rrent Trends in Industrial and Applied Mathematics
Manchanda, K. Ahmad and A.H. Siddiqi (Eds)

1. Role of Functional Analytic Methods in Imaging Science During the 21st Century

P. Manchanda and A.H. Siddiqi*

Department of Mathematics, Guru Nanak Dev University, Amritsar-143 005, India
e-mail: kmanch@jla.vsnl.net.in

*Department of Mathematical Sciences, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia
e-mail: ahasan@kfupm.edu.sa

Introduction

ınctional analytic methods were developed in the beginning of the 20th century ith the pioneering work of Stefan Banach and David Hilbert. Since then, these ethods have been applied in diverse fields of mathematical sciences and other eas of science and technology. The 21st century is the era of information chnology and therefore in this article we present the relevance of numerical ıd analytic aspects of functional analysis in the image processing, an important gredient of information technology. As Neunzert and Siddiqi [53] write in hapter 5 on Image Processing and Fourier-Wavelet Methods: "The image ·ocessing and signal analysis whose ingredients are modelling, transforms, noothing and sharpening, restoration, encoding (image data compression, image ansmission, feature extraction), decoding, segmentation, representation, archiving ıd description, have played a vital role in understanding the intricacies of ıture, for providing comforts, luxuries and pleasure to all of us, and even ·obing mission of a space-craft." Image processing is used in telecommunications elephone and television), photocopying machine, video camera, fax machine, ansmission and analysis of satellite images, medical imaging (echography, ımography and nuclear magnetic resonance), warfare, artificial vision, climatology, ıeteorology and film industry for imagery scenes. In short, it is one of the most nportant disciplines for industrial development and unveils the secrets of nature. 'ifferent kinds of techniques and tools like Fourier series, Fourier transform, /alsh-Fourier transform, Haar-Fourier transform, Hotelling transform, Hadamard ansform, entropy encoding and, more recently, wavelets, wavelet packets, and ·actal methodology have been used to tackle the problems of this field. It is ifficult to say authoritatively which method is superior to the other in a particular .tuation. However, a combination of the wavelets and fractal methods promises ɔr a great future.

Towards the later part of 2000, two major academic activities were organized: ne in the Abdus Salam International Centre of Theoretical Physics, Trieste,

Italy, School on Mathematical Problems in Image Processing 14–22 September 2000, and other in U.S.A., Yosemite Multiscale and Multiresolution Symposium. Multiscale and Multiresolution Methods, Yosemite National Park, October 29-November 1, 2000, sponsored by NSF, NASA, IBM, SCI, and MIT. Deliberations of these conferences clearly indicate that future developments in many areas of science, social science and technology are closely related to advances in numerical and functional analytic methods which are known by the name numerical and applied functional analysis, comprising variational methods, finite element, boundary elements, volume element and particle methods, algorithmic optimization, parallel algorithms and modeling, classical and refined and generalized Fourier analysis (Fourier analysis, Walsh dyadic harmonic analysis, wavelet analysis, wavelet packet analysis, ridgelets and curvelets) and fractals.

In Section 2, a model of imaging process proposed by Stan Osher and Leonid Rudin for a function of bounded variation is presented. Then the efficiency of wavelet-based algorithms related to the remarkable properties of wavelet expansions of functions with bounded variations ($BV(R^2)$) is discussed. Proper understanding of the behaviour of wavelet coefficients of functions of bounded variation is of vital importance for analyzing various developments in image processing. A brief introduction to fractal image compression is presented in Section 4 along with a fairly general theorem on the inverse problem.

Section 5 is devoted to miscellaneous problems like classification of images, watermarking of images and data analysis. It has been established that there is a close relationship between image processing and partial differential equations. An updated account of this aspect along with relevant details is lucidly presented in Guichard [35]. Due to limitation of space, we are unable to include this elegant aspect of image processing.

2. Modeling of Images

2.1 General Framework

An analogue image on a domain Ω can be viewed as a function $f(x_1, x_2) = f(x)$ belonging to the Hilbert space $L_2(\Omega)$. The energy of such an image is defined as

$$\int_\Omega |f(x)|^2 \, dx.$$

In order to sample this analogue image into a digital image, we need to fix a grid defined as $N^{-1}Z \times N^{-1}Z$ for some large N. A fine grid is a grid where $N = 2^j$. It is clear that if such grids are denoted by Γ_j, then $\Gamma_j \subseteq \Gamma_{j+1}$. A digital image f_j is a matrix indexed by points in Γ_j. One can also define sampling by computing the coefficients of images in some orthonormal basis Z. If D is the unit square $[0, 1] \times [0, 1]$, this digital image $f_j \in l_2(\Gamma)$ is now a huge matrix $c_{k,l} = f_j(x_1, x_2)$ where $x_1 = k2^{-j}$, $x_2 = l2^{-j}$, k and l ranging from 0 to $2^j - 1$. These entries $c_{k,l}$ are called pixels and each $c_{k,l}$ measures the gray level of the given image at $(k2^{-j}, l2^{-j})$. This is the case for a white and black image, and a color image has a similar definition with the difference that $c_{k,l}$ is now vector valued. A digital image can be viewed as a vector inside a 4^j-dimensional vector space. The gray

levels $c_{k,l}$ are finally quantized with a 8-bit precision which provides 256 gray levels. This discrete representation of an image needs to be compressed for efficient storing or transmission.

All scientists working on image processing agree on the possibility of compressing images but they have different viewpoints about the precise mathematical description of the models for natural images on which the compression algorithms crucially rely. In this article, we mainly confine ourselves to wavelet and fractal methods.

2.2 $u + v$ Models of Osher and Rudin

In $u + v$ model, images are assumed to be a sum of two components $u(x)$ and $v(x)$. The first component $u(x)$ models the objects or features of an image while the $v(x)$ term represents the texture and noise in the same figure. If a given image $f(x)$ is written as $f(x) = u(x) + v(x)$, then $u(x)$ will provide a sketchy approximation to the given image. The best method of image compression is that in which $u(x)$ provides the best approximation of $f(x)$ retaining its important features, while $v(x)$ is significantly small and could be neglected. u may be an element of a ball of a Banach space; for example, Banach space of the functions with bounded variation, BV or a Besov space, $B_{s,p}(\Omega)$.

Quantization is a crucial ingredient of a compression scheme. In a compression scheme, the first step maps a given image into a string of coefficients. These coefficients are real numbers which are replaced by some digital approximations depending on the computer precision. Once the quantization is performed, the image can be transmitted and the reconstructed image will be affected by the quantization. The error which affects the reconstructed image heavily depends on the orthonormal basis Z under consideration. We would like that this quantization error be less harmful to the $u(x)$ component of $f(x)$ than the $v(x)$ component. In the ideal case, $u(x)$ should be untouched while $v(x)$ might disappear. The quantization would de-noise the given image. An image is always distorted after quantization and transmission and so one would like to use a basis which gives the minimum distortion.

Thresholding is closely related to quantization and raises exciting mathematical problems. It may be accepted as the working principle that the behaviour of an image after quantization will be adjudged by its behavior after thresholding.

Let us define thresholding. A given signal or image $f(x)$ is decomposed in some orthonormal basis Z of a Hilbert space H, that is, $f(x) = \sum_{n=0}^{\infty} c_n l_n(x)$, where $Z = \{l_n(x)\}$ is an orthonormal basis in a Hilbert space H and $c_n = \langle f, l_n(x)\rangle$; $n = 0, 1, 2, \ldots$. All coefficients c_n such that $|c_n| < \varepsilon$ where ε is a given threshold are viewed as being insignificant and replaced by 0. In other words, we define $q_\varepsilon(x) = x$ if $|x| > \varepsilon$ and $q_\varepsilon(x) = 0$ otherwise. The thresholding operator θ_ε is defined by $\theta_\varepsilon(t) = \sum_{n=0}^{\infty} q_\varepsilon(c_n) l_n(x)$. The resulting error in the H norm is

$$R_\varepsilon = \| \theta_\varepsilon(f) - f \|_H \tag{2.1}$$

If the coefficients (c_n) are sorted out (rearranged) as a non-increasing sequence c_n^*, then

$$R_\varepsilon = \left(\sum_{n>N} (c_n^*)^2 \right)^{1/2} \tag{2.2}$$

where N is defined as the lowest index n such that $c_n^* < \varepsilon$. The error R_ε depends on the threshold ε and on the decay of c_n^*. This decay depends on the signal or image and also on the specific orthonormal basis. An interesting problem is to compute the error R_ε in the case of a more general setting, namely, in the case of a Banach space.

Let us assume that an orthonormal basis $\{l_n(x)\}$ can be found such that the sorted coefficients of $u(x)$ have a fast decay. Then thresholding the coefficients c_n of $f(x)$ in this basis, we will retain much of the energy of the $u(x)$ component and wipe out $v(x)$. The expansion of u will be compressed with few terms in this particular basis. It has been established that wavelet analysis performs much better than Fourier analysis under the following two conditions:

(a) A $u + v$ model based on specific Banach space X is adapted to the problem.

(b) The wavelet coefficients of functions in X should plunge much faster than the corresponding Fourier coefficients. For instance, the sorted wavelet coefficients of a function with bounded variation decay as $1/n$ (Theorem 3.2) while the sorted Fourier coefficients of functions with bounded variation may decay as badly as $n^{-1/2}$ if we ignore logarithmic factors. Theorem 3.2 explains the performances of wavelets in image compression for the Osher-Rudin model for still images. Osher-Rudin model is reminiscent of approximation theory and more precisely it mimics the theory of interpolation. In this context, one wants to write a generic function $f(x)$ as a sum of a "good function" $u(x)$ which is more regular than $f(x)$ plus a "bad function" $v(x)$ which is small in some sense. An example is the celebrated Calderón-Zygmund decomposition of an L_1 function $f(x)$ into a sum of L_2 function $u(x)$ plus an oscillating part $v(x)$ carried by a set with a small measure.

We explain now why the space $BV(R^2)$ is suited for $u + v$ images model. As we know in the case of image processing, we want to detect objects delimited by contours. Then these objects can be modeled by some planar domains Ω_1, $\Omega_2, \ldots, \Omega_n$ and the corresponding contours or edges will be modeled by their boundaries $\partial\Omega_1, \partial\Omega_2, \ldots, \partial\Omega_n$. Functions $u(x)$ in $u + v$ model is assumed to be smooth inside $\Omega_1, \Omega_2, \ldots, \Omega_n$ and with jump discontinuities across the boundaries $\partial\Omega_1, \partial\Omega_2, \ldots, \partial\Omega_n$. The sum of the lengths of these edges $\partial\Omega_1, \partial\Omega_2, \ldots \partial\Omega_n$ is approximately one of the two terms in the BV norm of $u(x)$. Thus we have a $u + v$ model where $u(x)$ belongs to $BV(R^2)$ and the energy norm of $v(x)$; that is,

$\int |v(x)|^2 \, dx$ is sufficiently small.

For a vector $\mu \in \mathbb{R}^2$, we define the difference operator Δ_μ in the directon μ by

$$\Delta_\mu(f, x) := f(x + \mu) - f(x).$$

Let Ω be any domain in $\mathbb{R}^2$. For functions f defined on Ω, $\Delta_\mu(f, x)$ is defined whenever $x \in \Omega(\mu)$, where $\Omega(\mu) := \{x: [x, x + \mu] \subset \Omega\}$ and $[x, x + \mu]$ is the line segment connecting x and $x + \mu$. Note that if Ω is bounded and μ is large enough, then $\Omega(\mu)$ is empty. Let $e_j, j = 1, 2$, be the two coordinate vectors in $\mathbb{R}^2$. We say that a function $f \in L_1(\Omega)$ is in $BV(\Omega)$ (space of functions of bounded variation) if and only if

$$V_\Omega(f) := \sup_{0<h} h^{-1} \sum_{j=1}^{2} \| \Delta_{he_j}(f, \cdot) \|_{L_1(\Omega(he_j))} = \lim_{h \to 0} \sum_{j=1}^{2} \| \Delta_{he_j}(f, \cdot) \|_{L_1(\Omega(he_j))}$$

is finite.

The quantity $V_\Omega(f)$ is the *variation* of f over Ω. It provides a semi-norm and norm for $BV(\Omega)$:

$$|f|_{BV(\Omega)} := V_\Omega(f); \; \|f\|_{BV(\Omega)} := |f|_{BV(\Omega)} + \|f\|_{L_1(\Omega)} \tag{2.3}$$

There are several variants of $u + v$ model (see Meyer [47]); we discuss here briefly some of these models. The first model which is discussed here is due to Osher and Rudin [54]. Its mathematical formulation is that a given function $f(x)$ is the sum $f(x) = u(x) + v(x)$ with explicit bounds on the BV norm of the unknown function $u(x)$ and on the L_2 norm of the unknown function $v(x)$, we want to recover these unknown functions $u(x)$ and $v(x)$. The explicit condition on $u(x)$ states $\| u \|_{BV} \leq C$ and the one on $v(x)$ reads $\| v \|_{L_2} < \varepsilon$. There is no uniqueness and some additional conditions are needed to find u and v. To prove existence and uniqueness of this optimal decomposition, it suffices to consider the closed subset K of $L_2(R^2)$ defined by $\| u \|_{BV} \leq C$ and to define the optimal u as the point in K which minimizes the L_2 distance to $f(x)$.

The second and related problem consists in finding a fast algorithm that would yield a sub-optimal decomposition $f(x) = g(x) + h(x)$ where the corresponding bounds for g and h might be enlarged by a fixed multiplicative amount (see Cohen et al. [22]).

A third approach to find u and v is as follows: Given a function $f(x)$ in $L_2(R^2)$, we want to solve the variation problem:

$$w(\lambda) = \inf \{J(u) = \| u \|_{BV} + \lambda \| v \|_{L_2}; \quad f = u + v\} \tag{2.4}$$

The tuning given by the large factor $\lambda = \varepsilon^{-1}$ implies that the L_2-norm of v should be of the order of magnitude of ε. It is clear that solving this variational problem yields a suboptimal decomposition of $f(x)$. One may be interested in relating the growth of $w(\lambda)$ as λ tends to infinity to some properties of the L_2 function $f(x)$. For example, the space of all functions $f(x)$ for which $w(\lambda) = o(\lambda^r)$ and λ tends to infinity will be characterized by Theorem 3.3.

A fourth approach to the decomposition $u + v$ was proposed by DeVore et al., (see DeVore et al., [26] and Chambolle et al., [16]) where BV norm is replaced by a Besov norm in the definition of J.

A fifth approach was proposed by Mumford and Shah (for details see [51, 52] and for further variants, see Meyer [47]).

3. Fourier Analysis to Wavelets, Ridgelets and Curvelets Analysis

It was quite clear from the very beginning of the study of signal processing that the classical Fourier analysis is not appropriate for real-life signals (see, for example, Meyer [45], Neunzert and Siddiqi [53]). In late forties, stalwarts like John von Neumann, Dennis Gabor, Leon Brillouix Eugene Wigner addressed the shortcomings of Fourier analysis and advocacted for a windowed Fourier transform. Gabor introduced the concept of short-time Fourier transform (called windowed Fourier transform or Gabor transform; for details see Mallat [39]). The computational hazard of this concept was realized soon by its inventor Gabor and his coworkers von Neumann et al,. There were some attempts to define optimal sampling but physicists Francis Low and Roger Balian proved that there exist L_2 functions that cannot be decomposed into a convergent series related to Gabor waves ($g_{w,s} = w(t-s)e^{iwt}$, $w(t)$ is a window function). Finally, Nobel Prize winner of 1984, Kenneth Wilson, reshaped Gabor transform and produced orthonormal time frequency atoms leading to fast algorithms for local Fourier analysis. His key idea was to alternate the DCT (see, for example, Neunzert and Siddiqi [53]) with the Discrete Sine Transform (DST) according to whether l is even or odd; l denotes the position of the interval and the DST uses the orthonormal basis of functions $\frac{1}{\sqrt{k}} \sin\left(\frac{k}{2}t\right)$, $k = 1, 2, 3, \ldots$.

Time-scale algorithms and wavelet analysis can be defined as an alternative to the classical windowed Fourier analysis and to time-frequency analysis. In wavelet analysis, one compares several magnifications of a signal with distinct resolutions. These magnifications are often called *zoomings*. A mother wavelet or simply wavelet is a function $\psi(t)$ which has a compact support (or a rapid decay at infinity) and satisfies the basic condition $\int_{-\infty}^{\infty} \psi(t)\, dt = 0$. This means that $\psi(t)$ is oscillating in some weak sense. Very often, $\psi(t)$ is assumed to be smooth and having an admissibility condition:

$$\int_0^{\infty} \frac{|\hat{\psi}(t)|^2}{t}\, dt < \infty.$$

Other wavelets are generated by the mother wavelet as

$$\psi_{a,b}(t) = \frac{1}{|a|^{1/2}} \psi\left(\frac{t-b}{a}\right),$$

where $a > 0$, $-\infty < b < \infty$.

$\psi_{a,b}(t)$ are building blocks of wavelet analysis. The parameter a measures the average width of the wavelet $\psi_{a,b}(t)$ while the parameter b gives the position or time. These dilations (by $1/a$) are precisely the magnifications we alluded to. The wavelet coefficients of a function $f(t)$ of the real variable t are the scalar products $W(a, b) = \langle f, \psi_{a,b} \rangle$. The original function $f(t)$ can be recovered as a linear combination of these wavelets $\psi_{a,b}(t)$ provided $f(t)$ satisfies the admissibility

condition (for details, see, for example, Neunzert and Siddiqi [53], Daubechies [25] and Meyer [44, 45]). More precisely, let $f(x) \in L_2(R^n)$ and $\psi(\cdot) : R^n \to R$ be a function satisfying $\int_0^\infty |\hat{\psi}(t\xi)|^2 \, dt/t = 1$, where $\hat{\psi}$ denotes the Fourier transform ($\xi \in R^n$). Then wavelet coefficients of f are defined by $F(x, t) = \langle f, \psi_{x,t} \rangle$, where

$$\psi_{x,t}(y) = t^{-n/2}\psi\left(\frac{y-x}{t}\right).$$

Caldéron's reproducing identity of Grossman-Morlet theorem states that $f(x)$ can be recovered and it is given by

$$f(x) = \int_0^\infty \int_{R^n} f(y, t)\psi_{y,t}(x)dy\, dt/t^{n+1}$$

This identity gives us a recipe for (i) measuring the local fluctuation coefficients of a given function f, around any point x, at any scale t and for (ii) reconstructing f with all these fluctuation coefficients. In other words, at any given scale $a > 0$, f is decomposed into the sum of a trend at the scale a and of a fluctuation around this trend. The trend is given by the contribution of this scale $t > a$ in Caldérons's reproducing identity and the fluctuation is given by the scales $t < a$.

In one-dimensional case, orthonormal wavelet bases are defined as follows. A mother wavelet satisfies the properties

(a) $\psi(t)$ is smooth function (with $r - 1$ continuous derivatives and a bounded derivative of order r).
(b) $\psi(t)$ together with its derivatives of order less than r has a rapid decay at infinity.
(c) The collection $\psi_{j,k}(t)$ defined by $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$, $j, k \in Z$ is an orthonormal basis for $L_2(R)$.

The first example of such a function $\psi(t)$ was invented by the Hungarian mathematician Alfred Haar in 1909 which is defined as $\psi(t) = 1$ on $[0, 1/2)$ and $\psi(t) = -1$ on $[1/2, 1)$ and $\psi(t) = 0$ elsewhere. In this case $r = 0$. In 1988, the young Belgian lady Ingrid Daubechies [25] proved that for each $r \geq 1$, one can construct a function $\psi(t)$ of class C^r with compact support and satisfying above conditions. Multiresolution analysis is an important concept of wavelet analysis. It is a family $\{V_j\}, j \in Z$ of closed subspaces V_j of $L_2(R)$ having the following properties:

(a) The intersection $\bigcap_j V_j, j \in Z$, is $\{0\}, \ldots V_{-1} \subseteq V_0 \subseteq V_1 \subseteq V_2 \subseteq V_3 \subseteq \ldots \subseteq V_j \subseteq V_{j+1} \ldots$.
(b) The union $\bigcup_j V_j, j \in Z$ is dense in $L_2(R)$.
(c) $f(t) \in V_j$ if and only if $f(2t) \in V_{j+1}$.
(d) There exists a smooth and localized function $\varphi(t)$ such that the collection $\varphi(t - k)$, $k \in Z$ is an orthonormal basis for V_0. $\varphi(t)$ is called the scaling function.

The relation between our wavelet basis and a multiresolution analysis is given by the condition that $\psi(t-k)$, $k \in Z$, is an orthonormal basis of the orthogonal complement W_0 of V_0 in V_1. By an obvious rescaling, one obtains the fact that $2^{j/2}\psi(2^j t - k)$, $k \in Z$, is an orthonormal basis for the orthonormal complement W_j of V_j into V_{j+1}. It is clear that the full collection $\{\psi_{j,k}\}$ is an orthonormal basis for $L_2(R)$. In this construction the mother wavelet is built from the scaling function $\varphi(t)$. The converse problem whether any wavelet basis $2^{j/2}\psi(2^j t - k)$, $j \in Z$, $k \in Z$ can be associated with a multiresolution analysis has a partial affirmative answer. That is, if $\psi(t)$ satisfies some reasonable smoothness localization properties, then it is a mother wavelet of a multiresolution analysis [37].

Multiresolution analysis is a mathematical concept which highlighted pyramidal algorithms and subband coding, two important concepts of image processing. The quadrature mirror filters used in subband coding are the matrix representation of the orthonormal decomposition $V_{j+1} = V_j \oplus W_j$.

The pyramidal algorithms and multiresolution analysis of $L_2(R^2)$ are closely related in the folowing way: If $\Gamma_j, j \in Z$, are the sampling grids, then the sampling operator $P_j : L_2(R^2) \to l_2(\Gamma_j)$ is nothing else but the orthogonal projection from $L_2(R^2)$ onto V_j. To clarify this correspondence, it suffices to associate each vector of the 2-D orthonormal basis $2^j\varphi(2^j x - k)$, $k \in Z^2$ of V_j to the corresponding point $k2^{-j}$ of the grid Γ_j. It will provide an isometric isomorphism between V_j and $l_2(\Gamma_j)$. The coarse-to-fine algorithm in the pyramidal algorithm reflects the canonical embedding of V_j inside V_{j+1} while the fine-to-coarse algorithm corresponds to the orthogonal projection from V_{j+1} onto V_j.

The two-dimensional wavelet is defined as

(i) $\psi_1(x_1, x_2) = \varphi(x_1)\psi(x_2)$
(ii) $\psi_2(x_1, x_2) = \psi(x_1)\varphi(x_2)$
(iii) $\psi_2(x_1, x_2) = \varphi(x_1)\psi(x_2)$

where $\psi(x)$ is one-dimensional mother wavelet and $\varphi(x)$ is the corresponding scaling function.

Theorem 3.1 For each positive exponent r, there exist three functions ψ_m, $m = 1, 2, 3$ with the following properties: each $\psi_m(x_1, x_2)$ is compactly supported and belongs to the Hölder space C^r and

$$2^j\psi_m(2^j x_1 - k_1, 2^j x_2 - k_2), j \in Z, \quad (k_1, k_2) \in Z^2$$

$m = 1, 2, 3$ is an orthonormal basis for $L_2(R^2)$.

In a recent paper Cohen, DeVore et al. [22] have explained nicely the role of numerical and functional analytic methods in image processing. They provide the solution of following problems:

Problem 3.1 Given a function (image) f defined on the unit square $\alpha = [0, 1)^2$ and parameter $t > 0$, find a function $g \in BV(Q)$ which attains the infimum

$$U(f, t) = \inf_{g \in BV(Q)} \| f - g \|^2_{L_2(Q)} + tV_Q(g), \tag{3.1}$$

where $BV(Q)$ is the space of functions of bounded variation on Q and $V_Q(f) = |f|_{BV}$ is the associated semi-norm, that is, the total variation of f.

Chambole et al. [16] have considered the related problem, namely,

Problem 3.2 Find $g \in B_1^1(L_1(Q))$ whih attains the infimum

$$\hat{U}(f,t) = \inf_{g \in B_1^1(L_1(Q))} \|f - g\|^2_{L_2(Q)} + t\,|g|_{B_1^1(L_1(Q))} \tag{3.2}$$

where the Besov space $B_1^1(L_1(Q))$ is taken in place of the (larger) space $BV(Q)$. Both $BV(Q)$ and $B_1^1(L_1(Q))$ are smoothness spaces of order one in $L_1(Q)$; that is, the space $BV(Q)$ is the same as $\mathrm{Lip}(1, L_1(Q))$. In contrast to BV, the $B_1^1(L_1)$ norm has a simple equivalent expression as the l_1 of the coefficients in a wavelet basis decomposition

$$f = \sum_{\lambda \in \Lambda} f_\lambda \psi_\lambda \tag{3.3}$$

(where Λ denotes the set of indices for the wavelet basis). One may get an equivalent discrete problem

$$\hat{U}((f_\lambda), t) = \inf_{(g_\lambda) \in l_1} \sum_{\lambda \in \Lambda} [\,|f_\lambda - g_\lambda|^2 + t\,|g_\lambda|\,], \tag{3.4}$$

whose solution (obtained by minimizing separately on each index λ) is exactly given by a soft thresholding procedure at level $t/2$:

$$g_\lambda = \mathrm{sgn}\,(f_\lambda)\,\max\,\{0, |f_\lambda| - t/2\}. \tag{3.5}$$

It may be observed that there is little distinction between Problems 3.1 and 3.2. However, BV seems to be more adaptive for real-life images, since it allows sharp edges (i.e., discontinuities on a line), which cannot occur in a bivariate function that belongs to the smaller space $B_1^1(L_1)$. One of the main tools for finding a solution of Problem 3.1 is the following result:

Theorem 3.2 [22] Let ψ_λ, $\lambda \in \Lambda$ be a two-dimensional orthonormal basis as described in Theorem 3.1. Then for every f in $BV(R^2)$, the wavelet coefficients $c_\lambda = \langle f, \psi_\lambda \rangle$, $\lambda \in \Lambda$ belong to weak $l_1(\Lambda)$.

Theorem 3.2 tells us that if $c_\lambda = \langle f, \psi_\lambda \rangle$ and $|c_\lambda|$, $\lambda \in \Lambda$ are sorted out by decreasing size, we obtain a non-increasing sequence c_n^* which satisfies $c_n^* \leq C/n$ for $1 \leq n$.

Theorem 3.3 [47] $\omega(\lambda)$ defined by the relation (2.4) satisfies $\omega(\lambda) = o(\lambda^r)$ as $\lambda \to \infty$ if and only if the sorted wavelet coefficients of $f(x)$ satisfy $c_n^* = o(n^{-\alpha})$ where $\alpha = 1 - \frac{r}{2}$.

It may be observed that numerical techniques for solving (3.1) based on partial differential equations have been developed and successfully applied to image processing, see for example [1, 6, 7, 8, 14, 15, 17, 18, 19, 35, 48, 54, 56,

83]. The advantage of these techniques is high performance and the disadvantage is that they are numerically intensive, and require in practice the approximation of the BV term in $U(f, t)$ by a quadratic term $\left(\text{e.g., } \int (\varepsilon + |\nabla f|^2)\right)$ in order to find a solution in reasonable computational time. For BV and Besov spaces, see for example [57] or [15] or [35].

The concepts of ridgelets and curvelets are still in infancy stages (for example, see references Cande's and Donoho [10]). As we know in many important imaging applications, images exhibit edges-discontinuities across curves. In traditional photographic imaging, this occurs whenever one object occludes another causing the luminance to undergo step discontinuities at boundaries. In biological imagery, this occurs whenever two different organs or tissue structures meet. Cande's [13] showed in his thesis that ridgelets and curvelets are appropriate tools to study these problems in higher dimensions. We briefly present these concepts and we refer to [10–13, 27, 28] for more details.

A ridgelet is a function of the form

$$\frac{1}{a^{1/2}}\psi\left(\frac{u \cdot x - b}{a}\right) \tag{3.6}$$

where a and b are scalar parameters and u is a vector of unit length. In the sequel, we assume that ψ satisfies $\int_R \frac{|\hat{\psi}(\zeta)|^2}{|\zeta|^d} d\xi = 1$. Of course, a ridgelet is a ridge function whose profile displays an oscillatory behavior (like a wavelet).

A ridgelet has a scale a, an orientation u, and a location parameter b. Ridgelets are concentrated around hyperplanes, roughly speaking the ridgelet (3.6) is supported near the strip $\{x, |u \cdot x - b| \le a\}$. Like wavelets, one can represent any function as a superposition of these ridgelets. Define the ridgelet coefficients

$$R_f(a, u, b) = \int f(x) a^{-1/2} \psi\left(\frac{u \cdot x - b}{a}\right) dx$$

then, for any $f \in L_1 \cap R_2(R^d)$, we have

$$f(x) = \left(\frac{1}{2\pi}\right)^{1-d} \int R_f(a, u, b) a^{-1/2} \psi\left(\frac{u \cdot x - b}{a}\right) d\mu(a, u, b)$$

where $d\mu(a, u, b) = da/a^{d+1}\, dudb$ (du being the uniform measure on the sphere). Furthermore, this formula is stable as one has a Parseval relation

$$\|f\|_2^2 = 2\pi\left(\frac{1}{2\pi}\right)^{1-d} \int |R_f(a, u, b)|^2 d\mu(a, u, b)$$

Discrete collection of ridgelets is defined as

$$\left\{\psi_{j,l,k}(x) = 2^{j/2}\psi(2^j u_{j,l} \cdot x - kb_0), j \ge j_0, u_{j,l} \in \sum_j, k \in Z\right\}$$

The scale a and location parameter b are discretized dyadically, as in the theory

of wavelets. However, unlike wavelets, ridgelets are directional variable u. This variable is sampled at increasing resolution, so that at scale j the discretized set $\sum_j$ is a net of nearly equispaced points at a distance of order 2^{-j}. In two dimensions, for instance, a ridgelet is of the form

$$\{2^{j/2}\psi(2^j(x_1 \cos \theta_{j,l} + x_2 \sin \theta_{j,l} - 2\pi k 2^{-j}))\} \ (j > j_0, l, k)$$

where the directional parameter $\theta_{j,l}$ is sampled with increasing angular resolution at increasingly fine scales such as $\theta_{j,l} = 2\pi l 2^{-j}$.

Curvelets provide a new multiresolution representation with several features that set them apart from existing representations such as wavelets, multiwavelets, etc. They are based on an anisotropic notion scaling. The natural scaling law that applies (at the origin) is of the form

$$f_a(x_1, x_2) = f(ax_1, ax_2)$$

which is known in Harmonic Analysis as a Parabolic Scaling. It is beyond the scope of this paper to describe this object which is based on combining several ideas: Ridgelets (described above), Multiscale Ridgelets (a pyramid of windowed ridgelets, renormalized and transported to a wide range of scales and locations) and Bandpass Filtering (a method of separating an object out into a series of disjoint scales).

4. Image Compression by Fractals

Image Compression Techniques and Methods are of vital importance in high technology, especially in information technology. These techniques and methods are related to communicating and storing images and data (information) in the shortest possible time (in minimum space on hard disk or CD-ROM) and retrieval with permissible distortion. Tools and techniques of the Fourier analysis like Fourier transform, convolution, Shannon sampling theorems and Walsh-Fourier methods have been successfully used for a long time. Matrix transform, optimization and variational techniques were also employed to study image and data compression. Wavelet theory entered this area in the late eighties and started replacing the role of DCT (JPEG) which was dominating the scene until that time in view of its high compression ratio and the quality of information after retrieval.

The concept of a fractal and the discipline of fractal geometry was introduced by Benôit Mandelbrot in the early eighties to study irregular shapes like coastal lines, mountains, clouds or rain fall. Fractals are complicated looking sets like Cantor set, Sierpinski gasket, Sierpinski carpet, von-Koch curve, Julia set, but they arise out of simple algorithms. By now it is a well-established discipline and a comprehensive and updated bibliography can be found in Barnsley [2, 3] Barnsley and Hurd [4], Lu [38] and references therein.

Barnsley [3] established a close connection between functional analysis, fractals and multimedia by demonstrating that fractals can be defined in terms of fixed points of mappings defined on an appropriate metric space into itself and image compression can be studied through this methodology achieving marvelous results.

He has already commercialized his achievements in the form of the top-selling multimedia encyclopedia *Encarta*, published by the Microsoft Corporation including, on one CD-ROM, seven thousand color photographs which may be viewed interactively on a computer screen.

There are diverse images like those of buildings, musical instruments, people's faces, baseball bats, ferns. This development is known as the IFS theory (Iterated Function System theory) for image compression. It has been established that for a fairly large class of images, the IFS theory provides a better compression ratio and quality of images after retrieval (see, for example, Fisher, Lu and recent papers on web resources) compared to the most popular method until now, DCT(JPEG). It is also expected that a combination of fractal and wavelet techniques may still yield better results. This is a fast-growing field in which, besides mathematicians, computer and information scientists, physicists, chemists and engineers are actively involved.

IFS theory

Let (X, d) be a complete metric space and let $T : X \to X$ be a mapping of X into itself. The iterates of T are mappings $T^{0n} : X \to X$, defined by $T^{00}(x) = x$, $T^{01}(x) = T(T^{00}(x)) = T(x)$, $T^{02}(x) = T(T^{01}(x))$ or $T^{02} = T \circ T^{01}$, . . . , $T^{0n}(x) = T(T^{0(n-1)}(x))$ or $T^{0n} = T \circ T^{0(n-1)}$. A mapping F on the set of real numbers R into itself is called an affine transformation if it is of the form $F(x) = ax + b$ for all $x \in R$, where a and b are constants. A mapping G on R^2 into itself is called an *affine transformation* if it can be written as

$$G(x, y) = (ax + by + e, cx + dy + f)$$

where a, b, c, d, e and f are real numbers. G can also be written as

$$G\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} = AX + B,$$

where $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is a 2×2 matrix and B is the column vector $\begin{pmatrix} e \\ f \end{pmatrix}$

The mapping T is called a Lipschitz mapping if there exists a constant $\alpha \geq 0$ such that $d(T(x), T(y)) \leq \alpha d(x, y)$ for all $x, y \in X$. T is called a *contraction mapping* if $0 \leq \alpha < 1$. α is called the contractivity factor of T. A Lipschitz continuous function T is called *eventually contractive* if there is a number n such that T^{0n} is a contraction map. Let Y be a non-empty subset of a compact metric space (X, d), then a mapping S on Y into X is called a *local (partitioned) contraction mapping* on (X, d) if there is a number s, $0 \leq s < 1$, such that $d(S(x), S(y)) \leq sd(x, y)$ for all $x, y \in Y$. A complete metric space (X, d) equipped with n contraction mappings $w_i : X \to X$, $i = 1, 2, \ldots, n$, denoted by $\{X, d, w_i, i = 1, 2, \ldots\}$ is called an *iterated function system (IFS)*. A complete metric space (X, d) equipped with n eventually contractive mappings $w_i : X \to X$, $i = 1, 2, \ldots, n$, denoted by $\{X, d, w_i, i = 1, 2, \ldots, n\}$ is called an *eventually iterated function system*. A local (partitioned) iterated function system (LIFS) is a compact metric space (X, d) equipped with n local contractive mappings $w_i : Y \to X$, $Y \subseteq X$.

A *recurrent iterated function system* (RIFS) is a collection $w_1, w_2, \ldots w_n$ of n Lipschitz maps in a complete metric space X and $n \times n$ matrix (a_{ij}) satisfying $\sum_j a_{ij} = 1$ for all i. The iterated function system (IFS) theorem states that:

For an IFS $\{X, d, w_i\}$ with the contractivity factors α_i, $i = 1, 2, \ldots n$, the mapping W defined on $H(X)$ = space of all compact subsets of X, into itself by $W(B) = \cup_{i=1}^{n} w_i(B)$ is a contraction mapping on the complete metric space $(H(X), h(., .))$, where

$$h(A, B) := \max \{d(A, B), d(B, A)\}$$

$$d(A, B) := \max \{d(x, B) \mid x \in A\}$$

$$d(x, B) := \min \{d(x, y) \mid y \in B)$$

with the contractivity factor $\alpha = \max_{1 \le i \le n} \alpha_i$, that is,

$$h(W(B), (W(C)) \le \alpha h(B, C),$$

for all $B, C \in H(X)$.

Consequently, W has a unique fixed point, say, $A \in H(X)$ which satisfies the relation $A = W(A) = \cup_{i=1}^{n} w_i(A)$, and is given by

$$A = \lim_{n \to \infty} W^{0n}(B) \quad \text{for any} \quad B \in H(X).$$

This fixed point A is called the *attractor* or *deterministic fractal* or *fractal*. $h(., .)$ is known as the Hausdorff metric.

Let $\{R^2, d, w_i, i = 1, 2, \ldots n\}$ be an IFS where w_i's are given by

$$w_i \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e_i \\ f_i \end{pmatrix}, i = 1, 2, \ldots, n$$

Then the following table is known as the IFS code:

IFS code

w	a	b	c	d	e	f	p
w_1	a_1	b_1	c_1	d_1	e_1	f_1	p_1
w_2	a_2	b_2	c_2	d_2	e_2	f_2	p_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
w_n	a_n	b_n	c_n	d_n	e_n	f_n	p_n

where
$$p_i \simeq \frac{|\det A|}{\sum_{i=1}^{N} |\det A_i|} = \frac{|a_i d_i - b_i c_i|}{\sum_{i=1}^{N} |a_i d_i - b_i c_i|},$$

for $i = 1, 2, \ldots n$; the symbol $\simeq$ means "approximately equal to". The numbers p_i's can be interpreted as probabilities for finding the attractor of an IFS using the Chaos Game Algorithm. An image can be treated as a closed bounded (compact) subset of R^2. The following result, known as the ***Collage Theorem***, is very

important for designing IFS whose attactors or fractals are close to a given image. Let (X, d) be a complete metric space and S be a given image, that is, $S \in H(X)$ and let $\varepsilon \geq 0$ be also given. Choose an IFS $\{X, d, w_i\}$, $i = 1, 2, \ldots n$, with contractivity factors α_i, $i = 1, 2, \ldots n$, such that

$$h(S, \cup_{i=1}^{n} w_i(S)) \leq \varepsilon$$

Then
$$h(S, A) \leq \frac{1}{1-s} h(S, W(S)) \leq \frac{\varepsilon}{1-s}$$

where
$$W(S) = \cup_{i=1}^{n} w_i(S) \quad \text{and} \quad s = \max_{1 \leq i \leq n} \alpha_i$$

and A is the attractor of the IFS. This theorem precisely tells us that if we can find an IFS code so that the Hausdorff distance between S and $W(S)$ is very small, the attractor of W will be very close to the target image S. There are algorithms for finding attractors of the given IFS like the Chaos Game Algorithm, the Photocopy Machine Algorithm (see Barnsley and Hurd [4] and Lu [38]). The basic technique of image compression through IFS is to find out appropriate affine contraction mappings $w_1, w_2, \ldots w_n$ such that the condition of the *Collage Theorem* is satisfied, namely, S is very close to $W(S)$ and so, instead of communicating/storing the image, we can communicate/store the fractal or attractor of IFS, that is, coefficients in the IFS code.

Inverse problem for images

The problem of representing a given image (or a function) by the IFSs or their variations is a typical inverse problem. This involves finding the IFS parameters of an image that is exactly generated via an IFS. In recent years, the *iterated function system with probabilities* (IFSP), *iterated fuzzy set systems* (IFZS) and IFS *with gray level maps* (IFSM) have been introduced and the corresponding inverse problems have been investigated. Such an inverse problem is related to the problem of finding the image/function as the fixed point element of a given iteration algorithm of the types IFS, IFSP and IFSM on function spaces like L_p, $H^{m,p}$ (the Sobolev space of order m), the weighted Sobolev space and the Besov space. In purely abstract mathematical terms, it comprises the following steps:

(i) Finding a suitable metric space X in which to represent the image (function).
(ii) Finding an appropriate metric $d(\cdot, \cdot)$ on X.
(iii) Finding an appropriate contraction map T on X into itself.

The fact that such problems have more than one solution motivated the search for different kinds of optimality. This problem has been studied in recent years by Forte and Vrscay (see, for example, Siddiqi, Ahmad, and Mukheimer [62] and Forte and Vrscay [31], Manchanda, Mukheimer and Siddiqi [40]).

Let contraction mappings $w_1, w_2, \ldots, w_n$ be associated with probabilities $p_1, p_2, \ldots, p_n$ with $\sum_{i=1}^{n} p_i = 1$. Furthermore, let $\mathscr{B}(X)$ denote the σ-algebra of Borel subsets of X and $\mathscr{M}(X)$ denote the set of all probability measures on $\mathscr{B}(X)$.

For a measure μ on $\mathscr{B}(X)$ and for any integer $p \geq 1$, let $L_p(X, \mu)$ denote the

vector space of all real-valued functions u such that $|u|^p$ is integrable on $(\mathscr{B}(X), \mu)$. $L_p(X, \mu)$ is a complete metric space with respect to the metric induced by the L_p norm; that is,

$$d(u, v) = \| u - v \|_{L_p} = \left(\int_X | u(x) - v(x) |^p \, d\mu(x) \right)^{1/p}, \quad (L_p(X, \mu), w_i, \Phi)$$

where $\Phi = \{\phi_1, \phi_2, \ldots \phi_n\}$ with ϕ_i: $R \to R$, known as the gray level maps, is called the *iterated function system with gray level maps* (IFSM).

An *operator* T can be defined on IFSM as

$$(Tu)(x) = \sum_{i=1}^{n}{}' \phi_i(u(w_i^{-1}(x))).$$

The prime (′) signifies that the sum operates on all those terms for which $w_i^{-1}(x)$ is defined. If $w_i^{-1}(x) = 0$ for all $i = 1, 2, \ldots, n$ then $(Tu)(x) = 0$. For $X \subset R^n$, let $m^{(n)} \in \mathscr{M}(X)$ denote the Lebesgue measure on $\mathscr{B}(X)$. The indicator function of a subset A of X denoted by $I_A(x)$ is defined by

$$I_A(x) = \begin{cases} 1, x \in A \\ 0, \text{ otherwise} \end{cases}$$

$$\text{Lip}\,(R) = \{\phi\colon R \to R \,\|\, \phi(t_1) - \phi(t_2) \| \le \beta \,|\, t_1 - t_2 \,|\},$$

$\forall\, t_1, t_2 \in R$ and for some $\beta \in [0, \infty)$. It can be verified that for any $u \in L_p(X, \mu)$, $1 \le p < \infty$, and $\phi \in \text{Lip}\,(R)$, $1 \le i \le n$, T is a mapping on $L_p(X, \mu)$ into itself. In fact, T becomes a contraction mapping under certain assumptions and hence has a unique fixed point as $L_p(X, \mu)$ is a complete metric space. Affine IFSM on $L_p(X, \mu)$ is that IFSM on $L_p(X, \mu)$ where ϕ_i are given by $\phi_i(t) = \alpha_i t + \beta_i$, $t \in R$, $i = 1, 2, \ldots n$. Let $X = [0, 1]$ and $\mu = m^{(1)}$ with $w_i(x) = s_i x + a_i$ and $\alpha_i = |s_i| < 1$, $1 \le i \le n$. If T is contractive with fixed point $\bar{u}$, then

$$\bar{u}(x) = \sum_{i=1}^{n}{}' \alpha_i \bar{u}\left(\frac{x - a_i}{s_i}\right) + \beta_i I_{w_i(X)}(x) = \sum_{i=1}^{n}{}' [\alpha_i \psi_i(x) + \beta_i \chi_i(x)]$$

This means that $\bar{u}$ may be expressed as a linear combination of both piece-wise constant functions $\chi_i(x)$ as well as functions $\psi_k(x)$ which are obtained by dilations and translations of $\bar{u}(x)$ and $I_X(x) = 1$, respectively. This reminds us of the role of scaling functions in the wavelet theory.

The Collage Theorem mentioned earlier can be rephrased as follows:
Let (X, d) be a complete metric space, and for a given $x \in X$ there exists a contraction map W: $X \to X$ with the contractivity factor α such that $d(x, W(x)) < \varepsilon$. Then

$$d(x, \bar{x}) < \frac{\varepsilon}{1 - \alpha}$$

where $\bar{x}$ is the fixed point of $W(W(\bar{x}) = \bar{x})$.

In view of this result, the inverse problem for approximation of functions in $L_p(X, \mu)$ by IFSM may be stated as follows: Given a target function $v \in L_p(X, \mu)$ and a $\delta > 0$, find an IFSM $(L_p(X, \mu), w_i, \phi_i)$ with the associated operator T such that $\| v - Tv \|_p \leq \delta$.

For $\mu \in M(X)$, a family $\mathscr{A}$ of subsets $A = \{A_i\}$ of X is called μ-dense in a family $\mathscr{N}$ of subsets B of X if for every $\varepsilon > 0$ and any $B \in \mathscr{N}$, there exists a collection $A \in \mathscr{A}$ such that $A \subseteq B$ and $\mu(A \backslash B) < \varepsilon$. Let $\{w_i\}$ be an infinite sequence of contraction maps on X into itself. We say that $\{w_i\}$ generates a "μ-dense and non-overlapping" (abbreviated as "μ-d-n") family $\mathscr{A}$ of subsets of X if, for every $\varepsilon > 0$ and every $B \subseteq X$, there exists a finite set of integers $i_k \geq 1$, $1 \leq k \leq n$, such that

(i) $A = \cup_{k=1}^{n} w_{i_k}(X) \subseteq B$,

(ii) $\mu(B/A) < \varepsilon$, and

(iii) $\mu(w_{i_k}(X) \cap w_{i_l}(X)) = 0$ if $k \neq l$.

If $\{w_i\}$ satisfies the above conditions on (X, d), then $\inf_{1 \leq i < \infty} \{\alpha_i\} = 0$, where α_i's are the contractivity factors of w_i's, independent of μ. If $X = [0, 1]$ and μ is the Lebesgue measure, then the wavelet type functions

$$w_{ij}(x) = 2^{-i}(x + j - 1), \quad i = 1, 2, \ldots, j = 1, 2, \ldots, 2^i$$

can form a "μ-d-n" family.

For each $i^* \geq 1$, the set of maps $\{w_{i^*j}, \ j = 1, 2, \ldots, 2^{i^*}\}$ provides a set of 2^{-i^*} contractions of $[0, 1]$ which tile $[0, 1]$. In 1995, Forte and Vrscay obtained the following result which provided the solution of the inverse problem:

Theorem A [Forte-Vrscay, 1995] Let $v \in L_p(X, \mu)$, $1 \leq p < \infty$, then

$$\liminf_{n \to \infty} \| v - T^n v \|_p = 0$$

provided the sequence of contraction maps w_i generates a "μ-d-n" family $\mathscr{A}$ of subsets of X and w_i's are also one-to-one, where

$$(T^n v)(x) = \sum_{i=1}^{n} \phi_i(u(w_i^{-1}(x)))$$

This theorem has been studied for local IFSM and special cases like $p = 2$ and ϕ_i's affine maps. Forte and Vrscay have also carried out an approximation of the target image "Lena", a 512×512-pixel gray scale image, with each fixed pixel having 256 possible values (8 bits, with values from 0 to 255, which are rescaled to values in $[0, 1]$). This type of approximation has been studied by Siddiqi et al. [62] for singer and bride. The correspondence between fractal-wavelet transforms and iterated function systems with gray-level maps has been systematically studied by Mendivil and Vrscay. A wavelet-based solution to the inverse problem for fractal interpolation functions has been investigated by Berkner (see, for example reference [53]). Manchanda, Mukheimer and Siddiqi [40] have extended Theorem A to the Besov space as follows:

For any $h \in R^2$ and Ω, a bounded smooth domain of R^2, we define

$$\Delta^0 hf(x) = f(x), \quad f: \Omega \to R$$

and
$$\Delta_h^{k+1} f(x) = \Delta_h^k f(x+h) - \Delta_h^k f(x), \quad k = 0, 1, 2, \ldots$$

For $r > 0$, $\Delta_h^r f(x)$ is defined for $x \in I_{rh} = \{x \in \Omega \mid x + rh \in \Omega\}$. The $L_p(\Omega)$ modulus of smoothness, $0 < p \leq \infty$ is defined as

$$\omega_r(f, t)_p = \sup_{|h| \leq t} \left(\int_{I_{rh}} |\Delta_h^r f(x)|^p \, dx \right)^{1/p}$$

with the usual change to an essential supremum when $p = \infty$. Given $\alpha > 0$, $0 < p \leq \infty$ and $0 < q \leq \infty$, choose $r \in Z$ with $r > \alpha \geq r - 1$. Then the Besov space seminorm is defined as

$$|f|_{B_q^\alpha(L_p(\Omega))} = \left(\int_0^\infty [t^{-\alpha} \omega_r(f, t)_p]^q \frac{dt}{t} \right)^{1/q}$$

again with a supremum when $q = \infty$.

The Besov space norm is

$$\|f\|_{B_q^\alpha(L_p(\Omega))} = |f|_{B_q^\alpha(L_p(\Omega))} + \|f\|_{L_p(\Omega)}$$

Thus the family of Besov spaces $B_q^\alpha(L_p(\Omega))$, $0 < \alpha < \infty$, $0 < p \leq \infty$, and $0 < q \leq \infty$ denote the set of those functions $f \in L_p(\Omega)$ such that

$$\|f\|_{B_q^\alpha(L_p(\Omega))} < \infty$$

Special cases yield familiar spaces, for example, $p = q = 2$, $B_2^\alpha(L_2(\Omega))$ is nothing but the Sobolev space $H^\alpha(L_2(\Omega))$. When $\alpha < 1$, $1 \leq p \leq \infty$, and $q = \infty$, $B_p^\alpha(\Omega)$ is the Lipschitz space Lip $(\alpha, L_p(\Omega))$. When $p < 1$ or $q < 1$, these spaces are not Banach spaces, but rather complete quasinormed linear spaces, that is, the triangle inequality may not hold, but for each space $B_q^\alpha(L_p(\Omega))$, there exists a constant C such that for all f and g in $B_q^\alpha(L_p(\Omega))$, we have

$$\|f + g\|_{B_q^\alpha(L_p(\Omega))} \leq C(\|f\|_{B_q^\alpha(L_p(\Omega))} + \|g\|_{B_q^\alpha(L_p(\Omega))})$$

$$\text{Con}(X) = \{ f: X \to X \mid d(f(x), f(y)) \leq cd(x, y), \quad c \in [0, 1]\}$$

$$\text{Con}_1(X) = \{ f \in \text{Con}(X) \mid f \text{ is one-one}\}$$

$$\text{Sim}(X) = \{ f: X \to X \mid d(f(x), f(y)) = cd(x, y), \quad c \in [0, 1)\}$$

$$\text{Lip}(Y) = \{\varphi: Y \to Y, Y \subseteq R \mid \varphi(t_1) - \varphi(t_2) | \leq k \mid t_1 - t_2 | \quad \forall\, t_1, t_2\}$$

Let μ be a finite measure on σ-algebra $M(X)$. A family $\mathscr{A}$ of subsets of X is called μ-dense in a family $\mathscr{B}$ of subsets of X if for every $\varepsilon > 0$ and any $B \in \mathscr{B}$, there exists a collection $A \in \mathscr{A}$ such that $A \subseteq B$ and $\mu(B \backslash A) < \varepsilon$.

Let $\mathscr{W} = \{w_1, w_2, \ldots\}$, $w_i \in \text{Con}(X)$ be an infinite set of contraction maps on X. We say that $\mathscr{W}$ generates a "μ-dense and non-overlapping" ("μ-d-n", for short) family $\mathscr{A}$ of X if for every $\varepsilon > 0$ and every $B \subseteq X$, there exists a finite set of integers $i_k \geq 1$, $1 \leq k \leq N$, such that

(i) $A = \bigcup_{k=1}^{B} w_{i_k}(X) \subseteq B$,

(ii) $\mu(B \backslash A) < \varepsilon$, and

(iii) $\mu(w_{i_k}(X) \cap w_{i_k}(X)) = 0$ if $k \neq \varepsilon$.

As mentioned in [31], a useful set of affine maps satisfying such a condition on $X = [0, 1]$ with respect to Lebesgue measure is given by the following "wavelet type" functions:

$$w_{ij}(x) = 2^{-i}(x + j - 1), \quad i = 1, 2, \ldots, j = 1, 2, 3, \ldots, 2^i$$

It may be observed that in practice X is taken as a subset of R^2.

Now let $\mathscr{W} = \{w_1, w_2, w_3, \ldots\}$, $w_i \in \text{Con}_1(X)$ be an infinite set of one-to-one contraction maps on X satisfying the μ-d-n property. Also, let

$$w^N = \{w_1, w_2, \ldots, w_N\}, \quad N = 1, 2, \ldots$$

denote N-map truncations of $\mathscr{W}$. For each $N \geq 1$, let

$$\Phi^N = \{\varphi_1, \varphi_2, \varphi_3, \ldots, \varphi_N\}$$

denote an associated N-vector of gray-level maps with the restriction that the $\varphi_i \in \text{Lip}(R)$. Let T^N denote the operator associated with the N-map IFSM (w^N, Φ^N). We look for a solution of the inverse problem, namely, for a given target function or image $v \in B^{\alpha}_{p,q}(L_p(\Omega))$ and an $\varepsilon > 0$, find an N-map IFSM (w^N, Φ^N) with the operator T^N such that

$$\| v - T^N v \|_{B^{\alpha}_{p,q}(L_p(\Omega))} < \varepsilon$$

More precisely, we prove the following theorem:

Theorem 4.1 For appropriate small values of α and $v \in B^{\alpha}_{p,q}(L_p(\Omega))$ we can find an IFS with infinite set of maps $\mathscr{W} = \{w_1, w_2, \ldots\}$, $w_i \in \text{Con}_1(X) \subset R^n$ generating a μ-d-n family λ of subsets of X such that

$$\lim_{N \to \infty} \inf \| v - T^N v \|_{B^{\alpha}_{p,q}(L_p(\Omega))} = 0$$

where $T^n v = \sum_{k=0}^{N} \varphi_k(t) w_k^{-1}(x))$

$$\varphi_k \colon R \to R, \ \varphi_k \in \text{Lip}(Y), \ \sigma_p < \alpha, \ \alpha < 1/p, \ \sigma/p = n \cdot \max\left(0, \frac{1}{p} - 1\right),$$

$$\varphi'_k \in L_\infty(R)$$

In the proof, we require the following lemmas (for complete proof, see Manchanda, Mukheimer and Siddiqi [40]).

It may be observed that the range of α is related to the smoothness of φ_k. A fairly large number of relevant results concerning decomposition operators on Besov spaces $B^{\alpha}_{p,q}$ for different classes of functions and different ranges of α are discussed in [57].

In the subsequent discussion, $A_1 \hookrightarrow A_2$ will denote continuous embedding, that is, there exists a constant c such that $\| a \|_{A_2} \leq c \| a \|_{A_1}$. We say that $B^{\alpha}_{p,q}$ is supercritical if $B^{\alpha}_{p,q} \hookrightarrow L_{\infty}$, and subcritical if $B^{\alpha}_{p,q} \not\hookrightarrow L_{\infty}$. For a continuous $G: R \to R$ and Lebesgue measurable function f the operator $T_G: f \to G(f)$ is called a composition operator.

Lemma 1 Let T_G be a composition operator, $G(0) = 0$, and $\sigma_p < \alpha < 1$, then T_G maps $B_{p,q}(\Omega)$, where Ω is a non-trivial bounded smooth domain, into itself if and only if either $G' \in L^{\mathrm{Loc}}_{\infty}(R)$ if $B^{\alpha}_{p,q}$ is supercritical or $G' \in L_{\infty}(R)$ if $B^{\alpha}_{p,q}$ is subcritical.

This is a consequence of Theorem I(ii) on page 335 and Remark 3 on page 267 in [57].

Lemma 2 [57, p. 355] Let $1 \leq p \leq \infty$, $0 < q \leq \infty$ and $0 < s < 1 + \frac{1}{p}$ and $S_G(f) = |f|$. Then S_G maps $B^{\alpha}_{p,q}$ into itself.

Lemma 3 [57, p. 200] Let $\Omega \subset R^n$. Suppose that $\alpha > \sigma_p$. Then the characteristic function of $\Omega \cdot \gamma_{\Omega}$ belongs to $B^{\alpha}_{p,q}$ if and only if either $\alpha < \frac{1}{p}$ or $\alpha = \frac{1}{p}$ and $q = \infty$.

Lemma 4 Let $\varphi_i(t) = \xi_i$, where $\xi_i \in R$, $1 \leq i \leq N$. Then for any $p \in [1, \infty)$ and $\mu \in M(X)$, the associated operator T^N is contractive on $B^{\alpha}_{p,q}$ with contractive factor $c = 0$ provided $\alpha < 1/p$ or $\alpha = 1/p$ and $q = \infty$. Furthermore, the fixed point $\bar{u}$ of T^N is given by

$$\bar{u}(x) = \sum_{i=1}^{N} \xi_i \chi_{w_i(X)}(x), \quad x \in X,$$

where $\chi_{w_i(X)}$ is the characteristic function of $w_i(X)$.

Lemma 5 Let $X \subset R^n$, $n \in \{1, 2, 3 \ldots\}$ and μ be the Lebesgue measure $m^{(n)}$. Let $(\mathscr{W}, \Phi)$, $\mathscr{W} = \{w_1, w_2, \ldots, w_N\}$, $\Phi = \{\varphi_1, \varphi_2, \ldots, \varphi_N\}$ on N-map IFSM such that

(i) $w_k \in \mathrm{Sim}_1(X) = \mathrm{Sim}(X) \cap \mathrm{Con}_1(X)$,
(ii) $\varphi_k \in \mathrm{Lip}(R)$, $1 \leq k \leq N$, and
(iii) $\varphi'_k \in L_{\infty}(R)$.

Then for a $p \in [1, \infty)$, $\sigma_p < \alpha$, $\alpha < 1/p$ or $\alpha = 1/p$ and $q = \infty$

$$\| T^N u - T^N v \|_{B^\alpha_{p,q}} \leq \lambda \| u - v \|_{B^\alpha_{p,q}}$$

where λ is a positive constant.

We note that λ may be less than 1 for an appropriate choice of constants in Lip (R). Sim_1 (X) and while applying Minkowski's inequality up to a constant. In this case T^N has a unique fixed point by the Banach contraction principle and this fixed point is given by $\bar{u}_n = \sum_{i=1}^{N} \xi_k \chi_{w_i(X)}(x)$ if we choose $\lambda = 0$, Lemma 4.

This is precisely the case when $\varphi_i(t) = \xi_i$ for all t.

In view of Lemma 2, Lemma 5 is valid without condition (iii) if $\varphi_k(t) = | t | \ \forall\ k$.

5. Miscellaneous (Image Classification, Watermarking and Data Analysis)

5.1 Image Classification

The choice of a compression technique for a given image is a difficult task. To facilitate it, Siddiqi and Ahmad [69] have studied the concept of sharp operator and have defined information function which measures oscillatory behaviour of images. Based on information function images are classified to adapt a specific compression technique. A comparative study of compression technique for different class of images has been carried out in their paper. The classification based on information function helps us in choosing a proper compression technique for a certain class of images. Wavelet image compression (WIC), fractal image compression (FIC) and JPEG techniques have been used in this study.

A few relevant concepts are mentioned here.

Definition 5.1 Let $\mathbb{R}^n$ be the n-dimensional Euclidean space and $f(x)$ a real-valued measurable function on $\mathbb{R}^n$. For such a function f on $\mathbb{R}^n$, its Hardy-Littlewood maximal function is defined by the formula

$$Mf(x) = \sup_Q \left\{ \frac{1}{\lambda(Q)} \int_Q | f(y) | \, dy : Q \subset \mathbb{R}^n, \ \ x \in Q \right\}, \tag{5.1}$$

where the supremum ranges over all finite cubes Q in $\mathbb{R}^n$ and $\lambda(Q)$ is the Lebesgue measure of Q.

The function $Mf(x)$ has the following properties:

(i) $0 \leq Mf(x) \leq \infty$,
(ii) $M(f + g)(x) \leq Mf(x) + Mg(x)$
(iii) $M(\alpha f)(x) = | \alpha | \, Mf(x)$,

where f, g are measurable functions on $\mathbb{R}^n$ and α is some scalar quantity.

It is easy to find a function whose maximal function is unbounded.

Theorem 5.1 For each function $f \in L_1(\mathbb{R}^n)$, we have

$$\lambda(\{x: Mf(x) > t\}) \leq 6^n t^{-1} \| f \|_1, \quad t > 0.$$

Definition 5.2 A measurable function f on $\mathbb{R}^n$ has bounded p-mean oscillation, $1 \leq p < \infty$, if

$$\| f \|_{BMO_p} = \sup_Q \left(\frac{1}{\lambda(Q)} \int_Q | f(x) - f_Q |^p \, dx \right)^{1/p} < \infty, \tag{5.2}$$

where the supremum ranges over all finite cubes Q in $\mathbb{R}^n$ and $f_Q = \frac{1}{\lambda(Q)} \int_Q f(x)\, dx$ is the mean value of the function f on the cube Q.

The set of all functions of bounded p-mean oscillation is denoted by $BMO_p(\mathbb{R}^n)$. $\| f \|_{BMO_p}$ is "almost" a norm since it has the following properties:

(i) $\| f + g \|_{BMO_p} \leq \| f \|_{BMO_p} + \| g \|_{BMO_p}$

(ii) $\| \alpha f \|_{BMO_p} = | \alpha | \cdot \| f \|_{BMO_p}$

(iii) $\| f \|_{BMO_p} = 0$ if and only if $f =$ constant almost everywhere,

where f, g are measurable functions on $\mathbb{R}^n$ and α is some scalar quantity.

If we define

$$\| f \|_{BMO'_p} = \left| \int_{\mathbb{R}^n} f(x)\, dx \right| + \sup_Q \left(\frac{1}{\lambda(Q)} \int_Q | f(x) - f_Q |^p \, dx \right)^{1/p}$$

we get a norm of f and BMO'_p becomes a Banach space. On the other hand, we can say, $\| f \|_{BMO_p}$ becomes a norm if we identify functions which differ by a constant. With this identification $BMO_p(\mathbb{R}^n)$ becomes a normed space, and ultimately a Banach space.

Fefferman and Stein introduced "sharp function" $f^{\#}$ that mediates between BMO_p and L_p spaces. It is defined as follows:

Definition 5.3 Let f be a locally integrable function on $\mathbb{R}^n$. The sharp function $f^{\#}(x)$ is represented by the formula

$$f^{\#}(x) = \sup_{Q: x \in Q} \left(\frac{1}{\lambda(Q)} \int_Q | f(y) - f_Q |^p \, dy \right)^{1/p} \tag{5.3}$$

Of course, $f \in BMO_p$ is identical with $f^{\#} \in L_\infty$. It is also observed that there are unbounded functions in $BMO_p(R)$.

The function $f(x) = \ln | x |$ on $\mathbb{R}$ is in $BMO_1(\mathbb{R})$.

After calculation it comes out to be $\| \ln | x | \|_{BMO_1} \leq 2$. So, the unbounded function $\ln | x |$ is in $BMO_1(\mathbb{R})$.

It is important to note that it does not matter in which L_p norm we measure the oscillation. It is clear from the following corollary.

Corollary 5.1 For each p, $1 \leq p \leq \infty$, there exists a constant C_p such that for each $f \in BMO_p(\mathbb{R}^n)$, we have

$$\|f\|_{BMO_1} \leq \|f\|_{BMO_p} \leq C_p \|f\|_{BMO_1}$$

In view of the above corollary, the spaces $BMO_p(\mathbb{R}^n)$ are equivalent for all p, $1 \leq p \leq \infty$.

Information Function

It is clear from the definition of the sharp function that for a pixel z in an almost uniform gray level region in an image, $f^{\#}(z)$ will be very small. However, for the contrast region we get large $f^{\#}(z)$ values. We count the number of pixels that lie in a small interval of $f^{\#}(z)$ range. The histogram of intervals of $f^{\#}(z)$ range versus number of pixels are plotted. We name this histogram of an image f as information function of the image, defined by In (f).

Definition 5.4 The information function of an image f is a piecewise constant function defined as follows:

$$\text{In}\,(f)(x) = \lambda(\{z \in \mathbb{R}: f^{\#}(z) \in [t\eta, (t+1)\eta)\}) \text{ if } x \in [t\eta, (t+1)\eta) \quad (5.4)$$

for $t = 0, 1, \ldots, n \in \mathbb{Z}_0^+$ (the set of positive integers with zero). Here, η is a fixed, positive number and λ is the Lebesgue measure.

In other words, if $\mathbb{R}$ is the range of $f^{\#}$, divide $\mathbb{R}$ equidistantly into n subintervals, i.e., $\mathbb{R} = \bigcup_{t=0}^{n} I_t = [t\eta, (t+1)\eta)$, η a fixed, positive number, then

$$\text{In}\,(f)(x) = \lambda(f^{\#^{-1}}(I_t)) \text{ for } x \in I_t$$

$f^{\#^{-1}}$ denotes the pre-image.

The choice of η is very important. In our study on various types of images we found that, e.g., $f^{\#}(z) \in [0, 70]$, so we partitioned this interval into 250 subintervals and fixed $\eta = 70/250$ (length of the subintervals). η cannot be an arbitrary small number, then there would not be any pixel in some subintervals and the very purpose of information function will become irrelevant. See [53 and 69] for more details.

5.2 Watermarking Technology for Multimedia

Multimedia watermarking technology has evolved very quickly during the last few years. A digital watermark is information that is imperceptibly and robustly embedded in the host data in a way that cannot be removed. A watermark typically contains information about the specifications of the data. This is essential for copyright protection, authenticity of the data, data monitoring and tracking. In a series of papers, the basic concepts of watermarking systems are investigated and watermarking methods for images, video, audio, text documents, and other media are proposed. Discrete Cosine Transform, Discrete Water Transform and Fractals are mathematical tools in these studies. This is an emerging area of

information technology, a comprehensive account of which can be found in references [77, 78].

5.3 Data Analysis

The analysis of experimental data that have been observed at different points of time leads to new and unique problems in modeling and inference. The systematic approach by which one resolves mathematical and statistical questions involved in the analysis is commonly known as time series analysis. The impact by time series analysis has been realized in diverse fields like economics, where we are continually exposed to daily stock market quotations; social sciences, where we encounter birth rates and school enrollments; medical sciences, where one may be interested in the number of influenza cases observed over some time period, blood pressure measurements traced over time or ECG recorded time series or functional magnetic resonance imaging of brain-wave time series. More sophisticated applications of time series analysis have been found in the physical and environmental sciences such as in warming trend in global temperature measurements, influence of level of pollution in mortality rate in a city, and geophysical time series which are produced by yearly depositions of various kinds having long range proxies for temperature, rainfall and wind speed. There are two, separate but not necessarily exclusive, approaches for time series analysis, known as the time domain approach and the frequency domain approach; for an updated literature see Robert H. Shumway and David S. Stoffer [59].

As discussed above, wavelet analysis, invented formally in eighties, is a rapidly developing area of mathematical and engineering sciences. The properties which establish its superiority over Fourier analysis are time-frequency localization, multirate filtering, scale-space analysis and high compression ratio. In view of this, wavelet analysis is being applied in geophysical process like atmospheric turbulence, space-time rainfall, ocean wind waves, seafloor bathmyetry, geological layered structures, climate change. For a comprehensive and updated references, we refer to Efi Foufoula-Georgioù and Praveen Kumar [29] and Percival and Walden [55].

In a series of papers, Siddiqi, Manchanda, Aslan, Tokgozlu, Khan and Rehman have analyzed different sets of data by wavelet methods [41, 67, 71, 74, 81, 82]. It may be observed that a forthcoming book by Gencay and Seluk [32] is devoted to wavelets in Finance and Economics.

6. Concluding Remarks

Meyer [47] has emphasized the importance of properties of wavelet coefficients of functions of bounded variation in R^2 (functions belonging to $BV(R^2)$ [22] in the context of image processing. Meyer [45], Coifman, Meyer and Wickerhauser [23], Wickerhauser [88, 89], Mallat [39] and Siddiqi [70] have introduced and studied the concept of wavelet packet (superposition of wavelets). Walsh function is the first example of wavelet packet as Haar is that of wavelet. Many results which are known now for wavelets were first proved for Haar wavelet. There exists a vast literature on Walsh functions, see for example Siddiqi [60], Maqusi

[43], Schipp, Wade and Simon [58], Golubov, Effimov and Skvortsov [34], Wade [85, 86] and references therein. Wade [85] has explained in a nice way the motivation for studying Walsh-Fourier Analysis. He also provides updated literature. We [65, 68] are at present examining how far the results of Cohen, DeVore, Petrushev and Xu [22] can be extended for Walsh function, in particular and wavelet packet, in general, for the space of functions with p-generalized bounded fluctuation and BMO (see Wade for references) and Schipp, Wade and Simon [58].

It has been observed [16] that problems in image processing are closely related to approximation in Besov spaces. Besov spaces in the context of the dyadic group have been investigated by Tateoka [79] and Tateoka and Wade [80]. It is worthwhile to examine to what extent the results of [16] can be put in the framework of dyadic Besov spaces.

Daubechies [25] discussed characterization of function spaces L_p, $1 < p < \infty$, $H^s(R)$ (Sobolev space), $C^s(R)$ (The Hölder spaces) and Z (Zygmund's class); See also Meyer [44] and Siddiqi [70]. Meyer [47] has explained that the space BV cannot be characterized by size properties on wavelet coefficients. Siddiqi, J.A. [75, 76] and Siddiqi, A.H. [72] have proved, respectively, characterization of a function in BV in terms of the $(C, 1)$ summability for Fourier coefficients and Walsh-Fourier coefficients. It is pertinent to examine whether such results hold for certain wavelet or wavelet packet coefficients.

A famous result of Hardy concerning Fourier coefficients states that if $\{a_k\}$ is a sequence of Fourier coefficients of a function belonging to L_p, $1 < p < \infty$, then $(C, 1)$ mean of $\{a_k\}$ is also Fourier coefficient of a function belonging to L_p. The converse of this problem was investigated by Siddiqi [73] and Izumi, S. and Masaka, M. [36]. Richard Bellman and Gunther Goes, and many others have worked on this problem. Such results for Walsh-Fourier coefficients were studied by Siddiqi [61]. One may investigate such properties for wavelet series.

As discussed by Donoho [27, 28], Candes [13] and Donoho and Candes [10–12] have introduced the concepts of ridgelets and curvelets to represent edges in images. As pointed out by them, wavelets can deal with points-like phenomena very effectively but are ineffective in dealing with line-like phenomena in dimension 2, and plane-like phenomena in dimension 3. Recently, Siddiqi [66] has studied ridgelet packets and has obtained certain results concerning the size of the ridgelet packet coefficients of functions belonging to $BV(R^2)$ and Besov spaces.

In this article, we have presented a glimpse of the theme which may attract the attention of researchers in different disciplines during the 21st century.

REFERENCES

1. Alvarez, L., Lions, P.L. and Morel, J.M.: Image selective smoothing and edge detection by nonlinear diffusion II, SIAM J. Numer. Anal., 29(1992), 845–866.
2. Barnsley, M.F.: Fractal Everywhere, Academic Press, SanDiego, 1988.
3. Barnsley, M.F.: Fractal Image Compression, Notices of the AMS, 43(1996), 657–662.

4. Barnsley, M.F., and Hurd, L.P.: Fractal Image Compression, AK Peters Ltd., Boston, 1993.
5. Bonami, A. and Poornima, S.: Non-multipliers of Sobolev spaces, J. Funct. Anal., 71(1977),175–181.
6. Bourdin, B.: Image segmentation with a finite element method, M2AN Math. Model. Numer. Anal., 33(2) (1999), 229–244.
7. Bourdin B. and Chambolle: Implementation of a finite elements approximation of the Mumford-Shah functional, Numer. Math., 85(4) (2000), 609–646.
8. Braides, A.: Approximation of Free Discontinuity Problems, Number 1694, Lecture Notes, Springer-Verlag, 1998.
9. Brislawn, Christopher, M.: Fingerprints go digital, Notices of the AMS,42(11) (1995), 1278–1283. Website http://www.c3.lanl.gov/brislawn.
10. Cande's, E.S. and Donoho, D.L.: Ridgelets: Key to higher-dimensional intermittency, Phil. Trans. R. Soc. London, A, 357(1999), 2495–2509.
11. Cande's E.S. and Donho, D.L.: Curvelets, Manuscript, http://www-stat.stanford.edu/~Reports/1998/curvelets.zip, 1999.
12. Cande's, E.J. and Donoho, D.L.: Curvelets-a surprisingly effective nonadaptive representation for objects with edges in curves and surfaces, L.L.S. et al., ed. Nashville, TN, (Vanderbilt University Press), 1999 (see Collected Lecture Notes on Multiscale and Multiresolution Methods, Yosemite National Park, Oct. 29-Nov. 1, 2000, Sponsored by: NSF, NASA, IBM, SGI, and MIT).
13. Cande's E.J.: Ph.D Thesis, Stanford University, 1998.
14. Chambolle, A.: Finite differences discretizations of the Mumford-Shah functional, M2AN Math. Model. Numer. Anal., 33(2) (1999), 261–288.
15. Chambolle, A.: Inverse problems in image processing and segmentation: Some mathematical and numerical aspects, Lect. Notes, School on Math Problems in Image Processing, the Abdus Salam, ICTP, Trieste, Italy, # SMR25711, 64 pages, September 2000.
16. Chambolle, A., DeVore, R.A., Lee, N.Y. and Lucier, B.: Nonlinear wavelet image processing, variational problems, compression and noise removal through wavelet shrinkage, IEEE Trans. Image Processing, 7(1998), 319–335.
17. Chambolle, A. and DalMaso, G.: Discrete approximation of the Mumford-Shah functional in dimension 2, M2AN Math. Model. Numer. Anal., 33(4) (1999), 651–672.
18. Chambolle, A. and Lions, P.L.: Image recovery via total variation minimization and related problems, Numer. Math., 75(2) (1997), 167–188.
19. Chan, T.F., Golub, G.H. and Mulet, P.: A nonlinear primal dual method for total variation based image restoration, SIAM J. Sci. Comput., 20(6) (1999), 1964–1977.
20. Chipot, M., March, R. and Vitulano, D.: Numerical Analysis of Oscillations in a Noncovex Problem Related to Image Selective Smoothing, Preprint 1999, CNR, Rome, Italy.
21. Cohen, A.: Numerical analysis of wavelet methods, in "Handbook of Numerical Analysis" (Ciarlet, P.G. and Lions, J.L. Eds.), 1999.
22. Cohen, A. et al.: Nonlinear approximation and the space $BV(R^2)$, American J. Math., 121(1999), 587–628.
23. Coifman, R.R., Meyer, Y. and Wickerhauser, M.V.: Wavelet analysis and signal processing, in "Wavelets and their Applications". (M.B. Ruskai et al., Eds), pp. 153–178, Jones and Bartlett Publishers, Boston, 1992.
24. DalMaso, G., Morel, J.M., and Solimini, S.: A variational method in image segmentation: Existence and approximation results, Acta Math., 68(1992), 89–151.

25. Daubechies, I.: Ten Lectures on Wavelets, SIAM, Philadelphia, 1992.
26. DeVore: Nonlinear approximation, Acta Numerica, 7(1998), 51–150.
27. Donoho, David L.: Orthogonal ridgelets and linear singularities, SIAM J. Math. Anal., 31(2000), 1062–1099.
28. Donoho, David L.: Tight frames of k-plane ridgelets and the problem of representing objects that are smooth away from d-dimensional singularities, Proce. Natl. Acad. Sci., USA. 96(1999), 1828–1833.
29. Efi Foufoula-Georgio and Kumar, P. (Eds.): Wavelets in Geophysics, Academic Press, 1994.
30. Falconer, K.J.: The Geometry of Fractal Sets, Cambridge University Press, 1985.
31. Forte, B. and Vrscay, E.R.: Solving the inverse problem for function and image approximation using iterated function systems, Dynamics of Continuous, Discrete and Impulsive Systems, 1(1995), 177–231.
32. Gencay, Ramazan, and Seluk, Faruk: An Introduction to Wavelets and other Filtering Methods in Finance and Economics, Academic Press, 2001.
33. Gonzalez, R.C. and Woods, R.E.: Digital Image Processing, Addison-Wesley Publishing Comp., Reading, Massachusetts, 1993.
34. Golubov, B., Effimov, A. and Skvortsow, V.: Wash Series and Transform, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1990 (See also Bulletin (New Series) of the American Mathematical Society Vol. 26(2), 1992, pp. 348–459 Book Review).
35. Guichard, F.: Image iterative smoothing and P.D.E.'s, School on Mathematical Problems in Image Processing, 4–22 September 2000, Abdus Salam, ICTP, Trieste, SMR 1237/2.
36. Izumi, S. and Masaka, M.: Hardy's Fourier coefficients, Proc. Japan Academy of Sciences, 44(1968), 418–423.
37. Kahane, J.P., and Lemarié-Rieusset, P.G.: Fourier Series and Wavelets, Gordon and Breach Science Publishers, 1996.
38. Lu, Ning: Fractal Imaging, Academic Press, San Deigo, London, 1997.
39. Mallat, S.: A Wavelet Tour of Signal Processing, Academic Press, 1998.
40. Manchanda, P., Mukheimer, A.A.S. and Siddiqi, A.H.: Certain results concerning the iterative function system, Numer. Funct. Anal. and Optimiz., 21(1 and 2) (2000), 217–225.
41. Manchanda, P.: Option pricing in India, Preprint (to be published).
42. Manchanda, P. et al: Pointwise convergence of wavelet expansions associated with dilation matrix, Applicable Analysis, Gordon and Breach Science Publisher, Vol. 76(3–4), 301–308.
43. Maqusi, Mohammad: Applied Walsh Analysis, Heyden and Sons Ltd., London-Philadelphia-Rheine, 1981.
44. Meyer, Y.: Wavelets and operators, Cambridge University Studies in Advanced Mathematics, 37 Cup (1992).
45. Meyer, Y.: Wavelets, Algorithm and Applications, Philadelphia, 1993.
46. Meyer, Y.: Wavelets, Vibrations and Scalings, CRM Monograph Series, Vol. 9(1998).
47. Meyer, Y.: Lecture Notes, School on Mathematical Problems in Image Processing, 4–22, Sept. 2000, the Abdus Salam International Centre of Theoretical Physics, Trieste, Italy, SMR 1237/3 and 4 (40 pagers and 35 pages).
48. Morel, J.M., Solimini, S.: Variational Methods in Image Segmentation, Birkhauser, Boston, 1995.
49. Moricz, F., Schipp, F. and Wade, W.R.: On the integrability of double Walsh series with special coefficients, Michigan Math. J., 37 (1990), 191–201.
50. Moricz, F. and Siddiqi, A.H.: Approximation by Nörlund means of Walsh-Fourier series, J. Approximation Theory, 70(3) (1992), 375–389.

51. Mumford, D. and Shah, J.: Boundary detection by piecewise minimizing functionals, I., In Proc. IEEE Conf. on Computer Vision and Pattern Recognition, San Francisco, 1985.
52. Mumford, D. and Shah, J.: Optimal approximation by piecewise smooth functions and associated variational problems, Comm. Pure. Appl. Math., 42(1989), 577–685.
53. Neunzert, H. and Siddiqi, A.H.: Topics in Industrial Mathematics—Case Studies and Related Methods, Kluwer Academic Publishers, Boston-Dordrecht-London, 2000.
54. Osher, S. and Rudin, L.: Total variation based image restoration with free local constraints. In Proc. IEEE, ICIP, Vol. 1, pp. 31–35, Austin (Texas), USA. 1994.
55. Percival, B. and Walden, A.T.: Wavelet Methods for Time Series Analysis, Cambridge University Press, 2000.
56. Rudin, L. Osher, S.J. and Fatemi: Nonlinear total variation based noise removal algorithms, Physica D., 60(1992), 259–279.
57. Runst, T. and Sickel, W.: Sobolev Spaces of Fractional Order, Nemytski Operators and Nonlinear Partial Differential Equations, Walter de Gruyter, Berlin, 1996.
58. Schipp, F., Wade, W.R. and Simon, P.: Walsh Series: An Introduction to Dyadic Harmonic Analysis, Adam Hilger Publ. Ltd., Bristol. New York, 1990.
59. Shumway, R.H. and Stoffer, D.S.: Time Series Analysis and its Applications, Springer-Verlag, 2000.
60. Siddiqi, A.H.: Walsh Functions, AMU Publication, AMU Press, Aligarh, 1978.
61. Siddiqi, A.H.: Approximation by Walsh series, in Mazhar et al. (eds.), Proc. Int. Conference on Math. Analysis and its Applications, Kuwait 1985, pp. 43–61, Pergamon Press, Oxford, New York, MR No. 90d: 42029.
62. Siddiqi, A.H. et al.: Certain recent developments in fractal image compression, in "Functional Analysis with Applications" (M. Brokate and A.H. Siddiqi, Eds.), pp. 177–191, Pitman Research Notes Series, Vol. 377(1998).
63. Siddiqi, A.H., S. Khan and S. Rehman: Wavelet based computer simulation for wind speed in Saudi Arabia, Proc. First Saudi Science Conference, 9–11 April 2001, KFUPM, Dharan, pp. 313–326.
64. Siddiqi, A.H. and Manchanda, P.: Approximation by Nörlund Means of Walsh-Fourier Series in Besov Spaces, Preprint, March 2001.
65. Siddiqi, A.H. and Manchanda, P.: Approximation of Functions in $BV(R^2)$ by the Walsh-Fourier Series, April 2001.
66. Siddiqi, A.H., Size of Ridgelet Packet Coefficients, Preprint 2001.
67. Siddiqi, A.H. and Khan, S.: Wavelet Methods for Time Series Analysis of Saudi Petroleum Production Data Set, Abstract p. 327, Proc. First Saudi Conference, 9–11 April, 2001, KFUPM, Dhahran, Saudi Arabia.
68. Siddiqi, A.H. and Manchanda, P.: Wavelet Packet Coefficients of Functions of Bounded Variation, Preprint 2001.
69. Siddiqi, A.H. and Ahmad, M.K.: Image Classification and Comparative Study of Compression Techniques, Preprint, 1999.
70. Siddiqi, A.H.: Certain problems concerning wavelets and wavelet packets, in "Industrial and Applied Mathematics" (A.H. Siddiqi and K. Ahmad, Eds.), pp. 12–42, Narosa Publishing House, New Delhi, London, 1998.
71. Siddiqi, A.H., Khan, S. and Rehman, S.: Wavelet-Based Computer Simulation for Wind Speed in Saudi Arabia, Proc. First Saudi Conference, KFUPM, Dhahran, 9–11 April, 2001, pp. 313–326.
72. Siddiqi, A.H.: On the summability of sequence of Walsh functions, J. Aust. Math. Soc. 10(1969), 385–394.

73. Siddiqi, A.H.: A note on Hardy's theorem for the arithmetic means of Fourier coefficients, Math. Students, 40(A) (1972), 14–113; MR No. 9223/48 (1974).
74. Siddiqi, A.H., Aslan, Z. and Tokgozlu, A.: Wavelet-Based Computer Simulation of Some Meteorological Parameters: Studies in Turkey, To appear in Proc. First International Conference on Industrial and Applied Mathematics—Emerging Trends in Industrial and Applied Mathematics, Kluwer Academic Publishers, Boston-Dordrecht-London, November 2001.
75. Siddiqi, J.A.: The Fourier series of continuous functions of bounded variation, Math. Ann., 143(1961), 103–108.
76. Siddiqi, J.A.: Coefficient properties of certain Fourier series, Math. Ann., 181(1969), 242–254.
77. Suhail, M.A., Siddiqi, A.H. and Simpson, A New Fast DCT-based Digital Watermarking Technique, in "Proc. International Conference on Emerging Areas of Industrial and Applied Mathematics," A.H. Siddiqi, and M. Kocvara, Eds.), Kluwer Academic Publishers, Boston-Dordrecth-London, 2001.
78. Swanson, M.D., M. Kobayashi, and Twefik, A.H.: Multimedia Data-Embedding and Watermarking Technologies Proc. IEEE, Vol. 86, Issue 6, pp. 1064–1087, June 1998.
79. Tateoka, J.: On the characterization of Hardy-Besov spaces on the dyadic group and its application, Studia Math., 110(1994), 127–148.
80. Tateoka, J. and Wade, W.R.: On the strong approximation and summability by Cesáro means on the Besov spaces over the 2-series field, Acta. Sci. Math, (Szged), 60(1995), 685–703.
81. Tokgozlu, A., Siddiqi, A.H. and Aslan, Z.: Denoising and Analysis of Flight Measurements, Preprint, 2001.
82. Tokgozlu, A., Siddiqi, A.H. and Aslan, Z.: De-Noising and Analysis of Flight Measurements, Preprint 2001.
83. Vogel, C.R. and Oman, M.E.: Fast, robust total variation-based reconstruction of noisy, blurred images, IEEE Trans. Image Process., 7(6)(1998), 813–824.
84. Wade, William, R.: Walsh Fourier series, International J. Math. Sci., 5(4) (1982), 625–637.
85. Wade, William, R.: Walsh Fourier series, Contemporary Mathematics, AMS, 208(1997), 313–350.
86. Wade, W.R.: Book review, Walsh series and transforms by Golubov, Effimov, and Skvortsov, Bulletin (New Series), Amer. Math. Soc., 26(No. 2) (1992), 348–359.
87. Wade, W.R.: Recent developments in the theory of Haar series, Colloquium Mathematicum, Vol. L11(1987), 213–238.
88. Wickerhauser, M.V.: Adapted Wavelet Analysis from Theory to Software, A.K. Peters, Wollesley, Massachusets, 1994.
89. Wickerhauser, M.V.: Acoustic Signal Compression with Wavelet Packets, in "Wavelets: A Tutorial in Theory and Applications" (C.K. Chui., Ed.), pp. 679–700, Academic Press, Boston, San Diego, 1992.