

WAVELETS, APPROXIMATION AND COMPRESSION**

MARTIN VETTERLI^{id}

TRANSLATORS: SARKOUT ABDI^{id}, ARAM AZIZI^{id}*, MAHMOUD SHAFIEE^{id} AND JAMSHID SAEIDIAN^{id}

ABSTRACT. Over the last decade or so, wavelets have had a growing impact on signal processing theory and practice, both because of their unifying role and their successes in applications. Filter banks, which lie at the heart of wavelet-based algorithms, have become standard signal processing operators, used routinely in applications ranging from compression to modems. The contributions of wavelets have often been in the subtle interplay between discrete-time and continuous-time signal processing. The purpose of this article is to look at recent wavelet advances from a signal processing perspective. In particular, approximation results are reviewed, and the implication on compression algorithms is discussed. New constructions and open problems are also addressed. Finding a good basis to solve a problem dates at least back to Fourier and his investigation of the heat equation. The series proposed by Fourier has several distinguishing features: The series is able to represent any finite energy function on an interval. The basis functions are eigenfunctions of linear shift invariant systems or, in other words, Fourier series diagonalize linear, shift invariant operators.

1. Introduction

Finding a good basis to solve a problem dates at least back to Fourier and his investigation of the heat equation [8]. The series proposed by Fourier has several distinguishing features:

- The series is able to represent any finite energy function on an interval.

Keywords: signal processing, filter banks, Fourier transform, wavelet transform, compress algorithm.

Article Type: Translation Paper.

Communicated by Alireza Amini Harandi.

*Corresponding author. Received: 31-05-2023, Accepted: 12-03-2024, Published Online: 19-10-2024.

The above abstract has been extracted by the translator from the original article (M. Vetterli, Wavelets, approximation and compression, *IEEE signal processing Magazine*, **18 no. 5 (2001) 35–71.)

<http://dx.doi.org/10.22108/msci.2024.137864.1581> .

- The basis functions are eigenfunctions of linear shift invariant systems or, in other words, Fourier series diagonalize linear, shift invariant operators.

Assume we have a space of functions S and we wish to represent an element $f \in S$. The space S can be, for example, integrable functions on the interval $[0, 1]$ with finite square integral

$$(1.1) \quad \int_0^1 |f(t)|^2 dt < \infty,$$

which we denote $L_2(0, 1)$. The first question is then to find a basis for S , that is a set of basis functions $\{\varphi_i\}_{i \in I}$ in S such that any element $f \in S$ can be written as a linear combination

$$(1.2) \quad f = \sum_{i \in I} a_i \varphi_i.$$

The example closest to the heart of signal processing people is certainly the expansion of bandlimited functions in terms of the sinc function. Given the representation of f in an orthonormal basis as in formula, its orthogonal projection into a fixed subspace of dimension N spanned by $\{\varphi_n\}_{n=0, \dots, N-1}$ is denoted $\hat{f}_N(t)$

$$(1.3) \quad \hat{f}_N = \sum_{n=0}^{N-1} \langle \varphi_n, f \rangle \varphi_n.$$

Given objects of interest and spaces in which they are embedded, we wish to know how fast an N -term approximation converges

$$(1.4) \quad \|f(t) - f_N(t)\|_2 \sim c t(N),$$

where $f_N(t)$ stands for an approximation of $f(t)$ which involves N elements, to be chosen appropriately. This immediately raises a number of questions. Different bases can give very different rates of approximation. Then, there are various ways to choose the N terms used in the approximation. A fixed subset (e.g., the first N terms) leads to a linear, subspace approximation as in (1.3). Adaptive schemes, to be discussed later, are nonlinear. Will different choices of the subset lead to different rates of approximation? Such questions are at the heart of approximation theory and are relevant when choosing a basis and an approximation method for a given signal processing problem. For example, denoising in wavelet bases has led to interesting results for piecewise smooth signals precisely because of the superior approximation properties of wavelets for such signals. We are now ready to address the last problem we shall consider, namely the compression problem. This involves not only approximation quality, but also description complexity. There is a cost associated with describing f_N , and this cost depends on the approximation method. Typically, the coefficient values and their locations need to be described, which involves quantization of the coefficients and indexing their locations.



2. Orthogonal Filter Banks

When thinking of filtering, one usually thinks about frequency selectivity. For example, an ideal discrete-time lowpass filter with cut-off frequency $\omega_c < \pi$ takes any input signal and projects it onto the subspace of signals bandlimited to $[-\omega_c, \omega_c]$. Orthogonal discrete-time filter banks perform a similar projection which we now review. Assume a discrete-time filter with finite impulse response $g_0[n] = \{g_0[1], g_0[1], g_0[L - 1]\}$, L even, and the property

$$(2.1) \quad \langle g_0[n], g_0[n - 2k] \rangle = \delta_k,$$

that is, the impulse response is orthogonal to its even shifts, and $\|g_0\|_2 = 1$. Denote by $G_0(z)$ the z -transform of the impulse response $g_0[n]$

$$(2.2) \quad G_0(z) = \sum_{n=0}^{L-1} g_0[n]z^{-n},$$

with an associated region of convergence covering the z -plane except the origin. Assume further that $g_0[n]$ is a lowpass filter, that is, its discrete-time Fourier transform has most of its energy in the region $[-\pi/2, \pi/2]$. Then define a high-pass filter $g_1[n]$ with z -transform $G_1(z)$ as follows:

$$(2.3) \quad G_1(z) = z^{-L+1}G_0(-z^{-1}).$$

3. Discrete-Time Polynomials and Filter Banks

Signal processing specialists intuitively think of problems in terms of sinusoidal bases. Approximation theory specialists think often in terms of other series, like the Taylor series, and thus, of polynomials as basic building blocks. We now look at how polynomials are processed by filter banks. A discrete-time polynomial signal of degree M is composed of a linear combination of monomial signals

$$(3.1) \quad \rho^{(m)}[n] = n^m, \quad 0 \leq m \leq M.$$

We shall now see that such monomial (and therefore polynomial) signals are eigensignals of certain multirate operators. We need to consider lowpass filters $G_0(z)$ that have a certain number $N > 0$ of zeroes at $z = 1$, or $\omega = \pi$ on the unit circle. That is, the filter factors as

$$(3.2) \quad G_0(z) = (1 + z^{-1})^N R_0(z).$$

Clearly, because of (2.3), the highpass $G_1(z)$ has N zeros at $z = 1$ or ($\omega = 0$), while $H_0(z)$ and $H_1(z)$ have N zeros at $z = 1$ and 1 , respectively because of formula.

4. Continuous-Time Polynomials and Wavelets

As is well known, a strong link exists between iterated filter banks and wavelets. For example, filter banks can be used to generate wavelet bases[4], and filter banks can be used to calculate wavelet series [9]. It comes thus as no surprise that the properties seen in discrete time regarding polynomial representation carry over to continuous time. While these properties are directly related to moment properties of wavelets and thus hold in general, we review them in the context of wavelets generated from orthogonal finite impulse response (FIR) filter banks. Assume again that the lowpass filter has N zeros at $\omega = \pi$, and thus, the highpass has N zeros at $\omega = 0$. From the two scale relation of scaling function and wavelet, we get that the Fourier transform of the wavelet can be factored as

$$(4.1) \quad \Psi(w) = \frac{1}{\sqrt{2}}G_1(l^{jw/2}) \cdot \phi\left(\frac{w}{2}\right).$$

where $\phi(w)$ is the Fourier transform of the scaling function.

5. Discontinuities in Filter Bank and Wavelet Representations

What happens if a signal is discontinuous at some point t_0 ? we know that Fourier series do not like discontinuities, since they destroy uniform convergence. Wavelets have two desirable properties as far as discontinuities are concerned. First, they focus locally on the discontinuity as we go to finer and finer scales. That is because of the scaling relation of wavelets, where the function set $\Psi_{m,n}(t)$ is defined as

$$(5.1) \quad \psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t - n) \quad m, n \in \mathbb{Z}$$

where $m \rightarrow -\infty$ corresponds to fine details. Thus, as m grows negative, the wavelet becomes “sharper.” If the discontinuity is isolated, and the surroundings are smooth, all wavelet inner products except the ones at the discontinuity will be zero, and around the discontinuity, L inner products are different from zero when the wavelet has support length L . Second, the magnitude evolution across scales of the nonzero wavelet inner products characterizes the discontinuity. This is a well-known characteristic of the continuous wavelet transform [5, 10] and holds as well for the orthonormal wavelet series.

6. Linear Approximation

Assume a space V and an orthonormal basis $\{g_n\}_{n \in N}$ for V . Thus, a function $f \in V$ can be written as a linear combination

$$(6.1) \quad f = \sum_{n \in N} \langle g_n, f \rangle g_n.$$



The best (in the squared error sense) linear approximation of f in the subspace V , denoted as \hat{f} , is given by the orthogonal projection of f onto a fixed subspace of V (see (1.3)). Assume it is a subspace W of dimension M , and spanned by the first M vectors of the basis, or

$$W = \text{span}\{g_0, g_1, g_2, \dots, g_{M-1}\}$$

. Then \hat{f} is given by

$$(6.2) \quad \hat{f}_M = \sum_{n=0}^{M-1} \langle g_n, f \rangle g_n.$$

and the squared error of the approximation is

$$\hat{\epsilon} = \|f - \hat{f}\|_2^2 = \sum_{n=M}^{\infty} |\langle g_n, f \rangle|^2.$$

Because the subspace W is fixed, independent of f , the approximation is linear.

7. Nonlinear Approximation in Orthonormal Bases

Consider the same set up as above, but with a different approximation rule. Instead of (6.1), where the first M coefficients in the orthonormal expansion are used, we keep the M largest coefficients instead. That is, we define an index set I_M of the M largest inner products, or

$$(7.1) \quad |\langle g_m, f \rangle| \geq |\langle g_n, f \rangle|, \quad m \in I_M, n \notin I_M.$$

Then, we define, the best nonlinear approximation as:

$$(7.2) \quad \hat{f}_M = \sum_{n \in I_M} \langle g_n, f \rangle g_n$$

which leads to an approximation error

$$\tilde{\epsilon}_M = \|f - \tilde{f}\|_2^2 = \sum_{n \notin I_M} |\langle g_n, f \rangle|^2$$

Clearly

$$(7.3) \quad \tilde{\epsilon}_M \leq \hat{\epsilon}_M.$$

(We could call this adaptive linear approximation or adaptive subspace approximation. However, the commonly used term is nonlinear approximation. More general nonlinear schemes could also be considered, but are beyond the scope of this article.)

8. Nonlinear Approximation and Compression

So far, we have considered keeping M elements from a basis, either a fixed set (the first M typically) or an adaptive set (corresponding to the largest projection). In compression, we have first to describe the coefficient set, which has zero cost when this set is fixed (linear approximation) but has a nontrivial cost when it is adaptive (nonlinear approximation). In that case, there are $\binom{N}{M}$ possible subsets, and the rate to describe the subset is equal to the entropy of the distribution of the subsets, or at most

$$(8.1) \quad \log_2 \binom{N}{M}$$

9. Compression of Piecewise Polynomial Signals

Let us return to one-dimensional piecewise smooth signals. Wavelets are well suited to approximate such signals when nonlinear approximation is allowed. To study compression behavior, consider the simpler case of piecewise polynomials, with discontinuities. To make matters easy, let us look again at the signal we used earlier to study nonlinear approximation, but this time include quantization and bit allocation. A simple analysis of the approximate rate distortion behavior of a step function goes as follows. Coefficients decay as $2^{m/2}$, so the number of scales J involved, if a quantizer of size Δ is used, is of the order of $\log_2(1/\Delta)$. The number of bits per coefficient is also of the order of $\log_2(1/\Delta)$, so the rate R is of the order

$$(9.1) \quad R \sim \log_2^2(1/\Delta) \sim J^2$$

The distortion or squared error is proportional to Δ^2 (for each coefficient), times the number of scales or, using $\Delta = 2^{-J}$

$$(9.2) \quad D \sim J \cdot 2^{-J}.$$

Using (9.1), we get

$$(9.3) \quad D_W(R) = C_1 \cdot \sqrt{R} \cdot 2^{C_2 \sqrt{R}}.$$

for the distortion-rate behavior of a wavelet scheme. Note that we ignored the cost of indexing the location. This cost turns out to be quite small (order J), because the coefficients are all gathered around the discontinuity.

10. True Two-Dimensional Bases

As the wavelet example shows, separable bases are not suited for “true” two-dimensional objects. What is needed are transforms and bases that include some form of “geometry” and that are truly two dimensional. (The notion of geometry is not easy to formalize in our context, but the intuition is that the dimensions are not independent, and certain shapes are more likely than others.) Besides

the two-dimensional Fourier and wavelet transforms, which are both separable, the Radon transform plays a key role. This transform, studied early in the 20th century [12], was rediscovered several times in fields ranging from astronomy to medical imaging (see [6] for an excellent overview). The Radon transform maps a function $f(x, y)$ into $RA_f(\theta, t)$ by taking line integrals at angle θ and location t

$$(10.1) \quad RA_f(\theta, t) = \iint f(x, y)\delta(x \cos \theta + y \sin \theta - t)dx dy$$

A key insight to construct directional bases from the Radon transform was provided by Candès and Donoho [2, 3], with the ridgelet transform. The idea is to map a one-dimensional singularity, like a straight edge, into a point singularity using the Radon transform. Then, the wavelet transform can be used to handle the point singularity.

11. Directional Filter Banks

To get directional analysis, one can alternatively use directional filter banks[1]. In such a case, the basis functions are given by the filter impulse responses and their translates with respect to the subsampling grid. Such filter banks can be designed directly, or through iteration of elementary filter banks. They lead to bases if critically subsampled, or frames if oversampled.

12. Two-Dimensional Bases and Compression

As we had seen in one dimension, a good N -term approximation is not yet a guarantee for good compression. While a powerful N -term approximation is desirable, it must be followed by appropriate compression. Thus, the topic of compression of two-dimensional piecewise smooth functions is still quite open. Several promising approaches are currently under investigation, including compression in ridgelet and curvelet domain, compression along curves using “bandelelets”[11] and generalization of footprints in two dimensions or edgeprints [7].

13. Conclusions

The interplay of representation, approximation, and compression of signals was reviewed. For piecewise smooth signals, we showed the power of wavelet-based methods, in particular for the one-dimensional case. For two-dimensional signals, where wavelets do not provide the answer for piecewise smooth signals with curve singularities, new approaches and open problems were indicated. Such approaches rely on new bases with potentially high impact on image processing, for such problems as denoising, compression and classification.

REFERENCES

- [1] R. H. Bamberger and M. J. T. Smith, A filter bank for the directional decomposition of images: Theory and design, *IEEE Trans. Signal Processing*, **40** (1992) 882–893.
- [2] E. Candès, Ridgelets: Theory and applications, Ph.D. dissertation, Dept. Statistics, Stanford University, Stanford, CA, 1998.
- [3] E. J. Candès and D. L. Donoho, Curvelets—A surprisingly effective nonadaptive representation for objects with edges, in *Curve and Surface Fitting*, A. Cohen, C. Rabutand, and L. L. Schumaker, Eds. Saint-Malo: Vanderbilt University Press, 1999.
- [4] I. Daubechies, Orthonormal bases of compactly supported wavelets, *Commun. Pure Appl. Math.*, **41** (1988) 909–996.
- [5] I. Daubechies, *Ten Lectures on Wavelets*, Philadelphia, PA: SIAM, 1992.
- [6] S. R. Deans, *The Radon Transform and Some of its Applications*, New York: Wiley, 1983.
- [7] P. L. Dragotti and M. Vetterli, Footprints and edgeprints for image denoising and compression, in *Proc. IEEE Int. Conf. Image Processing, ICIP 2001*, Patras, Greece, 2001.
- [8] J. Fourier, *Théorie Analytique de la Chaleur*, Paris, France: Gauthier-Villars, 1888.
- [9] S. Mallat, A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Trans. Pattern Recognition Machine Intell.*, **11** (1989) 674–693.
- [10] S. Mallat and W. L. Hwang, Singularity detection and processing with wavelets, *IEEE Trans. Inform. Theory (Special Issue on Wavelet Transforms and Multiresolution Signal Analysis)*, **38** (1992) 617–643.
- [11] E. Le Pennec and S. Mallat, Image compression with geometric wavelets, in *Proc. IEEE Int. Conf. Image Processing, ICIP 2000*, Vancouver, Canada, Sept. (2000) 661–664.
- [12] J. Radon, Ueber die bestimmung von funktionen durch ihre integralwerte längst gewisser mannigfaltigkeiten, *Berichte Sächsische Akademie der Wissenschaften, Leipzig*, (1917) 262–267.

Sarkout Abdi

Department of Mathematics, Rasht Branch, Islamic Azad University, Rasht, Iran

Email: sarkout_abdi2011@yahoo.com

Aram Azizi

Department of Mathematics, Payame Noor University, P.O. Box 19395-3697 Tehran, Iran

Email: a.azizi@pnu.ac.ir

Mahmoud Shafiee

Department of Mathematics, Rasht Branch, Islamic Azad University, Rasht, Iran

Email: hafiei@iaiu.ac.ir

Jamshid Saeidian

Faculty of Mathematical Sciences and Computer, Kharazmi University, 50 Taleghani avenue, Tehran 1561836314, Iran

Email: j.saeidian@khu.ac.ir

موجک‌ها، تقریب و فشرده‌سازی

مارتین وترلی^{id}

مترجمان: سهرکهوت عبدی^{id}، آرام عزیزی^{id}،* محمود شفیعی^{id} و جمشید سعیدیان^{id}

چکیده. نوشته حاضر ترجمه مقاله زیر است:

M. Vetterli, Wavelets, approximation and compression, *IEEE signal processing Magazine*, 18 no. 5 (2001) 59-73.

۱. مقدمه

در طول یک دهه‌ی گذشته یا بیشتر، موجک‌ها تأثیر فزاینده‌ای بر نظریه‌ی پردازش سیگنال و کاربردهای آن، هم به دلیل نقش یکپارچه‌کننده و هم موفقیت‌هایشان در کاربرد داشته‌اند (همچنین به [۴۲] و [۳۸] در این شماره بنگرید). فیلتر بانک‌ها که در قلب الگوریتم‌های مبتنی بر موجک‌ها قرار دارند، به عملگرهای استاندارد برای پردازش سیگنال بدل شده‌اند که به‌طور معمول در کاربردهای گوناگونی از فشرده‌سازی^۱ تا مودم (تلفیق و تفکیک) به‌کار می‌روند. کارکرد موجک‌ها اغلب در تقابل ظریف بین پردازش سیگنال‌های زمان‌گسسته و زمان‌پیوسته بوده است. هدف مقاله‌ی حاضر بررسی پیشرفت‌های اخیر موجک‌ها از دیدگاه پردازش سیگنال است. به‌طور خاص، نتایج تقریب مورد بررسی قرار گرفته شده و مفهوم الگوریتم‌های فشرده‌سازی مورد بحث قرار می‌گیرند. همچنین ساز و کارهای جدید و مسائل باز را نیز مورد توجه قرار می‌دهیم.

۲. پایه‌ها، تقریب و فشرده‌سازی

پیدا کردن یک پایه خوب برای حل یک مسئله، حداقل به زمان فوریه^۲ و بررسی معادله‌ی گرمایی او برمی‌گردد [۱۸]. سری پیشنهادی فوریه چندین ویژگی متمایز دارد:

- این سری می‌تواند هر تابع با انرژی محدود را در یک بازه نشان دهد.
- توابع پایه، توابع ویژه سیستم‌های خطی نسبت به انتقال مانا هستند، یا به عبارت دیگر، سری فوریه عملگرهای خطی نسبت به انتقال مانا را قطری می‌کند.

به‌طور مشابه، پایه‌های سینک به‌کار رفته در قضیه‌ی سامپلینگ می‌تواند هر تابع باند محدود را نمایش دهد. همچنین پردازش را می‌توان به‌جای خود تابع روی نمونه‌ها انجام داد. به‌طور خلاصه، یک پایه هم به‌دلیل توانایی آن برای نمایش یک شیء مورد نظر

عبارات و کلمات کلیدی: پردازش سیگنال، فیلتر بانک‌ها، تبدیل فوریه، تبدیل موجک، تبدیلات گسسته و پیوسته، الگوریتم فشرده‌سازی، تقریب خطی و غیرخطی. دبیرتخصصی رابط: علیرضا امینی‌هرندی
نوع مقاله: ترجمه‌ای

تاریخ دریافت: ۱۴۰۲/۰۳/۱۰ تاریخ پذیرش: ۱۴۰۲/۱۲/۲۲ تاریخ انتشار آنلاین: ۱۴۰۳/۰۷/۲۸

ارجاع به مقاله: مارتین وترلی، مترجمان: س. عبدی، آ. عزیزی، م. شفیعی و ج. سعیدیان، موجک‌ها، تقریب و فشرده‌سازی، نشریه ریاضی و جامعه، ۳ شماره ۹ (۱۴۰۳) ۷۱-۳۵.
<http://dx.doi.org/10.22108/msci.2024.137864.1581>

¹compression ²Fourier

(به‌عنوان مثال، تقریب خوب با ضرایب کم) و هم برای ارزش عملیاتی آن (به‌عنوان مثال، قطری کردن عملگرهای خاص) انتخاب می‌شود.

به‌طور دقیق‌تر، فرض کنید فضایی از توابع S داریم و می‌خواهیم عنصر $f \in S$ را نشان دهیم. برای مثال، فضای S می‌تواند توابعی انتگرال‌پذیر روی بازه‌ی $[0, 1]$ با انتگرال مجذوری متناهی باشد. یعنی

$$(1) \quad \int_0^1 |f(t)|^2 dt < \infty,$$

که آن را با $L_2(0, 1)$ نشان می‌دهیم. اولین سؤال این است که چگونه پایه‌هایی برای S پیدا کنیم که مجموعه‌ای از توابع پایه $\{\varphi_i\}_{i \in I}$ در S باشند به‌طوری‌که هر عنصر $f \in S$ را بتوان به‌صورت ترکیب خطی زیر نوشت

$$(2) \quad f = \sum_{i \in I} a_i \varphi_i.$$

ملموس‌ترین مثال برای افرادی که پردازش سیگنال انجام می‌دهند قطعاً بسط توابع باند محدود بر حسب توابع سینک است. فضایی حقیقی از توابع باند محدود در دامنه‌ی فوریه $[-\pi, \pi]$ و انتگرال مجذوری متناهی را نظر بگیرید. این فضا را با $BL_2(-\pi, \pi)$ نشان می‌دهیم. سپس قضیه سامپلینگ شانون [۲۸، ۳۶] می‌گوید هر تابعی در آن فضا را می‌توان به‌صورت زیر نوشت

$$(3) \quad f(t) = \sum_{n=-\infty}^{\infty} a_n \sin(t - n),$$

که $a_n = f(n)$ مقادیر $f(t)$ در اعداد صحیح هستند و

$$(4) \quad \sin C(t) = \frac{\sin(\pi t)}{\pi t}$$

و انتقال‌های صحیح آن‌ها، توابع پایه هستند.

یک پرسش خیلی مهم این است که در چه شرایطی (۳) درست است. $f_N(t)$ را به‌صورت تقریب N جمله‌ای که N یک عدد صحیح فرد است، در نظر می‌گیریم

$$\hat{f}_N(t) = \sum_{n=-(N-1)/2}^{(N-1)/2} a_n \sin C(t - n).$$

همچنین یک نرم برای اندازه‌گیری تقریب خطا نیاز داریم که در این مقاله ما صرفاً به L_2 -نرم خطا می‌پردازیم.

$$(5) \quad \|f(t) - \hat{f}_N(t)\|_2 = \left(\int_{-\infty}^{\infty} |f(t) - \hat{f}_N(t)|^2 dt \right)^{1/2}.$$

سپس ادعا می‌کنیم که (۳) در صورتی صحیح است که:

$$(6) \quad \lim_{n \rightarrow \infty} \|f(t) - \hat{f}_N(t)\|_2 = 0.$$

اگر مجموعه‌ای از توابع $\{\varphi_i\}_{i \in I}$ داشته باشیم به‌گونه‌ای که تمام توابع در S را بتوان به‌طور دلخواه دقیقاً مانند (۶) تقریب زد، آن‌گاه می‌گوییم که این مجموعه در S کامل است. علاوه بر این، اگر عناصر آن مستقل خطی باشند، می‌گوییم که $\{\varphi_i\}_{i \in I}$ پایه‌ای برای S است. در بین تمامی پایه‌های ممکن، پایه‌های متعامد از مطلوبیت خاصی برخوردارند. در این حالت، توابع پایه یا بردارهای پایه همگی متعامد هستند. با نرمال کردن آن‌ها به ۱، یا $\|\varphi_i\| = 1, i \in I$ ، مجموعه‌ای متعامد نرمال از بردارها به‌صورت زیر خواهیم داشت:

$$(7) \quad \langle \varphi_i, \varphi_j \rangle = \delta_{ij}$$

که ضرب داخلی $\langle \cdot, \cdot \rangle$ ، در زمان پیوسته به صورت

$$(۸) \quad \langle \varphi_i, \varphi_j \rangle = \int_{-\infty}^{\infty} \varphi_i^*(t) \varphi_j(t) dt$$

و در زمان گسسته به صورت زیر تعریف می‌شوند

$$(۹) \quad \langle \varphi_i, \varphi_j \rangle = \sum_{n \in \mathbb{Z}} \varphi_i^*[n] \varphi_j[n].$$

برای سیگنال‌های زمان گسسته، فضای دنباله‌های مجذور متناهی را که با $l_2(Z)$ نشان داده شده‌اند در نظر می‌گیریم. برای پایه‌های متعامد فرمول بسط (۲) به صورت زیر در می‌آید

$$(۱۰) \quad f = \sum_{i \in I} \langle \varphi_i, f \rangle \varphi_i.$$

این یک فرمول برای نمایش f در پایه‌های متعامد $\{\varphi_i\}_{i \in I}$ است. (فرمول‌های عمومی‌تر، برای مثال در پایه‌های دو متعامد یا در قاب‌ها نیز ممکن است، اما برای بحث ما مورد نیاز نیستند.) با توجه به نمایش f در یک پایه‌ی متعامد نرمال مانند (۱۰)، تصویر متعامد آن به یک زیرفضای ثابت با بعد N که با $\{\varphi_n\}_{n=0, \dots, N-1}$ بسط داده شده است به وسیله‌ی $\hat{f}_N(t)$ نشان داده می‌شود

$$(۱۱) \quad \hat{f}_N = \sum_{n=0}^{N-1} \langle \varphi_n, f \rangle \varphi_n.$$

این تقریب خطی f به وسیله‌ی تصویر بر روی یک زیرفضای ثابت با بعد N به دست می‌آید. این تقریب خطی است زیرا همان‌طور که به راحتی می‌توان بررسی کرد، تقریب $\alpha f + \beta g$ برابر است با مجموع وزنی تقریب‌ها، $\alpha \hat{f} + \beta \hat{g}$. اما این تنها یکی از تقریب‌های فراوان N جمله‌ای ممکن f است. بنابراین، مسأله تقریب در اینجا ظاهر می‌شود: با توجه به هدف مطلوب و فضاهایی که در آن‌ها تعبیه شده است، می‌خواهیم بدانیم یک تقریب N جمله‌ای با چه سرعتی همگرا می‌شود

$$(۱۲) \quad \|f(t) - f_N(t)\|_2 \sim fct(N),$$

که در آن $f_N(t)$ تقریبی از $f(t)$ است که شامل N عنصر است که باید به درستی انتخاب شود. این موضوع بلافاصله تعدادی پرسش را مطرح می‌کند. پایه‌های مختلف می‌توانند نرخ‌های تقریب بسیار متفاوتی را ارائه دهند. بنابراین، راه‌های مختلفی برای انتخاب N جمله‌ای مورد استفاده در تقریب وجود دارد. یک زیرمجموعه‌ی ثابت (به عنوان مثال، اولین N جمله) به یک تقریب خطی زیرفضا مانند (۱۱) منجر می‌شود طرح‌های تطبیقی، که بعداً مورد بحث قرار خواهد گرفت، غیرخطی هستند. آیا انتخاب‌های مختلف N جمله مورد نظر به نرخ‌های تقریب متفاوتی منجر می‌شود؟ چنین سؤالاتی در قلب تئوری تقریب قرار دارند و هنگام انتخاب یک پایه و روش تقریبی برای یک مسئله‌ی پردازش سیگنال معین ظاهر می‌شوند. برای مثال، نویززدایی با پایه‌های موجک، دقیقاً به دلیل خواص تقریبی برتر موجک‌ها برای چنین سیگنال‌هایی، به نتایج مطلوبی برای سیگنال‌های تکه‌ای هموار منجر شده است.

حال آمادگی لازم برای حل مسئله‌ی مطرح شده را داریم، یعنی مسئله‌ی فشرده‌سازی را بررسی کنیم. این نه تنها شامل کیفیت تقریب، بلکه پیچیدگی آن نیز می‌شود. برای بیان f_N بهایی وجود دارد که به روش تقریب بستگی دارد. به طور معمول، مقادیر ضرایب و مکان آن‌ها باید آورده شود، که شامل کمی‌سازی ضرایب و اندیس‌گذاری مکان آن‌ها است. با فراخوانی R به عنوان تعداد بیت‌های استفاده شده در توضیحات، می‌توانیم تابع $\hat{D}(R)$ را به صورت زیر تعریف کنیم:

$$(۱۳) \quad \hat{D}(R) = \|f - f_R\|_2,$$

که در آن f_R به صراحت نشان می‌دهد که R بیت برای تقریب f استفاده می‌شود. (در این مرحله، باید یک موضوع را در رابطه با کلاس توابعی که تقریب می‌زنیم معرفی کنیم. با توجه به متون، تقریب یک تابع منفرد در مفهوم نظری تئوری اطلاعات، پیچیدگی صفر دارد. برای سادگی، فعلاً این را به صورت ضمنی رها می‌کنیم.) البته، f_R به پایه‌ی انتخاب شده برای نشان دادن f بستگی دارد و بنابراین $\hat{D}(R)$ بسته به پایه، می‌تواند رفتارهای بسیار متفاوتی داشته باشد. در حالی که $\hat{D}(R)$ شبیه تابع نرخ پیچیدگی $D(R)$ تعریف شده در تئوری اطلاعات است، این دو تابع به هیچ وجه یکسان نیستند. $D(R)$ به عنوان مینیمم پیچیدگی در هر طرح کدگذاری با استفاده از R بیت تعریف می‌شود (برای مثال با استفاده از بلوک‌های طولانی دلخواه نمونه‌ها در زمان گسسته) در حالی که $\hat{D}(R)$ به طور خاص با تقریب توسط پایه‌ها تنظیم می‌شود. در موارد خاص، $\hat{D}(R)$ می‌تواند از همان مرتبه $D(R)$ باشد، در حالی که در موارد دیگر، مینیمم نظری داده شده توسط $D(R)$ می‌تواند بسیار بهتر باشد. با این حال، $\hat{D}(R)$ معمولاً در عمل با پیچیدگی مناسب قابل دستیابی است، در حالی که $D(R)$ اغلب در یک حد غیرقابل دستیابی باقی می‌ماند. (همچنین باید توجه داشت که تابع نرخ پیچیدگی واقعی فقط در موارد بسیار کمی شناخته شده است.) به عنوان جمع‌بندی این بخش، سه مسئله‌ی مرتبط با پایه‌ها در فضاها را بررسی کردیم. اولی وجود یک پایه، دومی قدرت تقریب یک پایه و سومی قدرت فشرده‌سازی یک پایه بود. در ادامه، رابطه‌ی این مسائل را، به ویژه برای پایه‌های موجک، بررسی خواهیم کرد. مقاله مروری [۱۵] همچنین رابطه‌ی بین آنالیز هارمونیک، روش‌های محاسباتی و فشرده‌سازی سیگنال را مورد بحث قرار می‌دهد.

۳. ویژگی‌های تقریبی فیلتر بانک‌ها و موجک‌ها

اگرچه فیلتر بانک‌های متعامد که در کدگذاری زیر باندها استفاده می‌شوند توسط یوسویچ^۳ مورد بحث قرار گرفته‌اند [۴۲]، در اینجا با تأکید بر ویژگی‌های تصویری و تقریبی آن‌ها را بررسی می‌کنیم.

۱.۳. فیلتر بانک‌های متعامد. وقتی سخن از فیلتر کردن باشد، معمولاً انتخاب فرکانس به ذهن می‌آید. به عنوان مثال، یک فیلتر پایین‌گذر^۴ زمان گسسته‌ی ایده‌آل با فرکانس قطع $\omega_c < \pi$ هر سیگنال ورودی را می‌گیرد و آن را در زیرفضای سیگنال‌های بانده محدود به $[-\omega_c, \omega_c]$ تصویر می‌کند. بانک‌های فیلتر زمان گسسته متعامد نیز طرح مشابهی را انجام می‌دهند که اکنون بررسی می‌کنیم. یک فیلتر زمان گسسته با پاسخ ضربه متناهی $\{g_0[n], g_0[n-1], \dots, g_0[n-L+1]\}$ با L زوج و ویژگی زیر را در نظر بگیرید

$$(14) \quad \langle g_0[n], g_0[n-2k] \rangle = \delta_k,$$

یعنی پاسخ ضربه نسبت به انتقال‌های زوج آن متعامد است و $\|g_0\|_2 = 1$ با $G_0(z)$ تبدیل z پاسخ ضربه‌ی $g_0[n]$ به صورت

$$(15) \quad G_0(z) = \sum_{n=0}^{L-1} g_0[n] z^{-n},$$

با یک ناحیه‌ی همگرایی مرتبط که صفحه‌ی z را به جز مبدأ پوشش می‌دهد. همچنین فرض کنید که $g_0[n]$ یک فیلتر پایین‌گذر است، یعنی تبدیل فوریه‌ی زمان گسسته‌ی آن بیشتر انرژی خود را در منطقه‌ی $[-\pi/2, \pi/2]$ دارد. سپس یک فیلتر بالاگذر $g_1[n]$ را با تبدیل $G_1(z)$ به صورت زیر تعریف می‌کنیم:

$$(16) \quad G_1(z) = z^{-L+1} G_0(-z^{-1}).$$

سه عملیات زیر در نظر گرفته شده است:

- $z \rightarrow -z$ مربوط به مدولاسیون توسط $(-1)^n$ است، به عبارت دیگر پایین‌گذر را به بالاگذر تبدیل می‌کند.
- $-z \rightarrow -z^{-1}$ وارون زمانی را به پاسخ ضربه بدل می‌کند.
- ضرب در z^{-L+1} ، وارون زمانی پاسخ ضربه را ایجاد می‌کند.

³Usevitch ⁴low-pass filter

به‌عنوان مثال، اگر $g_0 = \{a, b, c, d\}$ ، آن‌گاه $g_1[n] = \{-d, c, -b, a\}$. این روش ویژه برای به دست آوردن یک بالاگذر از یک پایین‌گذر، که به عنوان فیلترهای مربعی مزدوج (QCF) معرفی شده است [۲۷، ۳۹]، دارای ویژگی‌های زیر است:

$$(۱۷) \quad \langle g_1[n], g_1[n - 2k] \rangle = \delta_k,$$

یعنی $g_1[n]$ نیز نسبت به جابجایی زوج خود متعامد است و

$$(۱۸) \quad \langle g_0[n], g_1[n - 2k] \rangle = 0$$

یا $\{g_0[n], g_1[n]\}$ و انتقال‌های زوج آن‌ها دوبه‌دو متعامد هستند. آخرین مرحله نشان دادن این است که مجموعه‌ی متعامد

$$(۱۹) \quad \{g_0[n - 2k], g_1[n - 2l]\}_{k, l \in \mathbb{Z}}$$

یک پایه‌ی متعامد برای $l_2(\mathbb{Z})$ است. این امر مستلزم بررسی کامل بودن است که برای مثال می‌توان با تأیید رابطه‌ی پارسوال [۴۴] انجام داد. بنابراین، هر دنباله‌ای از $l_2(\mathbb{Z})$ را می‌توان به‌صورت زیر نوشت

$$(۲۰) \quad x[n] = \sum_{k \in \mathbb{Z}} \alpha_k \cdot g_0[n - 2k] + \sum_{l \in \mathbb{Z}} \beta_l \cdot g_1[n - 2l],$$

که

$$(۲۱) \quad \alpha_k = \langle g_0[n - 2k], x[n] \rangle \quad k \in \mathbb{Z},$$

$$(۲۲) \quad \beta_l = \langle g_1[n - 2l], x[n] \rangle \quad l \in \mathbb{Z},$$

حال چگونه فرمول بسط ((۲۰)) را پیاده‌سازی کنیم؟ به سادگی از یک فیلتر بانک چند نرخی همانطور که در شکل ۱ نشان داده شده است استفاده کنید. فیلترهای آنالیز $H_0(z)$ و $H_1(z)$ نسخه‌های زمان‌وارون $G_0(z)$ و $G_1(z)$ هستند، یا

$$(۲۳) \quad H_i(z) = \tilde{G}_i(z) = \tilde{G}_i(z^{-1}), \quad i = 0, 1.$$

تأیید اینکه تحلیل فیلتر بانک‌ها (قسمت سمت چپ شکل ۱) درحقیقت ضرایب α_k و β_l را محاسبه می‌کند و اینکه بسط فیلتر بانک (قسمت سمت راست شکل ۱) بازآرایی را مانند (۲۰) محقق می‌سازد، سخت نیست. به یاد داشته باشید که پیچش شامل برگشت زمان است.

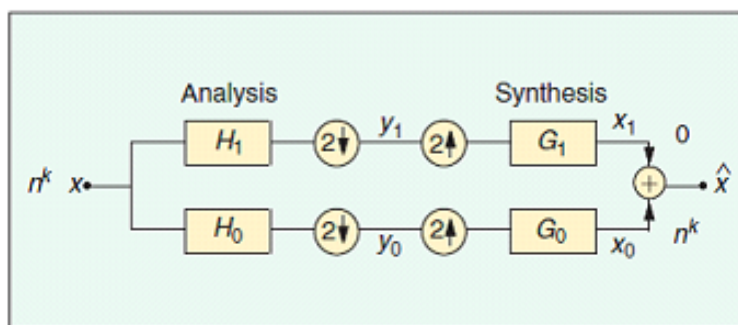
همان‌طور که در شکل ۲ نشان داده شده است، نگاه دومی به شاخه تحلیل/ترکیب پایین‌گذر خواهیم داشت. با استفاده از آنالیز استاندارد پردازش سیگنال چندنرخی، نتیجه‌ی چنین فیلترینگ با نمونه برداری کاهش‌ی-نمونه‌برداری افزایشی منجر به خروجی $y_0[n]$ با تبدیل z زیر می‌شود

$$(۲۴) \quad y_0(z) = \frac{1}{2} \cdot G_0(z)[H_0(z)X(z) + H_0(-z)X(-z)],$$

یعنی هم یک نسخه فیلتر شده $X(z)$ و هم یک نسخه دگرنام (الایزینگ) فیلتر شده شامل $X(-z)$ است. می‌توان نشان داد که عملگرهای از $x[n]$ تا $y_0[n]$ خطی، خودالحاقی و خودتوان هستند، به این معنی که این تصویر متعامد فضای ورودی $l_2(\mathbb{Z})$ بر روی یک زیرفضای بسط داده شده توسط $\{g_0[n - 2k]\}_{k \in \mathbb{Z}}$ است. ما این زیرفضا را V_1 می‌نامیم، که $\{g_0[n - 2k]\}_{k \in \mathbb{Z}}$ یک پایه‌ی متعامد را برای آن تشکیل می‌دهد.

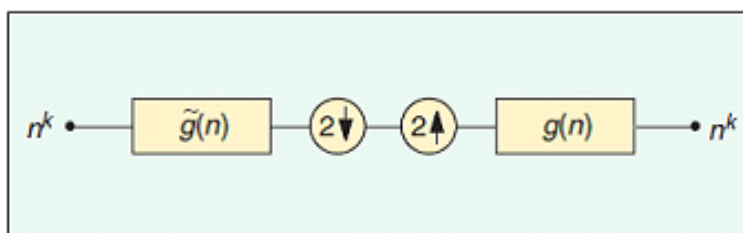
به‌طور مشابه با در نظر گرفتن شاخه‌ی پایین‌گذر، می‌بینیم که یک تصویر از $l_2(\mathbb{Z})$ را بر روی زیرفضای W_1 که توسط $\{g_1[n - 2k]\}_{k \in \mathbb{Z}}$ بسط داده شده‌اند، تولید می‌کند. فضای V_1 و W_1 متعامد ((۲۰)) را ببینید) هستند و جمع مستقیم آن‌ها $l_2(\mathbb{Z})$ را تشکیل می‌دهد، یعنی

$$(۲۵) \quad l_2(\mathbb{Z}) = W_1 + V_1.$$



شکل ۱. بانک فیلتر چند نرخیه و تحلیل دو کانال همراه با تلفیق‌کننده.

Figure 1: Multirate filter bank. Two channel analysis followed by synthesis.



شکل ۲. شاخه‌ی پایین‌گذر بانک فیلتر دو کاناله در شکل ۱. چندجمله‌ای زمان گسسته n^k یک سیگنال ویژه از این عملگر است.

Figure 2: Lowpass branch of the two channel filter bank in 1. The discrete-time polynomial n^k is an eigensignal of this operator.

بنابراین یک فیلتر بانک متعامد، فضای ورودی را به دو زیر فضای تقریب پایین‌گذر V_1 و جزء بالاگذر عمود بر آن W_1 تقسیم می‌کند. فضای V_1 مربوط به یک تقریب درشت است، در حالی که W_1 حاوی جزئیات اضافی است. این اولین مرحله در تجزیه و تحلیل چند ریزه‌سازی^۵ است که هنگام تکرار تقسیم بالاگذر/پایین‌گذر در شاخه پایین‌گذر به دست می‌آید (شکل ۳ را ببینید).

۴. چندجمله‌ای‌های زمان‌گسسته و فیلتر بانک‌ها

متخصصان پردازش سیگنال به طور شهودی به مسئله بر اساس پایه‌های سینوسی فکر می‌کنند. متخصصان نظریه تقریب اغلب به سری‌های دیگرمانند سری تیلور فکر می‌کنند و بنابراین چند جمله‌ای‌ها را بلوک‌های اساسی می‌دانند. اکنون به نحوه‌ی پردازش چندجمله‌ای‌ها توسط فیلتر بانک‌ها نگاه می‌کنیم. یک سیگنال چندجمله‌ای زمان‌گسسته با درجه M از ترکیب خطی سیگنال‌های تک جمله‌ای به صورت زیر تشکیل شده است:

$$(۲۶) \quad \rho^{(m)}[n] = n^m, \quad 0 \leq m \leq M.$$

اکنون خواهیم دید که چنین سیگنال‌های تک جمله‌ای (و در نتیجه چندجمله‌ای) سیگنال‌های ویژه (توابع ویژه) عملگرهای چند نرخیه خاص هستند. ما باید فیلترهای پایین‌گذر $G_0(z)$ را در نظر بگیریم که تعداد مشخص $N > 0$ ، صفر در $z = -1$ یا $\omega = \pi$ بر

⁵multiresolution

روی دایره واحد دارند. یعنی ضرایب فیلتر به‌عنوان

$$(27) \quad G_0(z) = (1 + z^{-1})^N R_0(z).$$

واضح است که به دلیل (۱۶)، بالاگذر $G_1(z)$ ، N صفر در $z = 1$ (یا $\omega = 0$) دارد، در حالی که به دلیل (۲۳)، $H_0(z)$ و $H_1(z)$ به ترتیب در $z = 1$ و $z = -1$ دارای N صفر هستند. سپس، تک‌جمله‌ای‌ها درجه $N - 1$ ، سیگنال‌های ویژه شاخه‌ی پایین‌گذر نشان‌داده‌شده در شکل ۲ هستند که با فیلتر کردن $H_0(z) = G_0(z^{-1})$ ، نمونه‌برداری کاهشی با ۲، نمونه‌برداری افزایشی با ۲ و در نهایت فیلتر کردن توسط $G_0(z)$ به‌دست آمده‌اند. این نتیجه پیامد قضیه‌ی نقطه ثابت استرانگ است (برای مثال [۴۱] را ببینید)، و ما در اینجا یک استدلال شهودی ارائه می‌دهیم. به دلیل اینکه شاخه‌ی بالاگذر دارای فیلتر تحلیل زیراست

$$(28) \quad H_1(z) = z^{L-1} G_0(-z) = z^{L-1} (1 - z^{-1}) \cdot R_0(-z),$$

با N صفر در $z = 1$ ، همه‌ی چندجمله‌ای‌ها تا درجه $N - 1$ در این شاخه صفر می‌شوند. این نتیجه قضیه گشتاور تبدیل فوریه است [۴۴، ۲۸]. از آنجایی که فیلتر بانک دارای خاصیت بازسازی کامل است، بنابراین این چندجمله‌ای‌ها به‌طور کامل تنها توسط شاخه‌ی پایین‌گذر تکرار می‌شوند و نشان می‌دهد که آن‌ها سیگنال ویژه هستند. شایان ذکر است که سینوسی‌ها، سیگنال‌های ویژه عملگرهای چندرخی نیستند. پیامد این خاصیت سیگنال ویژه این است که می‌توانیم بنویسیم

$$(29) \quad n^m = \sum_{k=-\infty}^{\infty} \alpha_k^{(m)} \cdot g_0[n - 2k] \quad m \leq N,$$

که $\alpha_k^{(m)}$ ضرایب مناسب هستند. این نشان می‌دهد که تک‌جمله‌ای‌ها تا درجه‌ی $N - 1$ در مجموعه تولید شده توسط $\{g_0[n - 2k]\}_{k \in \mathbb{Z}}$ هستند. (توجه داشته باشید که چندجمله‌ای‌ها در $l_2(\mathbb{Z})$ نیستند، اما تحدید آنها برای هر بازه متناهی در $l_2(z)$ هستند. بنابراین (۲۹) برای چندجمله‌ای‌های تصویر شده تا مرزها صادق است.) اکنون یک بانک فیلتر تکراری را همانطور که در شکل ۳ نشان داده شده است در نظر بگیرید. پس از J مرحله تجزیه و بازسازی، فیلتر پایین‌گذر بازسازی، $g_0^{(J)}[n]$ دارای تبدیل z به‌صورت زیر است

$$(30) \quad G_0^{(J)}(z) = \prod_{j=0}^{J-1} G_0(z^{2^j}).$$

همانطور که می‌توان این موضوع را با استفاده از برابری‌های چند مرتبه‌ای استاندارد نشان داد [۴۳، ۴۴]. حال در نظر بگیرید که چه اتفاقی می‌افتد اگر یک سیگنال تک‌جمله‌ای با درجه‌ی کوچک‌تر از N همان‌طور که در (۲۷) ارائه شده است وارد فیلتر بانک‌های تحلیلی سپس این سیگنال در تمام فیلترهای بالاگذر ناپدید می‌شود و بنابراین تنها توسط شاخه‌ی پایین‌گذر بازتولید می‌شود. به عبارت دیگر، ترکیبات خطی $g_0^{(J)}[n]$ و انتقال‌های خطی 2^J آن می‌توانند چندجمله‌ای‌ها را تا درجه $N - 1$ بازتولید کنند، یا مشابه با (۲۹)، می‌توانیم بنویسیم

$$(31) \quad n^m = \sum_l \alpha_{jl}^{(m)} \cdot g_0^{(J)}[N - 2^J l].$$

بنابراین دیدیم که چندجمله‌ای‌های زمان گسسته تا درجه $N - 1$ در فضای تقریب «درشت» فیلتر بانک‌های زمان گسسته قرار دارند هنگامی که فیلتر پایین‌گذر N صفر در $z = -1$ دارد.

۵. چندجمله‌ای‌ها و موجک‌های زمان پیوسته

همان‌طور که مشخص است، یک پیوند قوی بین فیلتر بانک‌های تکراری و موجک‌ها وجود دارد. به عنوان مثال، فیلتر بانک‌ها می‌توانند برای تولید پایه‌های موجک استفاده شوند [۹] و فیلتر بانک‌ها می‌توانند برای محاسبه‌ی سری موجک‌ها به‌کار روند [۲۳]. بنابراین جای تعجب نیست که ویژگی‌هایی که در زمان گسسته در رابطه با نمایش چندجمله‌ای دیده شد به زمان پیوسته نیز منتقل شوند. در حالی که این ویژگی‌ها مستقیماً با خواص گشتاور موجک‌ها مرتبط هستند و در نتیجه به‌صورت عمومی حفظ می‌شوند، ما آن‌ها را در زمینه‌ی موجک‌های تولید شده از فیلتر بانک‌های پاسخ ضربه متناهی متعامد (FIR) بررسی می‌کنیم. مجدداً فرض کنید که فیلتر پایین‌گذر دارای N صفر در $\omega = \pi$ است، و بنابراین، بالاگذر دارای N صفر در $\omega = 0$ است. از رابطه‌ی دو مقیاسی تابع مقیاس و موجک، به این نتیجه می‌رسیم که تبدیل فوریه‌ی موجک را می‌توان به‌صورت زیر نوشت:

$$(32) \quad \Psi(\omega) = \frac{1}{\sqrt{2}} G_1(l^{j\omega/2}) \cdot \phi\left(\frac{\omega}{2}\right).$$

که در آن $\Phi(\omega)$ تبدیل فوریه تابع مقیاس است. از آنجایی که $\Phi(0) = 1$ ، از (۳۲) می‌بینیم که $\Psi(\omega)$ دقیقاً N صفر در $\omega = 0$ ، یا N گشتاور صفر با توجه به قضیه گشتاور تبدیل فوریه دارد. بنابراین، حاصلضرب داخلی بین هر موجک (مقیاس و انتقال دلخواه) با چندجمله‌ای تا درجه $N - 1$ ، صفر خواهد بود. بنابراین، چندجمله‌ای‌ها را می‌توان به تنهایی به‌صورت ترکیب خطی توابع مقیاس نشان داد، یا مشابه با (۳۱)، ضرایب $\gamma_l^{(m)}$ وجود دارد به‌طوری‌که

$$(33) \quad t^m = \sum_l \gamma_l^{(m)} \varphi(t - l).$$

یک مثال، با استفاده از تابع مقیاس دابشیز بر اساس یک فیلتر چهار ضربه‌ای [۹]، در شکل ۴ در یک بازه‌ی زمانی متناهی نشان داده شده است. تابع خطی تا حد مرزها به‌طور کامل بسط داده شده است.

۶. ناپیوستگی‌ها در فیلتر بانک و نمایش موجک‌ها

اگر یک سیگنال در نقطه‌ای مانند T ناپیوسته باشد چه اتفاقی می‌افتد؟ می‌دانیم که سری‌های فوریه ناپیوستگی‌ها را دوست ندارند، زیرا همگرایی یکنواخت را از بین می‌برند. موجک‌ها از نظر ناپیوستگی دو ویژگی مطلوب دارند. اول این‌که آن‌ها همچنان که ما به مقیاس‌های ظریف و ظریف‌تر می‌رویم، به‌صورت محلی روی ناپیوستگی تمرکز می‌کنند. این به دلیل رابطه‌ی مقیاس بندی موجک‌ها است که در آن مجموعه تابع $\Psi_{m,n}(t)$ بدین صورت تعریف می‌شوند

$$(34) \quad \psi_{m,n}(t) = 2^{m/2} \psi(2^{-m}t - n) \quad m, n \in \mathbb{Z}$$

که در آن $m \rightarrow -\infty$ با جزئیات دقیق مطابقت دارد. بنابراین، با منفی شدن m ، موجک “تیزتر” می‌شود. اگر ناپیوستگی جدا شده باشد و محیط اطراف هموار باشد، همه‌ی ضرب‌های داخلی موجک به جز مواردی که در ناپیوستگی قرار دارند صفر خواهند بود و در اطراف ناپیوستگی، $L - 1$ ضرب داخلی مورد نظر، زمانی که موجک دارای تکیه‌گاه به طول L است غیر صفر خواهد بود. دوم این‌که فاصله بزرگ بین مقیاس‌های ضرب داخلی موجک‌های ناصفر، ناپیوستگی را مشخص می‌کند. این یک مشخصه‌ی شناخته شده تبدیل موجک پیوسته [۱۰، ۲۵] است و برای سری موجک متعامد ثرمال نیز به خوبی صادق است. برای مثال در تابع دیراک $f(t) = \delta(t)$ ، ضرایب موجک، که توسط رابطه زیر تعیین می‌شوند

$$c_{m,n} = \langle \psi_{m,n}, f \rangle,$$

همان‌طور که در (۳۴) مشاهده می‌شود بروی n های نزدیک صفر دارای نمو به‌صورت زیر هستند

$$(35) \quad |c_{m,n}| \sim 2^{-m/2},$$

برای یک تابع هوی‌ساید (پله‌ای) رفتار آن به صورت زیر است

$$(۳۶) \quad |c_{m,n}| \sim 2^{m/2},$$

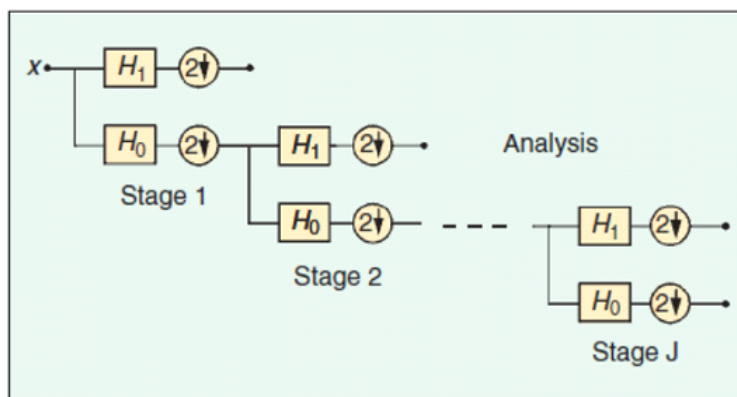
و به‌طور کلی، برای تکینگی مرتبه‌ی k رفتاری به صورت

$$(۳۷) \quad |c_{m,n}| \sim 2^{(k+\frac{1}{2})m},$$

برای $L - 1$ ضریب اطراف ناپیوستگی در مقیاس m خواهد داشت. تکینگی مرتبه k زمانی ظاهر می‌شود که k مشتق کراندار وجود داشته باشد، و مشتق $(k + 1)$ ام بیکران باشد. برای تابع پله‌ای $k = 0$ است، یک تابع پیوسته با یک نقطه‌ی مشتق‌ناپذیر دارای $k = 1$ است.

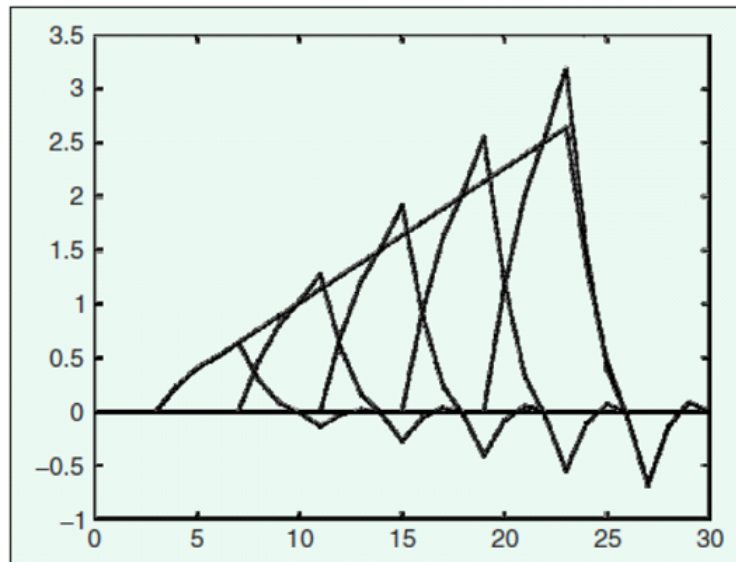
اکنون سیگنالی را در نظر بگیرید که عمدتاً هموار است، به جز در نقاط مجزا که دارای انواع مختلفی از تکینگی است. واضح است که انرژی ضرایب موجک در اطراف نقاط تکینگی متمرکز خواهد شد، با رفتاری که در (۳۷) نشان داده شده است. اگر سیگنال بین تکینگی‌ها چندجمله‌ای باشد، تمام دیگر ضرایب موجک صفر خواهند بود و تکه‌های چندجمله‌ای توسط توابع مقیاس "گرفته" می‌شوند. این در شکل ۵ نشان داده شده است.

بحث بالا بر سیگنال‌های زمان پیوسته متمرکز بود، پس در مورد یک سیگنال زمان گسسته چه اتفاقی می‌افتد؟ از نظر فنی، هیچ مفهومی از پیوستگی برای دنباله‌ها وجود ندارد. با این حال، از نظر کیفی، همان چیزی است که در زمان پیوسته اتفاق می‌افتد. علاوه بر این، دنباله می‌تواند قطعه‌ای چندجمله‌ای باشد که در این صورت یک برابری قطعه‌ای با بحث بالا داریم. در آن حالت، شاخه پایین‌گذر شاخه‌های (تقریب‌های) درشت را می‌گیرد، در حالی که کانال‌های بالاگذر و باندگذر به جز اطراف نقاط تکین صفر هستند. در این نقاط، $L - 1$ ضریب غیرصفر در هر مقیاس وجود خواهد داشت و اندازه آن‌ها مطابق با (۳۷) تعیین می‌شود. برای نتیجه‌گیری، بحث کردیم که چگونه توابع قطعه‌ای چندجمله‌ای‌های یا توابع تکه‌ای هموار دارای ضرایب بسط موجک هستند که بیشتر در اطراف نقاط تکین متمرکز شده‌اند، در حالی که توابع مقیاس یا فیلترهای پایین‌گذر چندجمله‌ای یا روندهای هموار را توصیف می‌کنند.



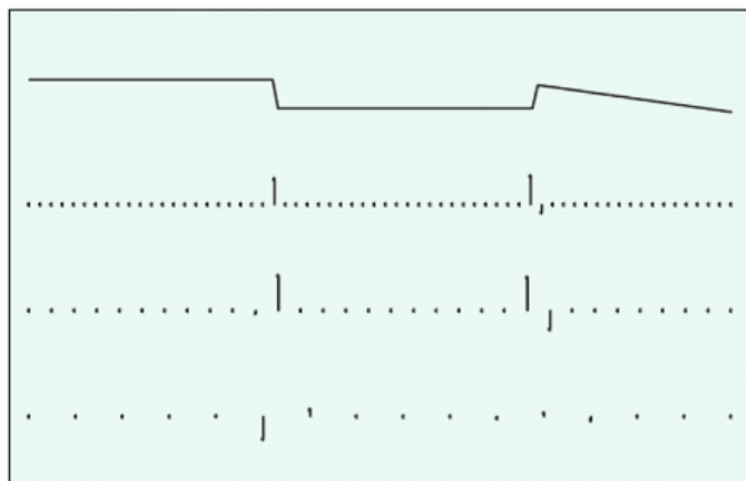
شکل ۳. بانک فیلتر پی‌درپی، شاخه‌ی پایین‌گذر به‌طور مکرر تقسیم می‌شود تا تبدیل موجک زمان گسسته به دست آید.

Figure 3: Iterated filter bank. The lowpass branch gets split repeatedly to get a discrete-time wavelet transform.



شکل ۴. بانک فیلتر پی‌درپی، شاخه‌ی پایین‌گذر به‌طور مکرر تقسیم می‌شود تا تبدیل موجک زمان گسسته به‌دست آید.

Figure 4: Reproduction of a linear function as a combination of scaling functions.



شکل ۵. تابع هموار تکه‌ای و بسط موجک آن. ضرایب موجک فقط در مجاورت ناپیوستگی با صفر متفاوت است، در حالی‌که رفتار درشت‌مقیاس تنها با ضرایب مقیاس‌بندی نشان داده می‌شود.

Figure 5: A piecewise smooth function and its wavelet expansion. The wavelet coefficients are different from zero only in the vicinity of the discontinuity, while the coarse behavior is represented by the scaling coefficients only.

۷. تقریب‌های خطی و غیرخطی

تا اینجا ساخت پایه‌ها به‌خصوص پایه موجک را بررسی کرده‌ایم. در آن‌جا، ما نگران نمایش دقیق هر سیگنال یا تابعی از یک فضای معین بودیم. اکنون به نمایش تقریب و مقایسه‌ی روش‌های خطی با غیرخطی می‌پردازیم. برای بحث عمیق در مورد چنین روش‌هایی، به‌ویژه در زمینه‌ی موجک، به توضیح عالی در [۲۴] اشاره می‌کنیم.

۱.۷. تقریب خطی. یک فضای V و یک پایه‌ی متعامد $\{g_n\}_{n \in \mathbb{N}}$ برای V در نظر بگیرید. بنابراین تابع $f \in V$ را می‌توان به‌صورت ترکیب خطی زیر نوشت

$$(۳۸) \quad f = \sum_{n \in \mathbb{N}} \langle g_n, f \rangle g_n.$$

بهترین تقریب خطی f (از نظر مربع خطا) در زیرفضای V ، که با \hat{f} نشان داده می‌شود، با تصویر متعامد f بر روی یک زیرفضای ثابت V به‌دست می‌آید (۱۱) را ببینید.

فرض کنید که یک زیرفضای W از بعد M است، که توسط اولین M بردارهای پایه یعنی

$$W = \text{span}\{g_0, g_1, g_2, \dots, g_{M-1}\}$$

بسط داده شده است. بنابراین f توسط

$$(۳۹) \quad \hat{f}_M = \sum_{n=0}^{M-1} \langle g_n, f \rangle g_n.$$

تقریب زده می‌شود و مربع خطای تقریب برابر است با

$$\hat{\varepsilon} = \|f - \hat{f}\|_V^2 = \sum_{n=M}^{\infty} |\langle g_n, f \rangle|^2.$$

از آنجایی‌که زیرفضای W مستقل از f ثابت می‌باشد، تقریب خطی است. این نتیجه از این واقعیت ناشی می‌شود که تقریب $f + g$ برابر با مجموع تقریب آن‌ها است. یک سؤال واضح این است که با توجه به دسته‌ای از اهداف مشخص چگونه بهترین پایه برای تقریب خطی (از نظر مربع خطا) را پیدا کنیم. به‌صورت دقیق‌تر، نتیجه‌ی کلاسیک را در تقریب خطی فرآیندهای برداری تصادفی مرور می‌کنیم. یک فرآیند برداری $X = [X_0, X_1, X_2, \dots, X_{N-1}]^T$ را در نظر بگیرید که در آن X یک فرآیند مستقل و با توزیع یکسان (iid) با میانگین صفر ($E[X_i] = 0$) و کوواریانس

$$(۴۰) \quad E[XX^T] = R_x,$$

که R_x از ابتدا معلوم است. از آنجایی‌که به تقریب حداقل مربعات علاقه‌مندیم، نیازی به مشخصه‌های آماری مرتبه بالاتر نخواهیم داشت. اکنون تقریب خطی را در پایه‌های متعامد نرمال $M < N$ ، $\{g_0, g_1, \dots, g_{M-1}\}$ در نظر می‌گیریم، که به‌صورت زیر است

$$(۴۱) \quad \hat{X}_M = \sum_{n=0}^{M-1} \langle g_n, X \rangle g_n.$$

مربع خطای مورد انتظار چقدر است یا $?\hat{\varepsilon}_M = E[\|X - \hat{X}_M\|^2]$

با بازنویسی $\|X - \hat{X}_M\|^2$ به‌صورت $\langle (X - \hat{X}_M) \cdot (X - \hat{X}_M) \rangle$ که پس از کمی دستکاری به‌صورت [۲۰]

$$(۴۲) \quad \|X - \hat{X}\|^2 = \sum_{m=M}^{N-1} g_m^T X X^T g_m,$$

در می‌آید، منجر به مربع خطای مورد انتظار می‌شود

$$(۴۳) \quad \hat{\varepsilon}_M = \sum_{m=M}^{N-1} g_m^T E[XX^T] g_m = \sum_{m=M}^{N-1} g_m^T R_x \cdot g_m.$$

اکنون سؤال این است: در تمام پایه‌های متعامد نرمال ممکن برای \mathbb{R}^N ، کدامیک (۴۳) را برای همه $M = 1, \dots, N-1$ کمینه می‌کند؟ با توجه به این که R_X نیمه‌معین مثبت است، پاسخ این است که بهترین پایه، از مجموعه بردارهای ویژه R_X با ترتیب کاهشی مقادیر ویژه متناظر به دست می‌آید. خطای تقریب به صورت

$$(۴۴) \quad \hat{\varepsilon}_M = \sum_{m=M}^{N-1} \lambda_m,$$

است که λ_m بزرگترین مقدار ویژه R_X است. شهود هندسی پشت این نتیجه این است که بردارهای ویژه، محورهای اصلی توزیع N -بعدی و بهترین زیرفضای M بعدی که بیشترین «انرژی» (یا نرم l_2 مورد انتظار) متناظر با محورهای اصلی با بزرگترین مقادیر ویژه را داراست (شکل ۶ را ببینید).

تبدیل متعامد داده شده توسط بردارهای ویژه R_X (که به صورت کاهش مقادیر ویژه مرتب شده) البته تبدیل معروف کارهونن لوه^۶ است که توسط گویال^۷ [۲۰] توضیح داده شده است. اما برای خلاصه کردن ویژگی اساسی آن مفید است: این بهترین پایه برای تقریب زیرفضای خطی یک فرآیند برداری است که توسط آمار مرتبه دوم مشخص شده است. دوباره فرض می‌شود که آمار مرتبه دوم از ابتدا مفهوم باشد، بنابراین زیرفضای آن نیز از ابتدا به صورت مطلوب انتخاب و فیکس می‌شود. (اگر ماتریس همبستگی باید غیر مستقیم تخمین زده شود، در این صورت وضعیت متفاوتی داریم که یک تقریب خطی نیست. برای مثالی از این نمونه [۲۲] را ببینید.) در مورد یک فرآیند بردار گاوسی مشترک، یک نتیجه‌ی قوی‌تر، یعنی بهترین تبدیل «فشرده‌سازی» (زمانی که ضرایب تبدیل کوانتیزه می‌شوند و آنتروپی کد می‌شوند) وجود دارد. برای بررسی این نتیجه به [۱۹، ۲۰، ۲۱] مراجعه کنید.

۲.۷. تقریب غیرخطی در پایه‌های متعامد. همان تنظیمات فوق را در نظر بگیرید، اما با یک قانون تقریب متفاوت. به جای (۳۸)، که در آن اولین ضرایب M در بسط متعامد استفاده می‌شود، بزرگترین ضرایب M را نگه می‌داریم. یعنی مجموعه شاخص I_M از بزرگترین M ضرب داخلی را تعریف می‌کنیم یا

$$(۴۵) \quad |\langle g_m, f \rangle| \geq |\langle g_n, f \rangle|, \quad m \in I_M, n \notin I_M.$$

سپس بهترین تقریب غیرخطی را به صورت زیر تعریف می‌کنیم:

$$(۴۶) \quad \hat{f}_M = \sum_{n \in I_M} \langle g_n, f \rangle g_n$$

که منجر به خطای تقریبی به صورت

$$\tilde{\varepsilon}_M = \|f - \hat{f}_M\|^2 = \sum_{n \in I_M} |\langle g_n, f \rangle|^2$$

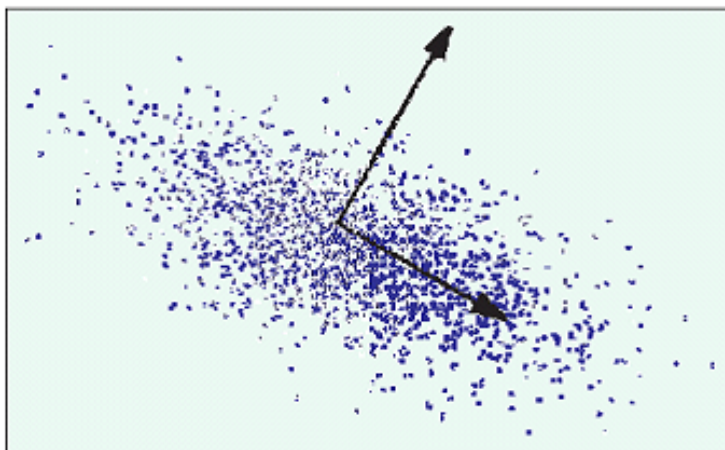
می‌شود. به وضوح

$$(۴۷) \quad \tilde{\varepsilon}_M \leq \hat{\varepsilon}_M.$$

(این تقریب را می‌توان تقریب خطی تطبیقی یا تقریب زیرفضای تطبیقی نامید. با این حال، اصطلاح رایج برای آن تقریب غیرخطی است. طرح‌های غیرخطی کلی‌تری را نیز می‌توان در نظر گرفت، اما از حوصله‌ی این مقاله خارج است.)

^۶Karhunen-Loève (KLT) ^۷Goyal

اجازه دهید نشان دهیم که اصلاح ساده‌ی انتخاب بزرگترین به‌جای اولین M ضرب داخلی، در واقع یک طرح غیرخطی ایجاد می‌کند. دو تابع f و h را در نظر بگیرید. به‌طور کلی، بزرگ‌ترین ضرایب M با مجموعه یکسانی برای f و h مطابقت نخواهد داشت و بنابراین تقریب تقریبی $f + h$ مطمئناً مجموع تقریب‌های f و h نیست که معمولاً بیشتر از M جمله است.



شکل ۶. توزیع دو بعدی با نمودار پراکندگی نشان داده شده است. محورهای اصلی نشان داده شده‌اند و بهترین تقریب یک بعدی، محور مربوط به بزرگترین گسترش است.

Figure 6: dimensional distribution indicated by a scatter plot. The principal axes are indicated, and the best one-dimensional approximation is the axis corresponding to the biggest spread.

۸. تقریب خطی در مقابل غیرخطی

حال، سؤال این است که آیا (۴۵) منجر به تفاوت‌های زیادی در کیفیت تقریب می‌شود؟ آیا پایه‌هایی وجود دارد که تفاوت آن‌ها بیشتر از سایر پایه‌ها باشد؟ مثال یک تابع ثابت تکه‌ای با ناپوستگی را در نظر بگیرید. با این‌که ساده است اما تفاوت‌های اساسی بین تقریب خطی و غیرخطی و همچنین بین پایه‌های فوریه و موجک را نشان می‌دهد. برای مشخص شدن این موضوع، بازه $[0, 1]$ را در نظر بگیرید و از سری فوریه یا سری موجک برای نشان دادن تابعی استفاده کنید که ثابت است به جز جهشی در یک نقطه‌ی تصادفی، به‌طور یکنواخت روی $[0, 1]$ توزیع شده است. ابتدا سری فوریه را در نظر بگیرید. به دلیل ناپوستگی، ضرایب سری فوریه به آرامی کاهش می‌یابد

$$(۴۸) \quad |\alpha_n|, |\beta_n| \sim \frac{1}{n}.$$

که α_n و β_n ضرایب کسینوس و سینوسی در فرکانس $2\pi n$ هستند. تقریب خطی توسط اولین M - جمله (یا اولین $M/2$ سینوس و کسینوس) منجر به خطایی از مرتبه‌ی

$$(۴۹) \quad \hat{\epsilon}_M^{(f)} \sim \sum_{N=M}^{\infty} \frac{1}{n^2} \sim \frac{1}{M}$$

می‌شود. انتخاب بزرگترین ضرایب، مرتبه‌ی تقریب را به دلایل زیر تغییر نمی‌دهد. در حالی‌که برخی از ضرایب α_n و β_n در جملات اولیه‌ی بسط، کوچک خواهند بود (سینوس و کسینوس در صورت ضرایب سری فوریه گاهی اوقات می‌توانند کوچک باشد)، بیشتر M ضرایب بزرگ، هنوز در اولین kM جمله جمع می‌شوند. که در آن k یک عدد صحیح کوچک است. بنابراین برای تقریب اول،

مجموعه‌ی بزرگ‌ترین M ضرایب دارای همپوشانی زیادی با مجموعه M ضرایب اول است و بنابراین خطای تقریب غیرخطی به صورت زیر است

$$(۵۰) \quad \tilde{\varepsilon} \sim \frac{1}{M}.$$

اکنون بسط موجک را با استفاده از موجک‌ها تعریف شده به صورت زیر در نظر بگیرید

$$(۵۱) \quad \Psi(t) = \begin{cases} 1 & 0 \leq t \leq 1/2 \\ -1 & 1/2 \leq t \leq 1 \\ 0 & \text{غیره} \end{cases}$$

و موجک‌ها به صورت (۳۴) با $m = 0, -1, -2, \dots$ تعریف می‌شوند، در حالی‌که چنین موجکی به نظر می‌رسد به طور ناعادلانه نزدیک به تابعی که باید نمایش داده شود، است. باید اشاره کنیم که هنوز تعداد نامتناهی ضرایب موجک غیرصفر وجود دارد، مگر اینکه t یک باینری منطقی باشد. علاوه بر این، رفتاری که می‌خواهیم نشان دهیم همچنان به صورت کلی باقی می‌ماند. ابتدا به یاد آورید که یک ضریب موجک غیرصفر در هر مقیاس m در اطراف ناپیوستگی وجود خواهد داشت و این ضریب دارای اندازه‌ی زیر است

$$(۵۲) \quad |c_{m,n}| \sim 2^{m/2}.$$

که در آن $m \rightarrow -\infty$. هنگام انجام تقریب خطی، ابتدا مقیاس‌های درشت‌تر (جایی که ضرایب بزرگ‌تر قرار دارند) را در نظر می‌گیریم و سپس به سمت مقیاس‌های ظریف‌تر می‌رویم. در حالی‌که همه‌ی ضرایب به جز یک صفر هستند، ما مجاز به تطبیق نیستیم. اگر $M = 2^J$ می‌توان $J-1$ مقیاس (از 0 تا $-J+2$) را در نظر گرفت که در آن نقطه تنها ضریب غیرصفر از مرتبه‌ی $2^{-1/2}$ است. خطای درجه دوم باقیمانده از مرتبه‌ی 2^{-J} است و بنابراین

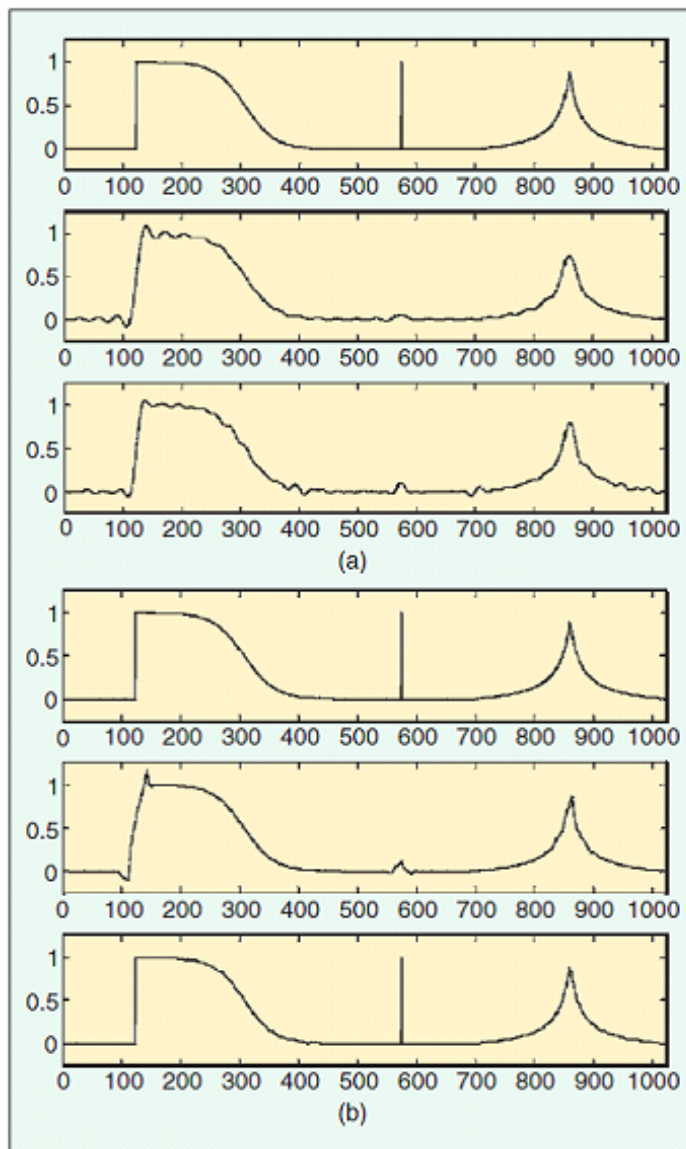
$$(۵۳) \quad \hat{\varepsilon}_M^{(W)} \sim \frac{1}{M}$$

که درست مانند فوریه است. با این حال، اگر به طرح تقریب غیرخطی برویم، همه چیز به طور چشمگیری تغییر می‌کند. انتخاب بزرگترین ضرایب M به ما اجازه می‌دهد تا از مقیاس‌های M عبور کنیم (زیرا فقط یک ضریب غیر صفر در هر مقیاس وجود دارد) که در آن نقطه ضریب از مرتبه $2^{-M/2}$ است. خطای درجه دوم پس از آن به صورت زیر است

$$(۵۴) \quad \tilde{l}_M^{(W)} \sim 2^{-M}.$$

بنابراین از میرایی $\frac{1}{M}$ در فوریه (خطی و غیرخطی) یا تقریب خطی موجک، به میرایی نمایی با تقریب غیرخطی در پایه‌های موجک رفته‌ایم.

آنچه که در بالا دیدیم در موارد کلی‌تر هم صادق است. اگر تابعی به صورت تکه‌ای هموار و با ناپیوستگی‌های مجزا باشد، تقریب فوریه به دلیل ناپیوستگی‌ها ضعیف است. در مورد موجک با موجکی که دارای گشتاورهای صفر کافی است به طوری که ضرب داخلی در ناحیه‌ی هموار کوچک هستند، ضرایب موجک کمی برای ثبت ناپیوستگی‌ها کافی است، و تقریب غیرخطی، از طرح‌های خطی چندین مرتبه‌ی بزرگ‌تر بهتر است. شکل ۷ یک مثال را نشان می‌دهد، جایی‌که می‌توان دید تقریب غیرخطی در مورد موجک بسیار برتر است، در حالی‌که تفاوت کمی در مورد فوریه ایجاد می‌کند. همچنین توجه داشته باشید که با تقریب خطی، تفاوت کمی بین فوریه و موجک وجود دارد. برای بررسی برخی از نتایج تقریب قوی برای بسط موجک توابع هموار و تکه‌ای هموار، [۱۲، ۲۴] را ببینید.



شکل ۷. فوریه در مقابل پایه موجک و تقریب خطی در مقابل غیرخطی. سیگنال زمان گسسته دارای طولی برابر ۱۰۲۴ است و تعداد $M = 64$ ضریب تعیین می‌شوند. (الف) مورد فوریه. اصل در بالا، تقریب خطی در وسط، و تقریب غیرخطی در پایین. MSE به ترتیب ۲/۷ و ۲/۴ است. (ب) حالت موجک، با شش سطح تجزیه و یک موجک دارای سه گشتاور صفر. اصل در بالا، تقریب خطی در وسط، و تقریب غیرخطی در پایین. MSE به ترتیب ۳/۵ و ۰/۱ است.

Figure 7: Fourier versus wavelet bases and linear versus nonlinear approximation. The signal is discrete time of length 1024, and $M = 64$ coefficients are retained. (a) Fourier case. Original on top, linear approximation in the middle, and nonlinear approximation at the bottom. The MSE is 2.7 and 2.4 respectively. (b) Wavelet case, with six levels of decomposition and a wavelet having three zero moments. Original on top, linear approximation in the middle, and nonlinear approximation at the bottom. The MSE is 3.5 and 0.01, respectively.

۹. تقریب و فشرده‌سازی غیرخطی

ما تاکنون عناصر M را از یک پایه، یا یک مجموعه ثابت (معمولاً اولین M) یا یک مجموعه‌ی تطبیقی (مطابق با بزرگترین تصویر) در نظر گرفته‌ایم. در فشرده‌سازی، ابتدا باید مجموعه ضرایب را توصیف کنیم که وقتی این مجموعه ثابت است، ارزش آن صفر است (تقریباً خطی) اما وقتی تطبیقی است (تقریباً غیرخطی) ارزش غیرمعمول دارد. در آن صورت، $\left(\frac{N}{M}\right)$ زیرمجموعه ممکن وجود دارد و سرعت توصیف زیرمجموعه برابر با آنتروپی توزیع زیرمجموعه‌ها یا حداکثر برابر با

$$(55) \quad \log_2 \left(\frac{N}{M} \right)$$

است. علاوه بر این لازم است که ضرب‌های داخلی مختلف را کوانتیزه و کدگذاری آنتروپی کنیم. به دلایل پیچیدگی، معمولاً به جای کوانتیزاسیون برداری و کدگذاری آنتروپی برداری، فقط کدگذاری اسکالر و آنتروپی ضرایب جداگانه در نظر گرفته می‌شود. با استفاده از نتایج معمول کوانتیزاسیون با نرخ بالا [۱۹، ۲۰]، صرف R بیت برای یک ضریب خاص منجر به هزینه نرخ پیچیدگی به شکل

$$(56) \quad D(R) = C_d 2^{-2R},$$

می‌شود که در آن C_d به توزیع ضریب بستگی دارد (به عنوان مثال σ^2 برای توزیع یکنواخت با واریانس σ^2 یا $(\pi e/6)\sigma^2$ برای متغیر گاوسی تصادفی با واریانس σ^2). با استفاده از نرخ‌های تقریب M جمله‌ای مثلاً در (۴۹)، (۵۰) (یا (۵۳)، (۵۴)) همراه با پیچیدگی توصیف مجموعه‌ی (۵۵) و نرخ پیچیدگی برای ضرایب (۵۶)، می‌توان کران‌های بالای قابل دستیابی را برای هزینه نرخ پیچیدگی طرح‌های تقریب خطی و غیرخطی استخراج کرد.

۱۰. مثالی از تقریب و فشرده‌سازی خطی در مقابل غیرخطی

مثالی که تفاوت اساسی بین تقریب خطی و غیرخطی در فشرده‌سازی را به خوبی مشخص می‌کند به این شرح است. یک پردازش برداری تصادفی به اندازه N را در نظر بگیرید که در آن هر تحقق فقط یک مقدار غیر صفر در یک مکان تصادفی k دارد، بقیه صفر هستند و یا

$$(57) \quad X[n] = \alpha \cdot \delta[n - k], \quad n \in \{0, \dots, N - 1\}$$

که در آن α ، $N(0, \sigma^2)$ گاوسی است و k یک عدد صحیح است که به‌طور یکنواخت روی $\{0, \dots, N - 1\}$ توزیع شده است. به عبارت دیگر، برای هر تحقق، مکانی را به‌طور یکنواخت بین 0 و $N - 1$ انتخاب می‌کنیم و در آن مکان، یک متغیر تصادفی گاوسی با میانگین صفر و کوواریانس σ^2 قرار می‌دهیم، همان‌طور که به صورت شماتیک در شکل ۸ نشان داده شده است. ابتدا مسئله‌ی تقریب خطی را در نظر بگیرید. ماتریس اتوکوواریانس برابر است با

$$(58) \quad R_x = \frac{\sigma^2}{N} \cdot I$$

و بنابراین با استفاده از رویکرد KLT، پایه‌ی استاندارد در این حالت بهترین پایه (نه تنها پایه) برای تقریب زیرفضای خطی است. بنابراین، R بیت به‌طور مساوی در بین موقعیت‌های N توزیع می‌شوند و رفتار کلی نرخ پیچیدگی به صورت زیر خواهد بود:

$$(59) \quad D_L(R) = C \cdot \sigma^2 \cdot 2^{-2R/N}.$$

توجه داشته باشید که این همان چیزی است که رویکرد خطی KLT به ما می‌گوید که انجام دهیم، حتی اگر به وضوح کمتر از حد بهینه است. یک رویکرد بهتر، حداقل اگر R به اندازه‌ی کافی بزرگ باشد، صرف $\log_2 N$ بیت به نقطه‌ی در مکانی است که در

تحقق داده شده فعال است و بقیه $R - \log_2 N$ بیت مانده را برای توصیف متغیر گاوسی تصادفی صرف کنید. این منجر به رابطه زیر می‌شود

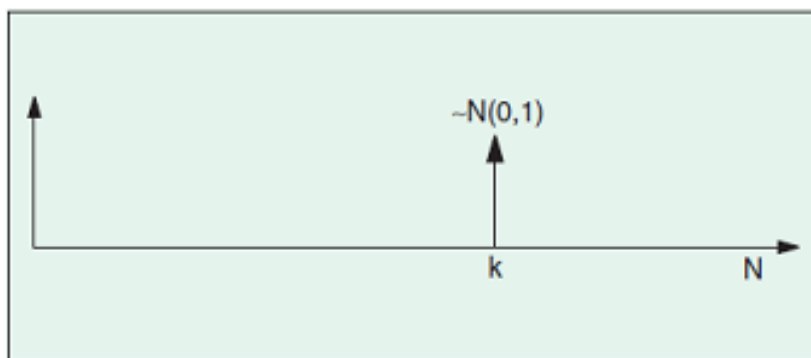
$$(۶۰) \quad D_{NL}(R) = C \cdot \sigma^2 \cdot 2^{-(R - \log_2 N)},$$

که برای $R \gg \log_2 N$ به‌وضوح بهتر از (۵۹) است. برای نرخ‌های پایین ($R \sim \log_2 N$ و کمتر از آن)، وضعیت دشوارتر است. برای تحلیل کامل به ویدمن [۴۵] مراجعه کنید. نکته مثال بالا این است که نظریه‌ی خطی زمانی گمراه‌کننده است که فرآیندها از گاوسی مشترک دور باشند. KLT منجر به همبستگی می‌شود، که تنها در مورد گاوسی به معنای استقلال است. این در حالی است که معلوم می‌شود بسیاری از سیگنال‌های مورد نیاز برای فشرده‌سازی گاوسی نیستند.

۱۱. کدگذارهای واقعی، تقریب غیرخطی انجام می‌دهند

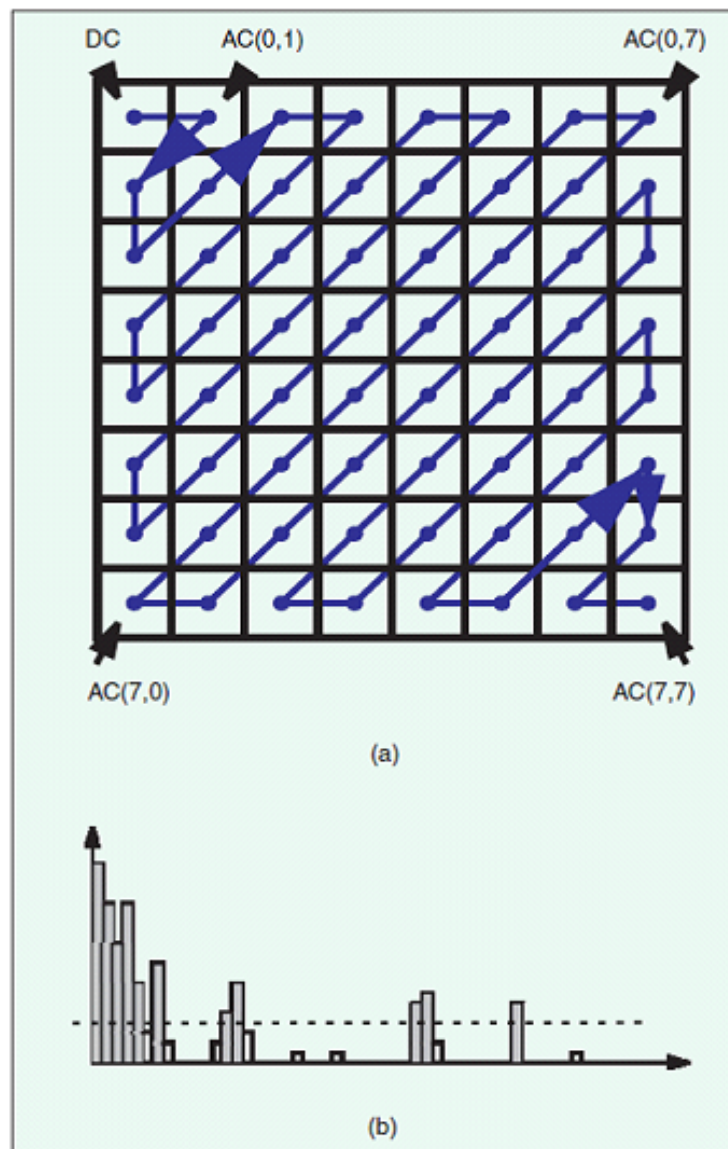
بیاید ببینیم کدگذاران “واقعی” چه می‌کنند. در واقع، مدت‌ها قبل از موجک‌ها و نظریه تقریب غیرخطی، مهندسان فشرده‌سازی، روش‌هایی را برای تطبیق نظریه مبتنی بر KLT خطی با سیگنال‌های دنیای واقعی توسعه دادند. دو جزء اصلی استفاده می‌شود، یعنی محلی بودن و سازگاری. ابتدا، سیگنال‌ها به زیر سیگنال‌های محلی تقسیم می‌شوند (مثلاً با استفاده از بلوک کردن مانند کدگذاری تبدیل مبتنی بر بلوک)، و این زیر سیگنال‌ها به‌صورت جداگانه تبدیل می‌شوند. سپس، این تبدیل‌های «محلی» بسته به ویژگی‌هایشان، به‌صورت تطبیقی کدگذاری می‌شوند. این یادآور مثال قبلی ما است: اگر فقط چند ضریب بزرگ اما در مکان‌های ناشناخته ظاهر شوند باید هزینه خطی سازی آن‌ها را داد. بدون وارد شدن به جزئیات زیاد، شکل ۹ یک مجموعه معمولی از ضرایب تبدیل کسینوس گسسته (DCT) را همان‌طور که در JPEG استفاده می‌شود و کوانتیزاسیون زیردنباله‌ای، طول اجرا و کدگذاری آنتروپی استفاده می‌شود. با این روش، زمانی که ضرایب بزرگ کمی وجود داشته باشد، می‌توان آن‌ها را با بیت‌های کمی نشان داد و ما در واقع یک طرح تقریب غیرخطی بر پایه‌های فوریه مانند محلی داریم. شایان ذکر است که هر طرح تطبیقی، باعث غیرخطی شدن می‌شود. به‌عنوان مثال، تخصیص بیت تطبیقی بین بلوک‌های تبدیل (بیت‌های بیشتر به بلوک‌هایی با فعالیت بالا، کمتر به بلوک‌های کم فعالیت) در حال حاضر یک رویکرد غیرخطی است. همچنین، بهترین پایه‌های تطبیقی [۶، ۳۴]، تخصیص‌های مبتنی بر برنامه‌ریزی پویا [۲۹، ۳۴]، بخش‌بندی تطبیقی [۳۱، ۳۲] و پیگیری انطباقی [۲۶] همگی نمونه‌هایی از طرح‌های غیرخطی مورد استفاده در فشرده‌سازی هستند.

شایان ذکر است که بررسی رفتار نرخ-پیچیدگی رمزگذارهای “واقعی” موضوعی دشوار است، هم به دلیل غیرخطی بودن تقریب و کمی‌سازی و هم به دلیل اینکه تعامل تقریب، کمی‌سازی و کدگذاری آنتروپی غیربدهی هستند.



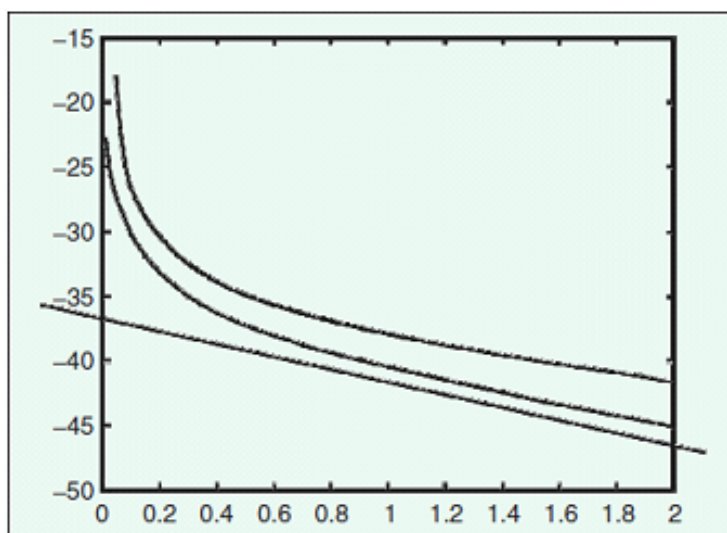
شکل ۸. فرآیند نقطه‌یابی، با انتخاب یکنواخت یک مکان واحد که در آن یک متغیر تصادفی گاوسی قرار می‌گیرد.

Figure 8: Pointing process, with uniform choice of a single location where a Gaussian random variable is placed.



شکل ۹. نمونه‌ای از ۶۴ ضریب DCT از یک بلوک 8×8 و پیمایش زیگزاگ. (الف) پیمایش زیگزاگ بلوک 8×8 ضرایب DCT برای به دست آوردن یک آرایه یک بعدی از ضرایب. (ب) نمونه‌ای از آرایه ۶۴ ضرایب (فقط اندازه) با آستانه T (کوانتایزر منطقه مرده) و کمی‌سازی ضرایب بالای T . پس از آن، کدگذاری طول اجرا برای نمایه‌سازی ضرایب بزرگ و کدگذاری آنتروپی برای نشان دادن مقدار آن‌ها استفاده می‌شود.

Figure 9: Example of 64 DCT coefficients from an 8×8 block and zig-zag scanning. (a) Zig-zag scanning of the 8×8 block of DCT coefficients to obtain a one-dimensional array of coefficients. (b) Example of an array of 64 coefficients (magnitude only) with a threshold T (deadzone quantizer) and quantization of the coefficients above T . Subsequently, run length coding is used to index the large coefficients, and entropy coding to represent their value.



شکل ۱۰. مربع خطا (در مقیاس ورود به سیستم) در مقابل نرخ بیت برای JPEG (منحنی بالایی) و SPIHT [۳۵] (منحنی پایینی). در نرخ‌های بالا، معمولی آشکار است، در حالی که در نرخ‌های بسیار پایین، یک فروپاشی تندتر معمولی است.

Figure 10: Squared error (on log scale) versus bit rate for JPEG (upper curve) and SPIHT [35] (lower curve). At high rates, the typical -6dB/bit is apparent, while at very low rates, a steeper decay is typical.

به‌ویژه، رفتار نرخ پایین گمراه کننده است. شکل ۱۰ لگاریتم مربع خطای JPEG و یک کدگذاری پایه موجک را نشان می‌دهد، اسپیت^۸ [۳۵]. در حالی که رفتار نرخ بالا همان‌طور است که انتظار می‌رود (یعنی ۶ دسی‌بل در هر بیت در مقیاس \log)، رفتار نرخ پایین متفاوت است، زیرا مربع خطا با نرخ‌های خیلی پایین بسیار سریعتر از بین می‌رود. یک توضیح شهودی این است که در نرخ‌های بسیار پایین، تخصیص نرخ کلاسیک برقرار نیست، زیرا بسیاری از ضرایب تبدیل هیچ نخ‌ری را دریافت نمی‌کنند. بنابراین تعداد اندک ضرایب باقیمانده دقت بیشتری دریافت می‌کنند که منجر به کاهش سریع‌تر خطا می‌شود. مالات^۹ و همکاران [۲۴] با استفاده از یک تحلیل متفاوت، رفتار $1/R$ را در نرخ‌های پایین برای کدهای تبدیل نشان داده‌اند.

۱۲. فشرده‌سازی سیگنال‌های چندجمله‌ای تکه‌ای

اجازه دهید به سیگنال‌های یک بعدی تکه‌ای هموار برگردیم. هنگامی که تقریب غیرخطی مجاز باشد، موجک‌ها برای تقریب چنین سیگنال‌هایی مناسب هستند. برای مطالعه رفتار فشرده‌سازی، حالت ساده‌تر چندجمله‌ای‌های تکه‌ای را با ناپوستگی در نظر بگیرید. برای سهولت کار، اجازه دهید دوباره به سیگنالی که قبلاً برای مطالعه تقریب غیرخطی استفاده کردیم نگاه کنیم، اما این بار همراه با کوانتیزاسیون و تخصیص بیت است. تحلیل ساده‌ی رفتار پیچیدگی نرخ تقریبی یک تابع پله‌ای به این شرح است. ضرایب به اندازه‌ی $2^{m/2}$ کاهش می‌یابد، بنابراین تعداد مقیاس‌های J درگیر، اگر از کوانتیزر اندازه Δ استفاده شود، از مرتبه $\log_2(1/\Delta)$ است. تعداد بیت‌ها در هر ضریب نیز از مرتبه $\log_2(1/\Delta)$ است، بنابراین نرخ R از مرتبه‌ی

$$(۶۱) \quad R \sim \log_2^2(1/\Delta) \sim J^2$$

^۸SPIHT ^۹Mallat

است. پیچیدگی یا مربع خطا متناسب با Δ^2 (برای هر ضریب) برابر تعداد مقیاس‌ها یا با استفاده از $\Delta = 2^{-J}$ داریم

$$(۶۲) \quad D \sim J \cdot 2^{-J}.$$

با به‌کارگیری (۶۱) خواهیم داشت

$$(۶۳) \quad D_W(R) = C_1 \cdot \sqrt{R} \cdot 2^{C_1 \sqrt{R}}.$$

برای رفتار نرخ پیچیدگی یک طرح موجک، توجه داشته باشید که هزینه نمایه‌سازی مکان را نادیده گرفتیم. به نظر می‌رسد که این هزینه بسیار کوچک است (مرتبه J)، زیرا ضرایب همه حول ناپیوستگی جمع شده‌اند. برای سیگنال‌های کلی‌تر، پراندونی [۳۱] نشان داده است که تقریب غیرخطی چندجمله‌ای‌های تکه‌ای با حداکثر درجه‌ی N ، با استفاده از فشرده‌سازی موجک که در آن موجک $N + 1$ گشتاور صفر دارد، منجر به تساوی زیر می‌شود

$$(۶۴) \quad D_W(R) = C_W''(1 + C_W' \sqrt{R}) \cdot 2^{-\sqrt{C_W R}}$$

که در آن ثابت‌ها به N و تعداد ناپیوستگی‌ها بستگی دارند. خبر ناراحت‌کننده، عبارت R در توان است که بسیار دور از رفتار نرخ بالای مورد انتظار است. چنین رفتاری به‌طور کلی برای توابع تکه‌ای هموار توسط کوهن و همکاران نشان داده شده است [۵]. (توجه داشته باشید که این یک کلاس بسیار گسترده‌تر از چندجمله‌ای‌های تکه‌ای است.) یک رویکرد مستقیم به فشرده‌سازی چندجمله‌ای‌های تکه‌ای، بر اساس یک پیش‌گویی که به ما می‌گوید ناپیوستگی‌ها کجا هستند، منجر به تساوی

$$(۶۵) \quad D_P(R) = C'_P$$

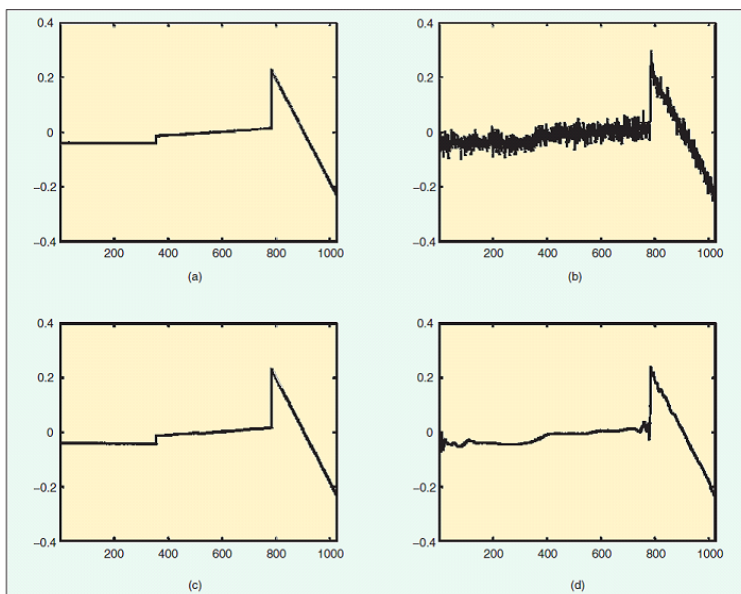
می‌شود و چنین رفتاری با استفاده از برنامه‌نویسی پویا قابل دستیابی است [۳۱]. واضح است که (۶۵) در نرخ‌های بالا بسیار بهتر از (۶۴) است. حال چرا فشرده‌سازی موجک به‌صورت غیربینه انجام می‌شود؟ دلیل شهودی ساده است: با فشرده‌سازی ضرایب در مقیاس‌ها به‌طور مستقل، دیگر از مزیت‌های وابستگی‌های بین مقیاسی استفاده نمی‌شود. چندین تلاش برای استفاده از چنین وابستگی‌هایی صورت گرفته است. به‌عنوان مثال، درختان صفر [۳۷] فقدان مقیاس‌های انرژی را پیش‌بینی می‌کنند و مدل‌های پنهان مارکوف [۸] وابستگی‌های بین مقیاس‌ها را نشان می‌دهند. در مورد چندجمله‌ای‌های تکه‌ای با ناپیوستگی، رفتار در مقیاس‌ها در واقع معین است و این می‌تواند با استفاده از الگوی موجک مدل شود [۱۶]. استفاده از الگو بر روی تجزیه موجک یک چندجمله‌ای تکه‌ای این امکان را فراهم می‌کند که در حالی‌که از تبدیل موجک استفاده می‌کند، به بهترین رفتار ممکن که در (۶۵) داده شده است، دست یابد. نمونه‌هایی از استفاده از الگو برای حذف نویز را می‌توان در شکل ۱۱ مشاهده کرد.

۱۳. مورد دوبعدی

با توجه به عملکرد خوب موجک‌ها برای توابع تکه‌ای هموار در یک بعد، می‌توان امیدوار بود که در دو بعد نیز نتایج خوبی حاصل شود که البته چنین نیست. در اصل، موجک‌ها در گرفتن تکنیکی‌های صفر بعدی خوب هستند، اما سیگنال‌های تکه‌ای هموار دوبعدی که شبیه تصاویر هستند، تکنیکی‌های یک بعدی دارند. یعنی نواحی صاف با لبه‌ها از هم جدا می‌شوند اما خود لبه‌ها معمولاً منحنی صاف هستند. به‌طور شهودی، موجک‌ها در جداسازی ناپیوستگی متعامد به لبه خوب هستند، اما صافی را در امتداد لبه نمی‌بینند. این‌را می‌توان با نگاه کردن به توان تقریبی یک موجک دو بعدی تأیید کرد. چنین مبنایی با حاصلضرب تانسور موجک‌های یک بعدی به‌دست می‌آید. تحلیلی مشابه آنچه در بخش تقریب خطی در مقابل غیرخطی انجام شد نشان می‌دهد که تقریب غیرخطی M جمله‌ای یک تابع ثابت تکه‌ای ساده با ناپیوستگی خطی، منجر به مربع خطای از مرتبه‌ی

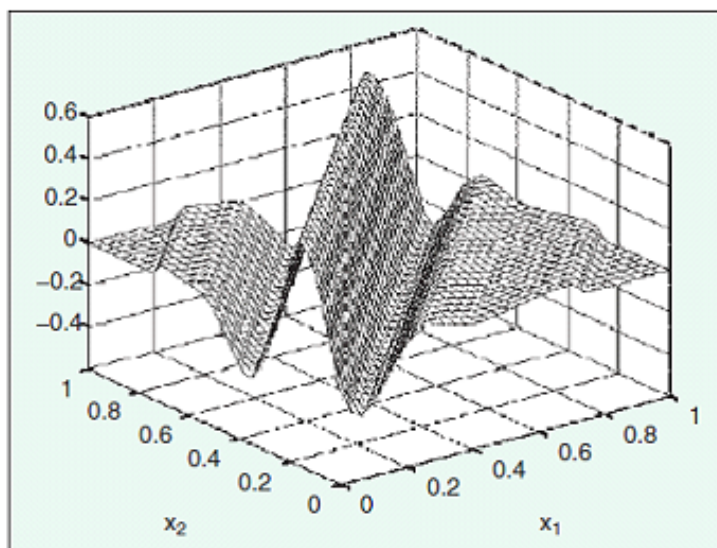
$$(۶۶) \quad \bar{\epsilon}_M \sim \frac{1}{M}$$

می‌شود این رفتار ناامیدکننده نشان می‌دهد که در ابعاد بالاتر به پایه‌های قوی‌تری نیاز است.



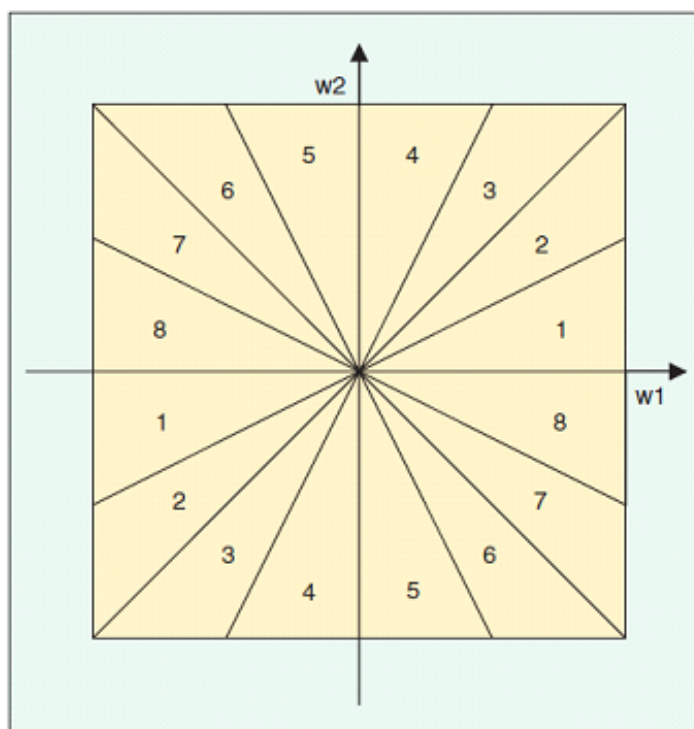
شکل ۱۱. نویز زدایی در حوزه موجک با استفاده از آستانه و الگو الف) اصل (ب) نسخه نویزدار (SNR = ۱۱٫۱ دسی‌بل). (ج) نویز زدایی با استفاده از الگوی اثر (SNR = ۳۱٫۴ دسی‌بل). (د) نویز زدایی با استفاده از آستانه موجک استاندارد (SNR = ۲۰٫۱ دسی‌بل).

Figure 11: Denoising in wavelet domain using thresholding and footprints. (a) Original. (b) Noisy version (SNR = 11.1 dB). (c) Denoising using footprints (SNR = 31.4 dB). (d) Denoising using standard wavelet thresholding (SNR = 20.1 dB).



شکل ۱۲. یک ریجالت نوعی، در زاویه θ با نمایه‌ای که توسط $\Psi_{a,b}(t)$ داده می‌شود.

Figure 12: Typical ridgelet, at angle θ with a profile given by $\Psi_{a,b}(t)$.



شکل ۱۳. تقسیم فرکانس یک فیلتر بانک جهت‌دار با ۸ باندها.

Figure 13: Frequency division of a directional filter bank with 8 bands.

۱۴. پایه‌های دو بعدی واقعی

همان‌طور که مثال موجک نشان می‌دهد، پایه‌های قابل تفکیک برای اشیاء دو بعدی “واقعی” مناسب نیستند. آنچه مورد نیاز است تبدیل‌ها و پایه‌هایی است که شامل نوعی “هندسه‌ی” واقعاً دو بعدی باشد. (مفهوم هندسه در زمینه‌ی ما رسمیت بخشیدن به آن آسان نیست، اما شهود این است که ابعاد مستقل نیستند و اشکال خاصی بیشتر از بقیه محتمل هستند) علاوه بر تبدیل فوریه و موجک دو بعدی که هر دو قابل تفکیک هستند، تبدیل رادون نقش کلیدی ایفا می‌کند. این تبدیل که در اوایل قرن بیستم مورد مطالعه قرار گرفت [۳۳]، چندین بار در زمینه‌های مختلف از نجوم تا تصویربرداری پزشکی دوباره کشف شد (برای یک نمای کلی عالی به [۱۱] مراجعه کنید). تبدیل رادون با گرفتن انتگرال‌های خط در زاویه θ و مکان t ,

$$(۶۷) \quad RA_f(\theta, t) = \iint f(x, y) \delta(x \cos \theta + y \sin \theta - t) dx dy$$

تابع $f(x, y)$ را به $RA_f(\theta, t)$ نگاشت می‌کند. بینش کلیدی برای ساخت پایه‌های جهت‌دار از تبدیل رادون توسط کاندس^{۱۰} و دونا هو [۲، ۳] با تبدیل ریجالت ارائه شد. ایده این است که با استفاده از تبدیل رادون، یک تکنیکی یک بعدی مانند یک لبه‌ی مستقیم را به یک تکنیکی نقطه‌ای ترسیم کنیم. سپس، تبدیل موجک را می‌توان برای رسیدگی به تکنیکی نقطه‌ای استفاده کرد. برای توسعه شهود، یک تابع دو بعدی مانند تابع پله‌ای را در نظر بگیرید که دارای مقدار ۱- در سمت چپ خط $\delta(x \cos \theta_o + y \sin \theta_o - t_o)_x$ و ۱ در سمت راست است. این تابع، لبه “بی‌نهایت” در زاویه θ_o و مکان t_o است.

به دلیل گستردگی نامتناهی، ما نمی‌توانیم انتگرال را در (۶۷) بدون محدودیت‌های اضافی (مانند اعمال یک پنجره صاف) محاسبه کنیم، اما به‌طور شهودی برای هر زاویه $\theta_o \neq \theta$ ، تصویر $RA_f(\theta, t)$ در t صاف است، در حالی که برای $\theta = \theta_o$ ، نتیجه یک تابع

¹⁰Candes

پله‌ای یک بعدی در t با تکنیکی در t_0 است. بنابراین طبیعی است که یک تبدیل موجک در امتداد t گرفته شود که منجر به تعریف تبدیل رجلت پیوسته [۲] به صورت زیر می‌شود

$$(۶۸) \quad RI_f(a, b, \theta) = \int \Psi_{a,b}(t) RA_f(\theta, t) dt,$$

که $\Psi_{a,b}(t) = a^{-1/2} \Psi(t - b/a)$ و $\Psi(t)$ یک موجک با داشتن حداقل یک گشتاور صفر است. «اتم‌های» آنالیز برجستگی‌های نامتناهی در زاویه θ ، مکان b و مقیاس a هستند، جایی‌که مشخصات برآمدگی توسط موجک داده می‌شود. یکی از این رجلت‌ها در شکل ۱۲ نشان داده شده است.

تبدیل در (۶۸) در تمام پارامترها پیوسته است. گسسته‌سازی و جایگزینی مناسب منجر به مجموعه‌هایی از برآمدگی‌های جهت‌دار موضعی می‌شود که می‌توانند به طور مؤثر توابع دو بعدی را با ناپیوستگی‌های لبه مانند نشان دهند (به طور دقیق‌تر، می‌توان نشان داد که یک نمایش چهارچوب مانند به دست می‌آید [۳]).

همچنین می‌توان طرح‌های گسسته‌ای را توسعه داد که مستقیماً روی داده‌های نمونه‌برداری شده با اندازه محدود کار کنند، در حالی‌که اصول تبدیل ریجلت را شبیه‌سازی می‌کنند. چنین طرح‌هایی معمولاً یک تبدیل گسسته رادون را در فضا یا حوزه فوریه (در مورد دوم، با استفاده از قضیه برش تصویر) و به دنبال آن تبدیل موجک گسسته انجام می‌شود. یک مثال در [۴۰] آورده شده است، جایی‌که قاب‌های ریجلت برای حذف نویز تصویر ساخته می‌شوند. مثال دیگر ساخت یک پایه رجلت متعارف از تبدیل محدود رادون و تبدیل موجک گسسته [۱۳] است.

نکته‌ی مهم در مورد ریجلت‌ها این است که برخلاف موجک‌ها، آن‌ها به تقریب سریع N - جمله‌ای اشیا با لبه‌های مستقیم دست می‌یابند [۳]. برای مثال یک تابع پله‌ای، یک تقریب غیرخطی M - جمله با

$$(۶۹) \quad \tilde{\epsilon}_M \sim 2^{-M}$$

به دست می‌دهد. بنابراین، آن‌ها می‌توانند به عنوان بلوک‌هایی در طرح‌های پیچیده‌تر برای تقریب اشیا با لبه‌های صاف استفاده شوند. یکی از این طرح‌ها از برآمدگی‌های موضعی با اندازه مناسب در امتداد لبه‌ها استفاده می‌کند (به عنوان مثال، با اندازه‌های ثابت یا تطبیقی [۴۰]). ریجلت‌ها را می‌توان با طرح‌های چند تفکیکی ترکیب کرد تا ریجلت‌های چند مقیاسی به دست آید. با ترکیب آن همراه با فیلتر باندگذر، باعث ایجاد منحنی‌ها می‌شود [۴].

۱۵. فیلتر بانک‌های جهت‌دار

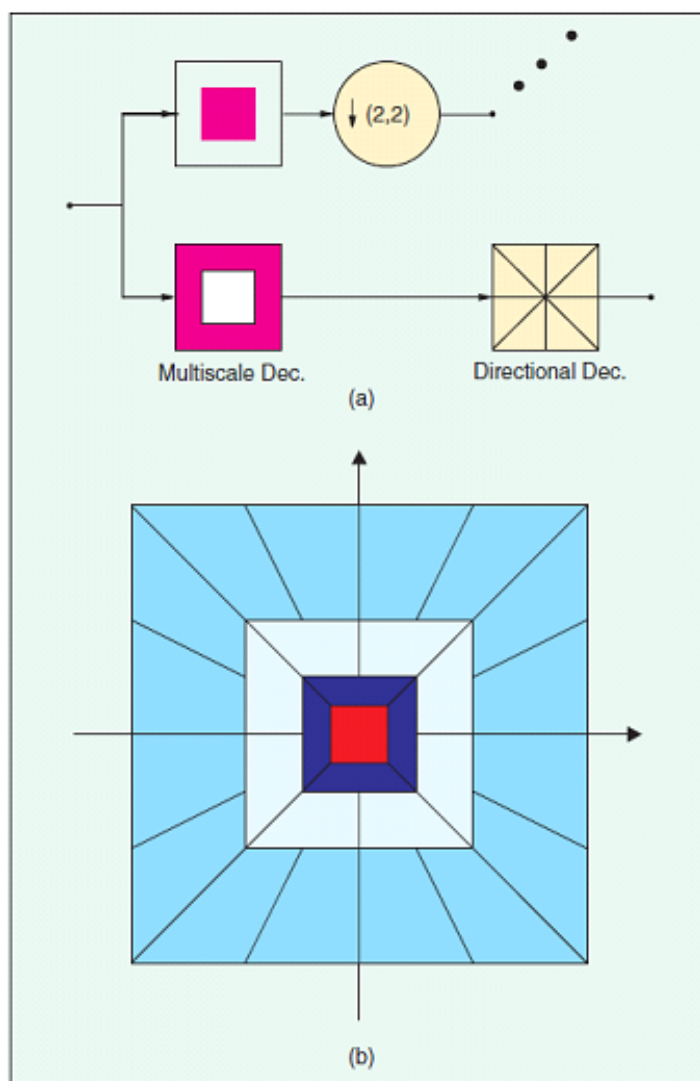
برای به دست آوردن تحلیل جهت‌دار، می‌توان از فیلتر بانک‌های جهت‌دار [۱] استفاده کرد. در چنین حالتی، توابع پایه توسط پاسخ‌های ضربه‌ای فیلتر و تبدیلات آن‌ها با توجه به گره‌های زیرشبکه‌ی نمونه‌گیری انجام می‌شود. چنین فیلتر بانک‌هایی را می‌توان به طور مستقیم یا از طریق تکرار فیلتر بانک‌های مقدماتی طراحی کرد. آن‌ها در صورت زیر نمونه‌گیری نه‌چندان مناسب به پایه‌ها یا اگر بیش از حد نمونه برداری شوند به قاب‌ها منتهی می‌شوند.

شکل ۱۳ تقسیم فرکانس بدست آمده توسط یک فیلتر بانک جهت‌دار ایده آل با هشت کانال را نشان می‌دهد. هنگامی که این با یک تجزیه چندریزه‌سازی ترکیب می‌شود، تجزیه منحنی مانند به دست می‌آید [۴، ۱۴]. به طور خاص، تجزیه هرمی به کانال‌های باندگذر (نگاه کنید به شکل ۱۴ (a)) توسط تحلیل جهت‌دار کانال‌های باندگذر ادامه می‌یابد. تعداد جهت‌ها با افزایش فرکانس افزایش می‌یابد و تقسیم فرکانس حاصل در شکل ۱۴ (b) نشان داده شده است. این سیستم فیلتر بانک جهت‌دار هرمی PDFB نامیده می‌شود. یک تصویر آزمایشی و تجزیه آن در شکل ۱۵ نشان داده شده است که نشان می‌دهد چگونه جهت‌های مختلف از هم جدا شده‌اند و این یک روش چندریزه‌سازی است. به عنوان مثال، یک الگوریتم نویززدایی ساده (بر اساس آستانه‌گذاری ضرایب با اندازه کوچک در تجزیه) هم در تجزیه موجک (این یک روش حذف نویز استاندارد است) و هم برای تجزیه فیلتر بانک جهت‌دار هرمی اعمال می‌شود. همان‌طور که در شکل ۱۶ مشاهده می‌شود، PDFB جهت‌گیری را به طور مؤثرتری دریافت می‌کند و نتایج

بصری دلپذیرتر و عملکرد SNR بهتری تولید می‌کند. این نشان‌دهنده پتانسیل چنین ایده‌های چندریزه‌سازی تفکیک ناپذیر جهت دار می‌باشد.

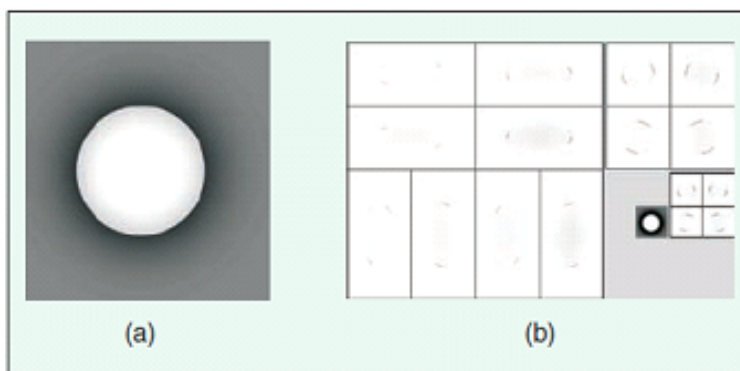
۱۶. پایه‌های دو بعدی و فشرده‌سازی

همان‌طور که در یک بعد دیده بودیم، یک تقریب $-N$ جمله‌ای خوب، تضمینی برای فشرده‌سازی خوب نیست. در حالی که یک تقریب $-N$ جمله‌ای مطلوب هرگاه با فشرده‌سازی مناسب همراه باشد.



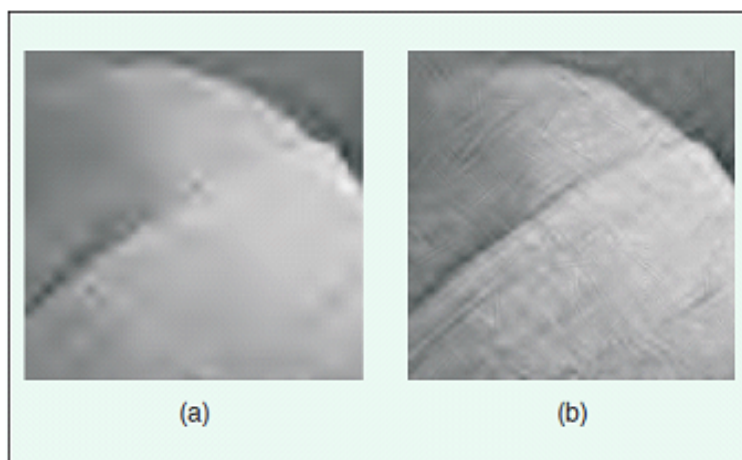
شکل ۱۴. بانک فیلتر جهت‌دار هرمی. (الف) یک تجزیه استاندارد چند مقیاسی به باندهای هشت‌تایی، که در آن کانال پایین‌گذر زیر نمونه برداری می‌شود، در حالی که بالاگذر نه. (ب) تقسیم فرکانس حاصل، که در آن تعداد جهت‌ها با فرکانس افزایش می‌یابد.

Figure 14: Pyramidal directional filter bank. (a) A standard multiscale decomposition into octave bands, where the lowpass channel is subsampled while the highpass is not. (b) Resulting frequency division, where the number of directions is increased with frequency.



شکل ۱۵. تجزیه بانک فیلتر جهت‌دار هرمی. (الف) تصویر آزمایشی ساده. (ب) تجزیه به تصاویر باند گذر جهتی (۸، ۴ و ۴) و تصویر پایین‌گذر.

Figure 15: Pyramidal directional filter bank decomposition. (a) Simple test image. (b) Decomposition into directional bandpass images (8, 4 and 4) and a lowpass image.



شکل ۱۶. مقایسه روش‌های حذف نویز مبتنی بر آستانه با استفاده از موجک (الف) و بانک فیلتر جهت‌دار هرمی (ب). SNR به ترتیب ۱۳٫۸۲ دسی‌بل و ۱۵٫۴۲ دسی‌بل است.

Figure 16: Comparison of threshold-based denoising methods using wavelets (a) and a pyramidal directional filter bank (b). The SNR is 13.82 dB and 15.42 dB, respectively.

بنابراین، موضوع فشرده‌سازی توابع تکه‌ای هموار دوبعدی هنوز در دست بررسی است. چندین رویکرد امیدوارکننده در حال حاضر در دست بررسی هستند، از جمله فشرده‌سازی در حوزه رجت و منحنی، فشرده‌سازی در امتداد منحنی‌ها با استفاده از "بندلیت" [۳۰] و تعمیم الگو در دو بعد یا لبه [۱۷].

۱۷. نتیجه‌گیری

تأثیر متقابل نمایش، تقریب و فشرده‌سازی سیگنال‌ها مورد بررسی قرار گرفت. برای سیگنال‌های قطعه‌ای هموار، قدرت روش‌های مبتنی بر موجک را به‌ویژه برای حالت تک بعدی نشان دادیم. برای سیگنال‌های دو بعدی، جایی که موجک‌ها پاسخی

برای سیگنال‌های قطعه‌ای هموار با تکنیکی منحنی ارائه نمی‌کنند، رویکردهای جدید و مشکلات در حال بررسی نشان داده شد. چنین رویکردهایی برای مسائلی مانند حذف نویز، فشرده‌سازی و طبقه‌بندی بر پایه‌های جدید با تأثیر بالقوه بالا بر پردازش تصویر متکی هستند.

قدردانی

نظرات داوران با کمال تشکر و قدردانی پذیرفته می‌شود. نویسنده از پی. ال. دراگوتی^{۱۱} و ام. دو^{۱۲} هر دو از EPFL برای نظرات روشن‌گرانه و برای ارائه مشکل‌ها، پیش‌نویس و پایه‌های جهت‌دار، تشکر می‌کند. بحث و گفتگو با بسیاری از افراد درگیر در این موضوع بین رشته‌ای بسیار مفید بوده است، به‌ویژه با دابشیز^{۱۳}، دیور^{۱۴}، دونوهو^{۱۵}، مالات^{۱۶} و انسر^{۱۷}. مارتین وترلی مدرک مهندسی خود را از Eidgenossische Technische Hochschule Zuerich (ETHZ) کارشناسی ارشد خود را از دانشگاه استنفورد، و دکترای خود را از Ecole Polytechnique Federale de Lozanne (EPFL) دریافت کرد. او قبل از پیوستن به بخش سیستم‌های ارتباطی مؤسسه فناوری فدرال سوئیس در لوزان، در دانشگاه کلمبیا، نیویورک و دانشگاه کالیفرنیا در برکلی بود. علایق تحقیقاتی او شامل پردازش سیگنال و ارتباطات، به‌ویژه، تئوری موجک و برنامه‌های کاربردی، فشرده‌سازی تصویر و ویدئو، کدگذاری مشترک منبع-کانال، و سیستم‌های ارتباطی خودسازماندهی‌کننده است. او برنده‌ی جوایز بهترین مقاله از EURASIP در سال ۱۹۸۴ و انجمن پردازش سیگنال IEEE در سال‌های ۱۹۹۱ و ۱۹۹۶ شده است. او به همراه جی. کواچویچ نویسنده‌ی کتاب درسی موجک‌ها و کدگذاری زیرباند (Wavelets and Subband Coding (Prentice-Hall, 1995)) هستند.

مراجع

- [1] R. H. Bamberger and M. J. T. Smith, A filter bank for the directional decomposition of images: Theory and design, *IEEE Trans. Signal Processing*, **40** (1992) 882–893.
- [2] E. Candès, Ridgelets: Theory and applications, Ph.D. dissertation, Dept. Statistics, Stanford University, Stanford, CA, 1998.
- [3] E. Candès and D. L. Donoho, Ridgelets: A key to higher-dimensional intermittency?, *Phil. Trans. R. Soc. London A.*, (1999) 2495–2509.
- [4] E. J. Candès and D. L. Donoho, Curvelets—A surprisingly effective nonadaptive representation for objects with edges, in *Curve and Surface Fitting*, A. Cohen, C. Rabutand, and L. L. Schumaker, Eds. Saint-Malo: Vanderbilt University Press, 1999.
- [5] A. Cohen, W. Dahmen, and I. Daubechies, Tree approximation and optimale ncoding, *Appl. Computational Harmonic Anal.*, to be published.
- [6] R. R. Coifman and M. V. Wickerhauser, Entropy-based algorithms for best basis selection, *IEEE Trans. Inform. Theory (Special Issue on Wavelet Transforms and Multiresolution Signal Analysis)*, **38** (1992) 713–718.
- [7] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [8] M. Crouse, R. D. Nowak, and R. G. Baraniuk, Wavelet-based signal processing using hidden Markov models, *IEEE Trans. Signal Processing (Special Issue on Wavelets and Filterbanks)*, (1998) 886–902.
- [9] I. Daubechies, Orthonormal bases of compactly supported wavelets, *Commun. Pure Appl. Math.*, **41** (1988) 909–996.
- [10] I. Daubechies, *Ten Lectures on Wavelets*, Philadelphia, PA: SIAM, 1992.

¹¹Dragotti ¹²M. Do ¹³Daubechies ¹⁴DeVore ¹⁵Donoho ¹⁶Mallat ¹⁷Unser

- [11] S. R. Deans, *The Radon Transform and Some of its Applications*, New York: Wiley, 1983.
- [12] R. A. DeVore, B. Jawerth, and B. J. Lucier, Image compression through wavelet transform coding, *IEEE Trans. Inform. Theory (Special Issue on Wavelet Transforms and Multiresolution Signal Analysis)*, **38** (1992) 719–746.
- [13] M. Do and M. Vetterli, Orthonormal finite ridgelet transform for image compression, in *Proc. IEEE Int. Conf. Image Processing, ICIP 2000*, Vancouver, Canada, Sept., (2000) 367–370.
- [14] M. Do and M. Vetterli, Pyramidal directional filter banks and curvelets, in *Proc. IEEE Int. Conf. Image Processing, ICIP 2001*, Patras, Greece, Oct. 2001.
- [15] D. Donoho, M. Vetterli, R. DeVore, and I Daubechies, Data compression and harmonic analysis, *IEEE Trans. Inform Theory (Special Issue, Information Theory: 1948-1998 Commemorative Issue)*, **44** (1998) 2435–2476.
- [16] P. L. Dragotti and M. Vetterli, Wavelet transform footprints: Catching singularities for compression and denoising, in *Proc. IEEE Int. Conf. Image Processing, ICIP 2000*, Vancouver, Canada, Sept., (2000) 363–366.
- [17] P. L. Dragotti and M. Vetterli, Footprints and edgeprints for image denoising and compression, in *Proc. IEEE Int. Conf. Image Processing, ICIP 2001*, Patras, Greece, 2001.
- [18] J. Fourier, *Théorie Analytique de la Chaleur*, Paris, France: Gauthier-Villars, 1888.
- [19] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Norwell, MA: Kluwer, 1992.
- [20] V. K. Goyal, Theoretical foundations of transform coding, *IEEE Signal Processing Mag.*, **18** (2001) 9–21.
- [21] V. K. Goyal, *Transform Coding*, SIAM, to be published.
- [22] V. K. Goyal, J. Zhuang and M. Vetterli, Transform coding with backward adaptive updates, *IEEE Trans. Inform. Theory*, to be published.
- [23] S. Mallat, A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Trans. Pattern Recognition Machine Intell.*, **11** (1989) 674–693.
- [24] S. Mallat, *A Wavelet Tour of Signal Processing*, San Diego, CA: Academic, 1998.
- [25] S. Mallat and W. L. Hwang, Singularity detection and processing with wavelets, *IEEE Trans. Inform. Theory (Special Issue on Wavelet Transforms and Multiresolution Signal Analysis)*, **38** (1992) 617–643.
- [26] S. G. Mallat and Z. Zhang, Matching pursuits with time-frequency dictionaries, *IEEE Trans. Signal Processing (Special Issue on Wavelets and Signal Processing)*, **41** (1993) 3397–3415.
- [27] F. Mintzer, Filters for distortion-free two-band multirate filter banks, *IEEE Trans. Acoust. Speech Signal Processing*, **33** (1985) 626–630.
- [28] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [29] A. Ortega, K. Ramchandran, and M. Vetterli, Optimal trellis-based buffered compression and fast approximations, *IEEE Trans. Image Processing*, **3** (1994) 26–40.
- [30] E. Le Pennec and S. Mallat, Image compression with geometric wavelets, in *Proc. IEEE Int. Conf. Image Processing, ICIP 2000*, Vancouver, Canada, Sept. (2000) 661–664.
- [31] P. Prandoni, Optimal segmentation techniques for piecewise stationary signals, Ph.D. dissertation, EPFL, Communications Systems, June 1999.
- [32] P. Prandoni and M. Vetterli, Approximation and compression of piecewise-smooth functions, *Phil. Trans. Roy. Soc. London*, **357** (1999) p. 1760.

- [33] J. Radon, Ueber die bestimmung von funktionen durch ihre integralwerte längst gewisser mannigfaltigkeiten, *Berichte Sächsische Akademie der Wissenschaften, Leipzig*, (1917) 262-267.
- [34] K. Ramchandran and M. Vetterli, Best wavelet packet bases in a rate-distortion sense, *IEEE Trans. Image Processing*, **2** (1993) 160-175.
- [35] A. Said and W. A. Pearlman, A new, fast, and efficient image codec based on set partitioning in hierarchical trees, *IEEE Trans. Circuits Syst. Video*, **6** (1996) 243-249.
- [36] C. B. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.*, **27** (1948).
- [37] J. M. Shapiro, Embedded image coding using zerotrees of wavelet coefficients, *IEEE Trans. Signal Processing (Special Issue, Wavelets and Signal Processing)*, **41** (1993) 3445-3462.
- [38] T. Skodras, C. Christopoulos, and T. Ebrahimi, The JPEG 2000 still image compression standard, *IEEE Signal Processing Mag.*, **18** (2001) 36-58.
- [39] M. J. T. Smith and T. P. Barnwell III, Exact reconstruction for tree-structured subband coders, *IEEE Trans. Acoust., Speech, and Signal Processing*, **34** (1986) 431-441.
- [40] J. L. Starck, E. Candès, and D. Donoho, The curvelet transform for image denoising, *IEEE Trans. Image Processing*, submitted for publication.
- [41] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Cambridge, MA: Wellesley-Cambridge, 1996.
- [42] B. Usevitch, Wavelet-based image compression, *IEEE Signal Processing Mag.*, **18** (2001) 22-35.
- [43] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [44] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding*, Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [45] C. Weidmann, Oligoquantization in low-rate lossy source coding, Ph.D. dissertation, EPFL, Communication Systems, July 2000.

سهرکهوت عبدی

گروه ریاضی کاربردی، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران
sarkout_abdi2011@yahoo.com

سهرکهوت عبدی دانشجوی دکتری گروه آموزشی ریاضی کاربردی دانشگاه آزاد اسلامی واحد رشت می باشد.



آرام عزیزی

گروه ریاضی کاربردی، دانشگاه پیام نور، تهران، ایران
a.azizi@pnu.ac.ir

آرام عزیزی استادیار گروه آموزشی ریاضی کاربردی دانشگاه پیام نور بوده و زمینه تحقیقاتی ایشان حل عددی معادلات دیفرانسیل و انتگرال به کمک موجک، منظم سازی دستگاه در حل معادلات خطی می باشد.



محمود شفیعی

گروه ریاضی کاربردی، واحد رشت، دانشگاه آزاد اسلامی، رشت، ایران

hafiei@iau.ac.ir

محمود شفیعی استادیار گروه آموزشی ریاضی کاربردی دانشگاه آزاد اسلامی واحد رشت بوده و زمینه تحقیقاتی ایشان روش‌های عددی در حل معادلات تابعی می‌باشد.



جمشید سعیدیان

گروه ریاضی کاربردی، دانشکده علوم ریاضی و کامپیوتر، دانشگاه خوارزمی، تهران، ایران

j.saeidian@khu.ac.ir

جمشید سعیدیان استادیار گروه آموزشی علوم کامپیوتر دانشکده علوم ریاضی و کامپیوتر دانشگاه خوارزمی تهران بوده و زمینه تحقیقاتی ایشان اسپلاین‌های حافظ شکل می‌باشد.

