

Available online at www.sciencedirect.com





C. R. Acad. Sci. Paris, Ser. I 336 (2003) 395-398

Mathematical Analysis Curvelets and Fourier Integral Operators Curvelets et Opérateurs Intégraux de Fourier Emmanuel Candès, Laurent Demanet

Applied and Computational Mathematics, California Institute of Technology, Mail Code 217-50, Pasadena, CA 91125, USA

Received 5 December 2002; accepted after revision 9 January 2003

Presented by Yves Meyer Dedicated to Yves Meyer

Abstract

A recent body of work introduced new tight-frames of *curvelets* E. Candès, D. Donoho, in: (i) Curvelets – a suprisingly effective nonadaptive representation for objects with edges (A. Cohen, C. Rabut, L. Schumaker (Eds.)), Vanderbilt University Press, Nashville, 2000, pp. 105–120; (ii) http://www.acm.caltech.edu/~emmanuel/publications.html, 2002 to address key problems in approximation theory and image processing. This paper shows that curvelets essentially provide optimally sparse representations of Fourier Integral Operators. *To cite this article: E. Candès, L. Demanet, C. R. Acad. Sci. Paris, Ser. I 336* (2003).

© 2003 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS. All rights reserved.

Résumé

Une série de récents articles ont introduit l'analyse en *curvelets* E. Candès, D. Donoho, in : (i) Curvelets – a surprisingly effective nonadaptive representation for objects with edges (A. Cohen, C. Rabut, L. Schumaker (Eds.)), Vanderbilt University Press, Nashville, 2000, pp. 105–120; (ii) http://www.acm.caltech.edu/~emmanuel/publications.html, 2002 : les curvelets offrent une représentation multi-échelle qui ouvre de nouvelles perspectives pour l'analyse de problèmes importants en théorie de l'approximation et en traitement de l'image. Cet article montre que les curvelets permettent une représentation optimale de la classe des opérateurs intégraux de Fourier. Par «optimale », nous entendons par exemple, la plus économe. *Pour citer cet article : E. Candès, L. Demanet, C. R. Acad. Sci. Paris, Ser. I 336 (2003).*

© 2003 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS. Tous droits réservés.

1. Introduction

This paper is concerned with the representation of a large class of operators, namely, Fourier Integral Operators (FIOs) in the newly introduced tight frames of *curvelets* [3,4]. Curvelets are a new multiscale construction for representing bivariate functions and were originally introduced in connection with central problems in approximation theory and statistical estimation; since then, curvelets have also made their way in image processing as an alternative to other classical image representations. Recall that a collection of functions $(f_{\mu})_{\mu}$ is said to be a tight frame for $L_2(\mathbb{R}^2)$ if it obeys the Parseval relation

E-mail addresses: emmanuel@acm.caltech.edu (E. Candès), demanet@acm.caltech.edu (L. Demanet).

¹⁶³¹⁻⁰⁷³X/03/\$ – see front matter © 2003 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS. Tous droits réservés. doi:10.1016/S1631-073X(03)00095-5

E. Candès, L. Demanet / C. R. Acad. Sci. Paris, Ser. I 336 (2003) 395-398

$$\sum_{\mu} \left| \langle f, f_{\mu} \rangle \right|^2 = \left\| f \right\|_{L_2(\mathbb{R}^2)}^2, \quad \forall f \in L_2(\mathbb{R}^2).$$

$$\tag{1}$$

This relation implies, by standard arguments, that we have available the reconstruction formula $f = \sum_{\mu} \langle f, f_{\mu} \rangle f_{\mu}$, with equality holding in an L_2 sense.

To introduce the concept of representation of a linear operator, suppose we are given a linear transformation Tacting on square integrable functions $f \in L_2(\mathbb{R}^2)$. We may want to think of T via its action on the frame elements f_{μ} and introduce the infinite matrix

$$T_F(\mu,\mu') = \langle f_{\mu}, Tf_{\mu'} \rangle; \tag{2}$$

in other words, T_F maps the coefficients of an object f into those of T f. Needless to say, the datum of the matrix $T_F(\mu, \mu')$ completely specifies the operator T as an operator from $L_2(\mathbb{R}^2)$ to itself since $(f_{\mu})_{\mu}$ is a tight frame for $L_2(\mathbb{R}^2).$

An operator T is said to be a Fourier Integral Operator (FIO) if it is of the form

$$Tf(x) = \int e^{i\Phi(x,\xi)} a(x,\xi) \hat{f}(\xi) \,\mathrm{d}\xi.$$
(3)

Here Φ is a phase function and a is an amplitude which we suppose obey the following standard assumptions: the phase $\Phi(x,\xi)$ is C^{∞} , homogeneous of degree 1 in ξ , i.e., $\Phi(x,\lambda\xi) = \lambda \Phi(x,\xi)$ for $\lambda > 0$, and with $\Phi_{x\xi} = \nabla_x \nabla_\xi \Phi$, obeys the nondegeneracy condition $|\det \Phi_{x\xi}(x,\xi)| > c > 0$, uniformly in x and ξ ; the amplitude a is a symbol of order *m*, which means that *a* is C^{∞} , and obeys $|\partial_{\xi}^{\alpha}\partial_{x}^{\beta}a(x,\xi)| \leq C_{\alpha\beta}(1+|\xi|)^{m-\alpha}$.

Wavelets are known to provide sparse representations of pseudo-differential operators; that is when the phase function Φ is linear in both variables x and ξ , $\Phi(x,\xi) = x \cdot \xi$. It is also known that both Fourier and wavelet bases do not provide sparse representations of FIOs.

2. Curvelets

By now, there are several types of curvelet frames [3,4] and we now briefly discuss the curvelet frame as introduced in [5,4]. We let μ be the triple (j, ℓ, k) : here, $j = 0, 1, 2, \ldots$ is a scale parameter; $\ell = 0, 1, \ldots, 2^{\lfloor j/2 \rfloor} - 1$ is an orientation parameter ($\lfloor x \rfloor$ is the integer part of x); and $k = (k_1, k_2), k_1, k_2 \in \mathbb{Z}$, is a translation parameter. Introduce

- (1) the *parabolic scaling* matrix $D_j = \text{diag}(2^j, 2^{\lfloor j/2 \rfloor})$ which is diagonal and whose entries are 2^j and $2^{\lfloor j/2 \rfloor}$; (2) the *rotation angle* $\theta_J = 2\pi \cdot 2^{-\lfloor j/2 \rfloor} \cdot \ell$, with *J* indexing the scale/angle pair $J = (j, \ell)$;
- (3) and the *translation parameter* $k_{\delta} = (k_1 \cdot \delta_1, k_2 \cdot \delta_2)$ (δ_1, δ_2 are some universal numerical quantities, e.g., $\delta_1 = 14/3$ and $\delta_2 = 10\pi/9$, see [4] for details).

With these notations, we define curvelets as functions of $x \in \mathbb{R}^2$ by $\gamma_{\mu}(x) = 2^{3j/4} \gamma_{(j)}(D_j R_{\theta_j} x - k_{\delta})$. The profile $\gamma_{(i)}$ actually depends on the scale parameter j but in a non essential way; for each j, $\gamma_{(i)}$ is smooth and oscillatory in the horizontal direction and bell-shaped (nonoscillatory) along the vertical direction and is well-localized in space; e.g., for each $m = 0, 1, ..., \gamma_{(j)}(x)$ and its derivatives obey $|\gamma_{(j)}(x)| \leq C_m (1+|x|)^{-m}$, uniformly in j.

The frequency-side description of a curvelet is equally important to understand our main results. In the frequency domain, curvelets are compactly supported and each element $\hat{\gamma}_{\mu}$ is localized near the symmetric wedge

$$W_J = \left\{ \pm \xi, \ 2^j \leqslant |\xi| \leqslant 2^{j+1}, \ |\theta - \theta_J| \leqslant \pi \cdot 2^{-\lfloor j/2 \rfloor} \right\},\tag{4}$$

i.e., curvelets are supported inside symmetric wedges of length about 2^{j} and width about $2^{j/2}$. This explains their oscillatory nature: at scale 2^{-j} , a curvelet is a little needle whose envelope is a specified 'ridge' of effective length $2^{-j/2}$ and width 2^{-j} , and which displays an oscillatory behavior across the main 'ridge'. Note that this frequency localization idea is known in the literature as the the Second Dyadic Decomposition, see [6].

396

As in wavelet theory, we also have coarse scale elements which are of the form $\varphi_{k_1,k_2}(x) = \varphi(x - k_{\delta}), k_1, k_2 \in \mathbb{Z}$, i.e., translates of a waveform $\varphi(x_1, x_2)$ that we shall take to be bandlimited and rapidly decaying. Augmented with this layer of coarse scale elements, the system $(\gamma_{\mu})_{\mu}$ obeys the Parseval relation (1) and the corresponding reproducing formula.

3. Main result

A distinguished feature of the curvelet transform is that the action of an FIO on curvelet elements is in some sense very 'simple'. Roughly speaking, a curvelet γ_{μ} is mapped into another curvelet at a corresponding index $t(\mu)$.

There are many ways to establish a formal index correspondence and we only present a possible approach. Let γ_{μ} be a curvelet with scale $2^{-j_{\mu}}$, location x_{μ} and codirection ξ_{μ} (the unit vector in the frequency plane along which $\hat{\gamma}_{\mu}$ is supported); and *T* be an FIO with phase Φ . Now let the subscripts *x* and ξ denote partial differentiation and define

$$\phi_{\mu}(x) = \Phi_{\xi}(x,\xi_{\mu}), \quad y_{\mu} = \phi_{\mu}^{-1}(x_{\mu}), \quad \text{and} \quad A_{\mu} = \Phi_{x\xi}(y_{\mu},\xi_{\mu}),$$
(5)

and put $\tau_{\mu} = A_{\mu}^{T} \xi_{\mu} / ||A_{\mu}^{T} \xi_{\mu}||$. With these notations, we introduce the index mapping *t* defined as follows: $\mu' = t(\mu)$ with (1) $j_{\mu'} = j_{\mu}$, (2) $\xi_{\mu'}$ is the closest point to τ_{μ} on our discrete lattice, and (3) $x_{\mu'}$ is the closest point to y_{μ} on the Cartesian lattice induced by the pair $(j_{\mu'}, \theta_{\mu'})$.

Finally, introduce the notional distance ω between pairs of indices (μ, μ') :

$$\left|\omega(\mu,\mu')\right| = 2^{|j_{\mu}-j_{\mu'}|} \cdot \left(1 + \min\left(2^{j_{\mu}}, 2^{j_{\mu'}}\right) \left[|\theta_{\mu} - \theta_{\mu'}|^2 + |x_{\mu} - x_{\mu'}|^2 + \left|\langle \xi_{\mu}, x_{\mu} - x_{\mu'} \rangle\right|\right]\right).$$
(6)

We see that ω increases as the distance between the scale, angular, and location parameters increases. Note that the extra term $|\langle \xi_{\mu}, x_{\mu} - x_{\mu'} \rangle|$ induces a non-Euclidean notion of distance between x_{μ} and $x_{\mu'}$. Equipped with this definition, we may now state the main results of this paper.

Theorem 3.1. Let T be a FIO with a symbol of order m = 0 so that T is a bounded L_2 -operator which obeys the above assumptions. Then for each $N \ge 0$, the matrix entries in a curvelet frame obey

$$\left|T_{F}(\mu,\mu')\right| \leqslant C_{N} \cdot \left|\omega\left(\mu,t(\mu')\right)\right|^{-N},\tag{7}$$

for some constant $C_N > 0$.

A corollary of this result is as follows:

Theorem 3.2. Under the same assumptions of Theorem 3.1, the matrix T maps boundedly ℓ_p into ℓ_p for every $0 . For <math>p \leq 1$, this says that

$$\|T_F\|_{\ell_p \to \ell_p}^p = \sup_{\mu'} \sum_{\mu} |T_F(\mu, \mu')|^p \leqslant C_p^p.$$
(8)

We give an equivalent formulation of (8). Let (a_{μ}) be either a row or a column of T_F , and $|a|_{(n)}$ be the *n*-largest entry of the sequence $|a_{\mu}|$. Then for each r > 0, $|a|_{(n)}$ obeys $|a|_{(n)} \leq C_r \cdot n^{-r}$, where the constant C_r does not depend on the row or column index. In short, the row or column entries of the matrix T_F decay nearly exponentially, i.e., faster than any negative polynomial.

The above two theorems say that the curvelet matrix is both *sparse* and *well-organized*. Roughly speaking, if we think about the wavelet matrix of a pseudo-differential operator as being *almost* diagonal, then we may think of the curvelet matrix of an FIO as being *almost* a permutation. Mathematically speaking, consider the approximate or 'compressed' operator T^B with at most B elements per row and column – those indices which are the closest to $t(\mu)$ in the sense of (6) – and setting the others to zero. Then it is not hard to conclude that T^B would obey $||T - T^B||_{L_2 \to L_2} \leq C_m \cdot B^{-m}$ for each m > 0. Note that our results do not imply a characterization of the operators for which the decay estimates (7), (8) would hold.

4. Importance of the parabolic scaling

Because of space constraints, we cannot possibly give a proof of Theorems 3.1 and 3.2 and refer the reader to the companion paper [2]. However, there is a very interesting phenomenon which occurs here and we now highlight. Instead of curvelets, we may want to consider general scaling matrices of the form $D_j = \text{diag}(2^j, 2^{j\alpha}), 0 \le \alpha \le 1$. We would then obtain tight frames whose elements would be needles with length about $2^{-j\alpha}$ and width 2^{-j} . We might then consider representing a FIO with basis elements exhibiting such arbitrary scaling ratios. The point here is that no scaling other than $\alpha = 1/2$ yields sparse representations of FIOs. To understand this fact, consider that the action of an FIO on a curvelet γ_{μ} is that of a pseudo-differential operator followed by a smooth change of coordinates (which are both index-dependent); and curvelets provide sparse decompositions of pseudo-differential operators and smooth change of coordinates *simultaneously*. Such sparse decompositions hold, if and only if, the scaling is indeed parabolic.

5. Discussion

The potential for sparsity is of course wide-ranging. Let us only mention that it may allow the design of fast matrix multiplication and fast matrix inversion algorithms. For instance, [1] had a large impact by showing how to use the wavelet transform to compute certain types of singular integrals in a number of operations of the order of $C(\varepsilon) \cdot N \log N$ where $C(\varepsilon)$ is a constant depending upon the desired accuracy ε . Fourier Integral Operators are also tightly connected to linear hyperbolic partial differential equations. In some sense the solution operators for such equations are almost FIO's and turn out to be equally sparse in a curvelet frame. Work in progress attempts to exploit this feature to develop fast multiscale solvers (based on fast digital curvelet transforms) for classes of wave propagation problems. We hope to report on this in a future publication.

While working on this project, we became aware of the work of Smith [7,8] which addresses topics such as the description of Hardy spaces for FIOs and the construction of parametrices for nonsmooth second-order linear wave equations. Especially, [8] alludes to estimates similiar to those developed in Theorem 3.1 although we have not been able to find proofs of such results. We find the connection with this line of research in pure harmonic analysis nevertheless stimulating. Our agenda is of course very different here and corresponds to the viewpoint of Computational Harmonic Analysis: namely, we are interested in a remarkable mathematical statement which says that curvelets provide optimally sparse representations of FIOs.

Acknowledgements

This research was supported by a National Science Foundation grant DMS-0140540 and by an Alfred P. Sloan Fellowship. E.C. thanks the Institute for Pure and Applied Mathematics at UCLA and especially Mark Green and Eilish Hathaway for their warm hospitality.

References

- [1] G. Beylkin, R. Coifman, V. Rokhlin, Fast wavelet transforms and numerical algorithms, Comm. Pure Appl. Math. 44 (1991) 141-183.
- [2] E. Candès, L. Demanet, Curvelets, warpings and optimally sparse representations of Fourier Integral Operators, Manuscript, 2002.
- [3] E. Candès, D. Donoho, Curvelets a suprisingly effective nonadaptive representation for objects with edges, in: A. Cohen, C. Rabut, L. Schumaker (Eds.), Curves and Surface Fitting: Saint-Malo 1999, Vanderbilt University Press, Nashville, 2000, pp. 105–120.
- [4] E. Candès, D. Donoho, New tight Frames of curvelets and optimal representations of objects with C² singularities, submitted, http://www.acm.caltech.edu/~emmanuel/publications.html, 2002.
- [5] E. Candès, F. Guo, New multiscale transforms, minimum total variation synthesis: applications to edge-preserving image reconstruction, Signal Processing 82 (2002) 1519–1543.
- [6] C. Fefferman, A note on spherical summation multipliers, Israel J. Math. 15 (1973) 44-52.
- [7] H. Smith, A Hardy space for Fourier integral operators, J. Geom. Anal. 7 (1997).
- [8] H. Smith, A parametrix construction for wave equations with $C^{1,1}$ coefficients, Ann. Inst. Fourier (Grenoble) 48 (3) (1998) 797–835.