero, e.g. $m \neq 0$ then by the zero sum property of the bases, $\sum_{l \in Z_p} w_m^{(k)}(l) = 0$, it leads to $\langle \rho_{k,m}, \rho_{k',m'} \rangle = 0$.

Finally note that $\bigcup_l L_k(l) = Z_p^2$, for all direction k (see (10)). Therefore all of the FRIT basis functions $\rho_{k,0}$ correspond to the mean of the input image so we only need to keep one of them (at any direction) and denote it as ρ_0 . So we just proved the following result.

Theorem 1 Given $p + 1$ orthonormal bases in $l^2(Z_p)$ (which can be the same): $\{w_m^{(k)}(\cdot), m \in Z_p\}$ with $w_0^{(k)} = const$, $0 \leq k \leq p$ then

$$\{\rho_{k,m} : k = 0, \dots, p; m = 1, \dots, p - 1\} \cup \{\rho_0 = const\}$$

is a orthonormal basis in $l^2(Z_p^2)$, where $\rho_{k,m}$ are defined in (15).

Remarks

1. We prove the above result for the general setting where different transforms can be applied on different FRAT projections. This could be useful for example when one uses wavelet packets or adaptive bases on each projection. Furthermore, due to the “wrap around” effect of the FRAT, its projections could have strong periodicity components so for some projection one could use a Fourier-type transform like DCT. Note [7] that if we apply the 1-D Fourier transform on all of the FRAT projections then we would obtain the 2-D Fourier transform.
2. All of these results also hold for any other affine planes rather than Z_p^2 . Especially using Galois planes constructed by finite field $GF(p^r)$ we can build orthogonal FRIT's for images of size $N \times N$ where N is a power of a prime, $N = p^r$.

5. NUMERICAL RESULTS

Figure 1 displays some basis images for the 7×7 FRIT using 1-D Haar wavelets. As can be seen from the figure, FRIT basis images have elongated linear structure which closely resemble the continuous ridge functions.

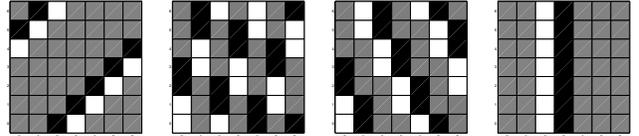


Fig. 1. Examples of 7×7 FRIT basis functions using Haar wavelets. Black, grey and white pixels corresponding respectively to negative, zero and positive values. Notice the “wrap-around” effect due to periodization of FRAT.

Following the study of the efficiency of the ridgelet transform in the continuous domain on the half-plane truncated Gaussian function, $f(x_1, x_2) = 1_{\{x_2 > 0\}} e^{-x_1^2 - x_2^2}$, $\mathbf{x} \in \mathbb{R}^2$ [4], we first perform a numerical comparison on the 127×127 discrete image of the same function using four competitive transforms: DCT, DWT, FRAT and FRIT. The comparison is measured by the non-linear approximation power, i.e. the ability of reconstructing the original image (measured in term of signal-to-noise ratios) using the N largest magnitude transform coefficients. Fig. 2 plots the results. Clearly that the FRIT achieves the best performance, as expected from the continuous theory.

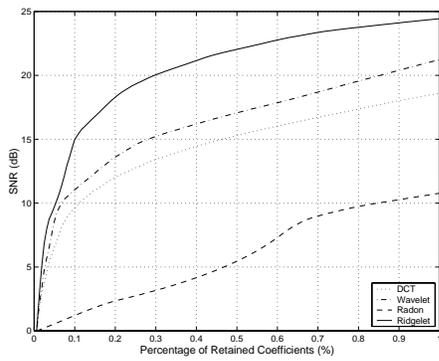


Fig. 2. Comparison of non-linear approximations on the 127×127 half-plane truncated Gaussian image.

Our next test is a real image of size 127×127 with linear discontinuities. Figure 3 displays the image together with its FRIT. In the FRIT image, each column corresponds to one direction.

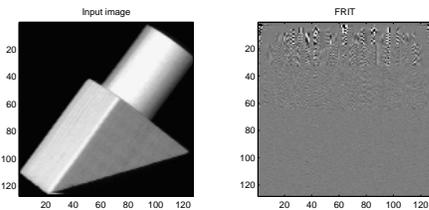


Fig. 3. Example of FRIT of an image.

To have an insight of the FRIT, Figure 4 plot the top five FRAT projections that contain most of the energy. Those correspond to the directions that have discontinuities, plus the horizontal ($k = 0$) and vertical ($k = 127$) directions. So we see that at first the FRAT compact most of the energy of the image into a few projections, where the linear discontinuities create “jumps”. Next by applying 1-D DWT on those projections that are mainly smooth, energy is further compacted into a few FRIT coefficients.

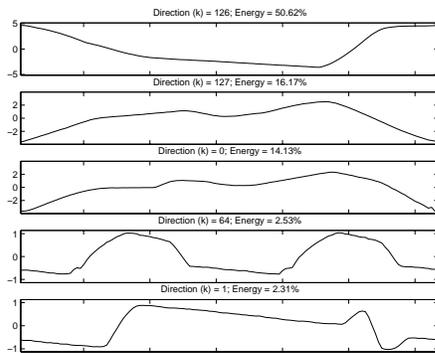


Fig. 4. Top five FRAT projections that contain most of the energy.

Finally, Figure 5 shows the images obtained from non-linear approximations using FRIT.

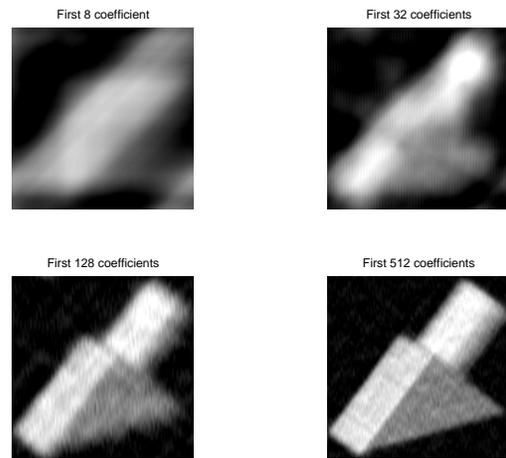


Fig. 5. Non-linear approximations using FRIT on the 127×127 test image in Figure 3.

6. DISCUSSION AND FURTHER WORK

The FRIT presented in this paper is our first step in achieving a new scheme which can deal efficiently with natural images that are typical piecewise smooth away from singularities along edges. Since ridgelets are specially adapted only to straight singularities, a more practical transform would first utilize a quad-tree division of images into localized pieces where edges looks straight and then apply FRIT to each piece. Furthermore, for bit-level compression applications, one need to code the position of the significant ridgelet coefficient in an efficient way, probably via some embedded tree-structured significant maps. These topics are under investigation; we plan to report the results in forthcoming paper.

Acknowledgment: The authors would like to thank Prof. Attila Kuba and Prof. Gabor Herman for pointing us to references [7, 8].

7. REFERENCES

- [1] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*, Prentice-Hall, Englewood Cliffs, NJ, 1995.
- [2] D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies, “Data compression and harmonic analysis,” *IEEE Trans. Info. Theory*, vol. 44, no. 6, pp. 2435–2476, October 1998.
- [3] E. Candès, *Ridgelets: theory and applications*, Ph.D. thesis, Department of Statistics, Stanford University, 1998.
- [4] E. Candès and D. L. Donoho, “Ridgelets: a key to higher-dimensional intermittency?,” *Phil. Trans. R. Soc. Lond. A.*, pp. 2495–2509, 1999.
- [5] G. T. Herman, *Image Reconstruction from Projections: The Fundamentals of Computerized Tomography*, Academic Press, 1980.
- [6] E. D. Bolker, “The finite Radon transform,” in *Integral Geometry (Contemporary Mathematics, Vol. 63)*, S. Helgason R. L. Bryant, V. Guillemin and R. O. Wells Jr., Eds., pp. 27–50. 1987.
- [7] F. Matúš and J. Flusser, “Image representation via a finite Radon transform,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, no. 10, pp. 996–1006, Oct 1993.
- [8] P. M. Salzberg and R. Figueroa, “Tomography on the 3D-torus and crystals,” in *Discrete Tomography: Foundations, Algorithms and Applications*, G. T. Herman and A. Kuba, Eds., pp. 417–434. Birkhäuser, 1999.