RIDGELET TRANSFORM APPLIED TO MOTION COMPENSATED IMAGES

Lorenzo Granai[1], Fulvio Moschetti[2], Pierre Vandergheynst[1]

[1] ITS Signal Processing Institute, EPFL Swiss Federal Institute of Technology, Lausanne, Switzerland
[2]NTT DoCoMo, Inc. Multimedia Signal Processing Laboratory, 3-5 Hikarinooka, Yokosuka, Kanagawa, Japan

ABSTRACT

Wavelet transform is a powerful instrument in catching zero-dimensional singularities. Ridgelets are a powerful instrument in catching and representing mono-dimensional singularities in bidimensional space.

In this paper we propose a hybrid video coder scheme using ridgelet transform for the first approximation of line-edge singularities in displaced frame difference images. We demonstrate the potential of ridgelets and results show substantial improvements when compared to wavelet only based coder.

1. INTRODUCTION

Images are generally described via orthogonal, non-redundant transforms like wavelet or discrete cosine transform. The good performance of wavelets in mono-dimensional domain is lost when they are applied to images using 2D separable basis since they are not able to efficiently code mono-dimensional singularities.

The ridgelet transform achieves very compact representation of linear singularities in images [1,2,3]. Instrumental in the implementation of the ridgelet is the Radon transform, which is a powerful tool to extract lines in edge dominated images. Therefore, they can offer an important contribution in order to detect and represent edges, which are fundamental structures in natural images and particularly relevant from a visual point of view.

Concerning video coding, the motion compensation procedure, commonly employed in traditional hybrid coding schemes, produces a displaced frame difference (dfd) that appears as an edge dominated image. Intuition suggests that ridgelet transform can be a good tool for the first approximations of dfd images.

2. RIDGELET THEORY

2.1. Continuous Ridgelet Transform

We start by briefly reviewing the continuous ridgelet transform, defined by Candès and Donoho in [1], stemming from the Radon transform, instrumental in its implementation. Given an integrable bivariate function $f(x_1,x_2)$, its Radon transform (RDN) is defined by:

$$RDN_f(\theta,t) = \int_{R^2} f(x_1,x_2)\delta(x_1\cos\theta + x_2\sin\theta - t)dx_1dx_2 \,. \quad (1)$$

Basically the Radon operator maps the spatial domain into the projection domain (θ, t), in which each point corresponds to a straight line in the spatial domain; conversely, each point in the spatial domain becomes a sine curve in the projection domain.

The Continuous Ridgelet Transform (CRT) is simply the application of a mono-dimensional wavelet ($\psi_{a,b}(t) = a^{-1/2}\psi((t-b)/a)$) to the slices of the Radon transform:

$$CRT(a,b,\theta) = \int_R \psi_{a,b}(t)RDN_f(\theta,t)dt =$$

$$= \int_{R^2} \psi_{a,b,\theta}(x_1,x_2)f(x_1,x_2)dx_1dx_2 \,, \quad (2)$$

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

$$\psi_{a,b,\theta}(x_1,x_2) = a^{-1/2}\psi((x_1\cos\theta + x_2\sin\theta - b)/a) \,. \quad (3)$$

This shows that the ridgelet function is constant along the lines where $x_1\cos\theta + x_2\sin\theta = const.$

Comparing ridgelets with wavelets we observe that the parameters of the former are scale factor and line position (respectively a and (b,θ) in (2)), while the latter uses scale factor and point position. As a consequence, wavelets are very effective in representing objects with isolated point singularities, while ridgelets are very effective in representing objects with singularities along straight lines.

It is interesting to notice that in the Radon domain if we apply a 1D Fourier transform along t instead of a wavelet we will obtain the 2D Fourier transform of f. This result is known as projection-slice theorem [4].

2.2. Finite Ridgelet Transform

In order to apply ridgelets to digital images a discrete transform is needed, and this leads to the research of a discrete Radon transform. In [5,6], Do and Vetterli propose a new procedure that results to be invertible, orthogonal and achieves perfect reconstruction: the *Finite*

RIdgelet Transform (FRIT). FRIT is based on the *Finite RAdon Transform* (FRAT) [7], which is defined as summations of image pixels over a certain set of "lines". Those lines are defined in a finite geometry in a similar way as the lines for the continuous Radon transform in the Euclidean geometry. Denote $Z_p = \{0,1,.., p-1\}$, where p is a prime number. Note that Z_p is a finite field with modulo p operations. The FRAT of a real discrete function f on the finite grid Z_p^2 is defined as:

$$FRAT_f(k,l) = \frac{1}{\sqrt{p}} \sum_{(i,j)\in L_{k,l}} f(i,j). \qquad (4)$$

Here $L_{k,l}$ denotes the set of points that make up a line on the lattice Z_p^2, i.e.

$$L_{k,l} = \begin{cases} \{(i,j) : j = (ki+l)(\mathrm{mod}p), i \in Z_p\} \; if \; 0 \le k < p \\ \{(l,j) : j \in Z_p\} \; if \; k = p \end{cases} .(5)$$

The inverse transform is obtained through the *Finite Back-Projection operator* (FBP) defined as a sum of Radon coefficients of all the lines that go through a given point. Here f is supposed to be a zero-mean image:

$$FBP_r(i,j) = \frac{1}{\sqrt{p}} \sum_{(k,l)\in P_{i,j}} FRAT_f(k,l). \qquad (6)$$

From (5) it can be found that $P_{i,j}$ is:

$$P_{i,j} = \{(k,l) : l = (j-ki)(\mathrm{mod}\ p), k \in Z_p\} \cup \{(p,i)\}.$$

Substituting (6) into (4) we obtain that

$$FBP_r(i,j) = \frac{1}{p} \sum_{(k,l)\in P_{i,j}} \sum_{(i',j')\in L_{k,l}} f(i',j') =$$

$$= \frac{1}{p}\left(\sum_{(i',j')\in Z_p^2} f(i',j') + p \cdot f(i,j) \right) = f(i,j) \qquad (7)$$

and so the perfect reconstruction is achieved.

It is easy to compute that the number of operations required by FRAT is p^3 addictions and p^2 multiplications, so almost comparable with other transforms like 2D-FFT. A drawback of this discrete implementation is the "wrap-around" effect, already observed by Do.

3. FRIT ON TEST IMAGES

FRIT needs an input image of size *pxp,* where p is a prime number, and this is an important limitation of this algorithm. Moreover wavelets usually require a dyadic length signal and this is absolutely incompatible with the FRAT output (that is a matrix px(p+1)). In the test we made, we extended the length of the signal from p to m, where m is defined as:

$$m = \min\{n \in N : (n > p) \text{ and } (n = 2^d), d \in N\}.$$

Typically we took p=31 or p=127 and so m=32 or 128 respectively.

Fig.1: Artificial image: original (left), reconstructed with 20 wavelet coefficients (center) and with 20 ridgelet coefficients (right).

Figure 1 represents an image 31x31 reconstructed with wavelets (Daubechies 9,7) and ridgelets. In this case, favorable to the second transform, the FRIT is able to detect the line structures even using a very small number of coefficients.

4. FRIT APPLIED TO MOTION COMPENSATED IMAGES

In natural video coding, displaced frame differences are characterized by a different range of luminance values when compared to natural images and by higher frequency components. Usually, as already observed, they present many edges.

We aim to catch these features exploiting the directional information given by the Radon transform and consequently by the FRIT.

Applying the ridgelet transform to a dfd and comparing with wavelets, we notice an interesting result: when we consider the very first coefficients the error obtained with the FRIT is lower (see figure 2). Afterwards wavelets have a better behavior and this is due to the fact that Ridgelets represent efficiently straight lines but not curves. Moreover a noise-like effect appears in the FRIT reconstructed images as can be seen in figures 1 and 3b. This is related to the wrap-around effect that occurs when choosing the lattice as in (5). In [5] Do and Vetterli propose a new ordering for the FRAT coefficients; we tested the old and the new one on motion compensated images observing that results do not significantly differ. In fact this new optimal ordering is designed for natural images and not for dfd.

Now studying the graph in figure 2 we notice that it can be split into two parts. This suggests that there is a potential for a base layer.

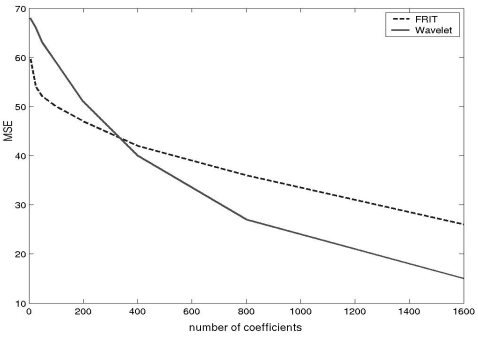

Fig.2: Mse decay using an increasing number of coefficients to reconstruct the image. Comparison between wavelets and ridgelets.

These results still hold for smaller blocks as can be seen in Table 1 where a block 31x31 of the same sequence is studied.

Coefficients	*Frit*	*Wavelet*
5	304	334
10	270	291
50	191	130
100	139	81

Tab.1: Mse for a block 31x31 taken from the sequence "Stefan".

5. HYBRID CODING

From what observed in the previous section a possibility to exploit the advantages of FRIT is to apply it to a block of a dfd and check if it is able to detect lines with the first coefficients. We propose a method composed of two stages. First the FRIT is used to catch line patterns with its biggest coefficients, afterwards the reconstructed image is subtracted from the original one and the residual is passed to the second stage. Here a classical wavelet decomposition is performed (2D-DWT). Figure 4 illustrates the scheme of this hybrid algorithm. The input dfd image has been obtained using a motion estimation algorithm similar to MPEG-4. The inverse algorithm is very simple: wavelet and FRIT coefficients are decoded separately and then the two images are added in order to obtain the reconstructed dfd.

This technique, that exploits the potential base layer shown in figure 2, offers a double advantage: first, lines are represented in the optimal way (i.e. using ridgelets), second the residual is easier to code for wavelets since some edges were eliminated in the previous stage.

This hybrid coding method does not require to split the image in small blocks and it can be applied to the whole image too. Only the features that are really efficiently represented by ridgelets are considered, all the rest is transferred to the second step.

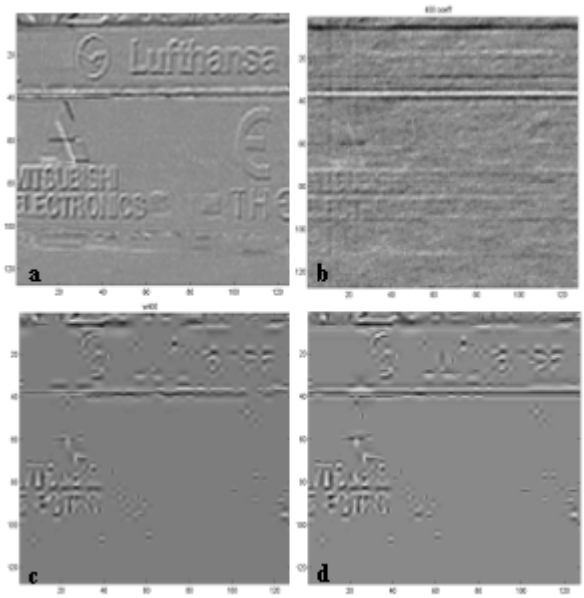

Fig.3: Motion compensated block extracted from the sequence "Stefan". Original (a) and reconstructed with FRIT(b), wavelets (c) and hybrid method (d). 400 coefficients.

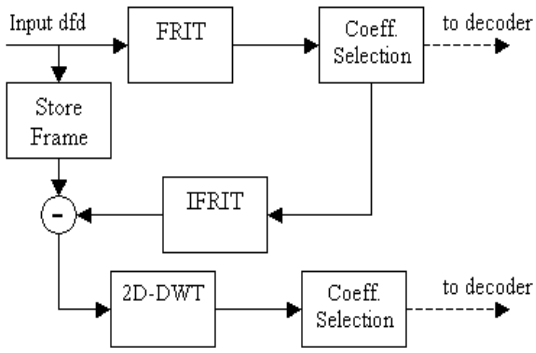

Fig.4: Scheme of the hybrid algorithm.

6. RESULTS

We consider now the case of a block of 127x127 pixels, extracted from the sequence "Stefan". Figure 3 shows the original block and the ones reconstructed with FRIT, 2D-DWT and the hybrid method. It should be noted that the pixels values are shifted and scaled, being the original range of the frame [-255, 255].

Table 2 illustrates the mse obtained with different number of coefficients: at high compression rates the advantage is evident. The number of coefficients for the hybrid encoder is the sum of the FRIT (fixed to five) and the wavelet ones.

coefficients	*wavelet*	*FRIT+ wavelet*
100	58.7	50.4
200	51	44
400	40.1	34.9

Tab.2: Mse obtained applying wavelets and the hybrid method on a block 127x127 from the sequence "Stefan".

Another interesting comparison can be done computing how many coefficients are necessary to reach a certain mse. Tables 3 and 4 examine this situation for two different frames taken from sport sequences The last column gives the percentage of coefficients that the new method saves compared with wavelets.

In all the examples reported the 2D wavelet transform is performed using biorthogonal Daubechies 9,7 functions.

Target MSE	*FRIT+wavelet*	*wavelet*	*var (%)*
10	100	110	10%
8	200	225	12.5%
6	400	450	12.5%

Tab.3: Number of coefficients needed to reach a given mse. Data refers to a block 127x127 taken from the soccer sequence "J-league" in format QCIF.

Target MSE	*FRIT+wavelet*	*wavelet*	*var (%)*
50	100	210	110%
44	200	315	57%
35	400	530	32%

Tab.4: Number of coefficients needed to reach a given mse. Data refers to a block 127x127 taken from the sequence "Stefan" in format CIF.

7. CONCLUSIONS

Ridgelet transform turns out to be optimal for representing discontinuities along straight lines [2]. In order to employ it with complex images a more elaborated structure is needed. For example one can first utilize a quad-tree division [8,9] of an image in small blocks and then apply ridgelets or use the curvelet transform [10] based on a localized application of ridgelets.

Here we propose a different approach, suitable for motion compensated images, that exploit ridgelets ability to find and represent edges, employing them for a base-layer coding. This method is based on the idea that an image is built of several components: one technique can be adopted to represent straight lines while the rest can be represented employing other transforms.

Moreover, from a visual point of view, the fact that lines are well reconstructed (even though the image is highly compressed) turns out to be very important.

There are video sequences that have many lines that can be well treated with ridgelets (a lot of sport sequences for example) and others in which this technique is useless. It could be interesting to develop an adaptive version of the proposed algorithm that finds out how many ridgelets coefficients are necessary to efficiently code edges on a certain dfd before passing the residual to the second step.

In the hybrid scheme that we presented the final stage is entrusted to wavelets but other techniques could also be used. It should be noticed that if the coder utilized in the second stage is block-based the capacity of detecting lines on the whole frame is especially useful.

Finally, it could be particularly interesting to use a matching pursuit algorithm [11,12] instead of wavelets as it seems to be better compatible with this hybrid coding implementation.

8. REFERENCES

[1] E.J. Candès, and D.L. Donoho, "Ridgelets: a key to higher dimensional intermittency?," Phil. Trans. R. Soc. Lond. A. , pp.2495-2509, 1999.

[2] E.J. Candès "Ridgelets and Their Derivatives: Representation of Images with Edges," Curves and Surfaces, L.L. Schumaker et al. (eds), Vanderbilt University Press, Nashville, TN, 1999.

[3] E.J. Candès, "Ridgelets and the Representation of Mutilated Sobolev Functions". SIAM J. Math. Anal., 1999.

[4] A. Rosenfeld and A.C. Kak , *Digital Picture Processing*, Academic Press, 2nd edition, 1982.

[5] M.N. Do and M. Vetterli, "Finite Ridgelet Transform for Image Representation," IEEE Transactions on Image Processing, 2002.

[6] M.N. Do and M. Vetterli, "Orthonormal finite ridgelet transform for image compression," in Proc. IEEE Int. Conf. on Image Processing (ICIP), 2000.

[7] F. Matus and J. Flusser, "Image representation via a finite Radon transform," IEEE Trans. Pattern Anal. Machine Intell., vol. 15, no. 10, pp. 996–1006, 1993.

[8] M.N. Do, P.L. Dragotti, R. Shukla, and M. Vetterli, "On the compression of two dimensional piecewise smooth functions," in Proc. of IEEE Int. Conf. on Image Processing (ICIP), 2001.

[9] R.M. Figueras i Ventura, L. Granai and P. Vandergheynst, "R-D Analysis of Adaptive Edge Representations," IEEE International Workshop on Multimedia Signal Processing MMSP, December 2002.

[10] E.J. Candès and D.L. Donoho, "Curvelets - A Surprisingly Effective Nonadaptive Representation for Objects With Edges," Curves and Surfaces, L. L. Schumaker et Al. (eds), Vanderbilt University Press, Nashville, TN, 1999.

[11] O.K. Al-Shaykh, et al., "Video Compression Using Matching Pursuit," IEEE Trans. Circuits Syst. Video Technol., vol. 9 (n.1): p. 123-143, 1999.

[12] F. Moschetti, L. Granai, P. Vandergheynst and P. Frossard, "New Dictionary and Fast Atom Searching Methods for Matching Pursuit Representation of Displaced Frame Difference," in Proc. of IEEE ICIP, 2002.