Monografías Matemáticas *García de Galdeano* **42**, 79–93 (2019)

MATHEMATICAL ASPECTS OF COMPUTERIZED TOMOGRAPHY: COMPRESSION AND COMPRESSED COMPUTING

Benedikt Diederichs, Tomas Sauer and A. Michael Stock

Abstract. Modern industrial tomography can produce such huge amounts of data that they cannot be handled any more by normal computers. To overcome this problem, the data can be represented and even further compressed by means of a sparse representation with thresholding, as, for example, a three dimensional tensor product wavelet representation. This approach, on the other hand, requires that all operations are realized in the sparse basis. After introducing the basic concepts behind this approach, we show one explicit example, namely how to compute the correlation of two objects by means of sparse representations.

Keywords: Tomography, wavelets, compression.

AMS classification: AMS classification codes.

§1. Introduction

While computerized tomography (CT) is a standard method in medical diagnosis, it is less widely known that tomography is also applied quite frequently in industrial applications. These applications comprise metrology and reverse engineering, documentation and digitalization, for example of cultural heritage, as well as nondestructive testing in manufacturing processes. In contrast to medical applications, these scans are not restricted in size, materials and nature of the objects. A so-called XXL–CT can scan even a full size car with the help of a particle accelerator, a micro or nano CT may scan a "normal sized" object with an extremely high resolution and an inline CT may scan one object per second.

What all these applications have in common is the fact that they produce a huge amount of data: large objects and high-resolution scans can easily reach one Terabyte and more of voxel data after reconstruction, and even if the individual inline scans are usually of moderate size, they come in a large number, typically hundreds of scans every day. Clearly, these circumstances provide new challenges for image processing. The large variety of objects and tasks that occur in industrial CT require advanced and extremely flexible algorithms for segmentation, object separation and information extraction. In medical applications a lot of a priori knowledge can be applied: for most organs, for example, location, size and shape are roughly known and can often been modeled quite efficiently by combining geometric primitives like ellipsoids, cylinders and cones. In industrial CT, the effort of these methods is usually too large for complex separation tasks and methods from machine learning have to be applied so that the system automatically extracts the relevant aspects of different parts.

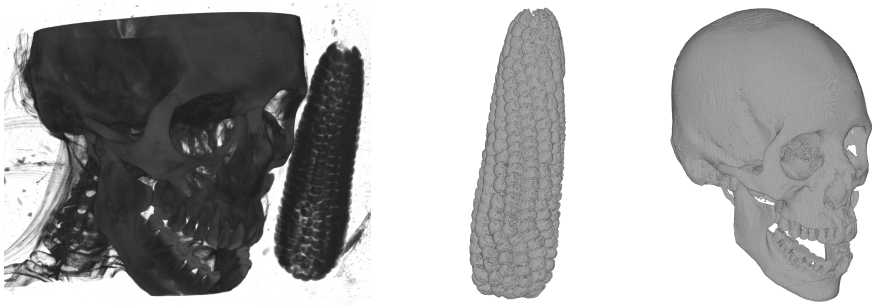

Figure 1: Cultural heritage CT bigture (Peruvian *mummy*, about 6th century AD, courtesy of *Lindenmuseum, Stuttgart*): *(Left)* region of interest containing skull and corncob, *(middle)* segmented corncob, *(right)* segmented skull.

An example for the complexity of such segmentation tasks in cultural heritage can be seen in Fig. 1.

But the main obstacle, of course, is the sheer size of data which leads us to an important concept.

Definition 1. An image is called a *bigture* (with respect to a certain representation) if it cannot be handled in the main memory of a computer any more.

The image on the left hand side of Fig. 1 is an example of a bigture: the size of the original image is about 170GB.

Whether an image is a bigture or not depends on two aspects: the size of the computer memory and the representation of the image. The trivial solution of the *bigture problem* would be to increase the computer memory and to rely on out of core memory techniques; while this is possible to a certain extent, it is not really practical since even if we take the very optimistic and somewhat unrealistic point of view that loading and access times only scale linearly with the amount of data, there is a significant slowdown, especially if we take into account that the amount of data scales *cubically* with the resolution: if we double the resolution of the image, the data increases by a factor of eight.

The more promising approach is to another *representation* of the image which is *sparse*. This means that we represent the same image, i.e., the same *information*, by a significantly smaller amount of *data*. Fortunately, such bases are known in many instances and as a general purpose tool, *wavelets* are still one of the best bases for sparse representation of discrete data on a rectangular grid, especially when this data is locally constant. This is one of the reasons why they were integrated into the *JPEG2000* standard. The drawback of sparse representations, however, is that now all operations have to be implemented in terms of the sparse basis as reconstruction of the full image would result in a bigture and thus render the algorithm useless. This paper will deal with some special case, namely computing the *correlation* of two bigtures using only their sparse representation. Correlations are important to register images and thus a fundamental operation in image processing.

The paper is organized as follows: we first give a short recapitulation of some of the basic concepts of computerized tomography, then recall the basics of the discrete wavelet

transformations, based on which we derive a method to compute the correlations depending on different shifts.

§2. Tomography basics

Tomography reconstructs objects from lower dimensional projections. Here, we focus on X-ray tomography which is based on the *attenuation* of X-rays by different materials. Let $f : \mathbb{R}^d \to \mathbb{R}$ be the function that describes the spatial distribution of this attenuation, where the standard cases are $d = 2$ and $d = 3$. Moreover, let L denote the *straight line* that connects the radiation source and the detector pixel. Then the energy I_D arriving at the detector pixel has the form

$$\frac{I_D}{I_S} = \int_L f(x)\, dx. \tag{1}$$

Note that (1) is only a first order model of the physical process that does not take into account effects like scattering and beam hardening. On the other hand, provided that the intersection of the support of f with L is contained in the line segment between source and detector, X-ray attenuation measures, in $\mathbb{R}^2$, the value of the *Radon transform*

$$L \mapsto Rf(L) := \int_L f(x)\, dx, \qquad L = L(v, s) = \left\{x \in \mathbb{R}^d : v^T x = s\right\} \in \mathcal{L}, \tag{2}$$

where $\mathcal{L}$ denotes the set of all lines in $\mathbb{R}^2$. In $\mathbb{R}^3$ the situation is more intricate as the Radon transform is then defined for planes and in the general s-dimensional case for hyperplanes.

In 2D slices the classical reconstruction is based on the *filtered backprojection* formula

$$(R^* g) * f = R^*(g * Rf), \qquad f \in L_1(\mathbb{R}^2), \quad g \in \mathcal{S}(\mathcal{L}), \tag{3}$$

where $\mathcal{S}(\mathcal{L})$ is the Schwartz space and

$$R^* g(x) = \int_{\|v\|=1} g(v, v^T x) dv, \qquad x \in \mathbb{R}^2, \quad g \in \mathcal{S}(\mathcal{L}) \subset \mathcal{S}(\mathbb{S}^1 \times \mathbb{R}),$$

stands for the *dual Radon transform*, where we cover $\mathcal{L}$ by $\mathbb{S}^1 \times \mathbb{R}$. For the application of (3), one chooses g such that $R^* g$ is close to the Dirac delta functional, resulting in $(R^* g) * f \approx f$. Typically, g is a radial function, i.e., $g(v, v^T x) = g(|v^T x|)$ and constructed such that convolving with g acts as a low-pass filter, hence the name filtered back projection. This formula is then discretized.

Similar inversion formulas exist in the three dimensional case. However, while in the two dimensional case typically a rather dense sampling of all lines passing through the support of f is available, the scanning geometries used in practice for the three dimensional case are more limited. Specific approximate inversion formulas, tailored to different geometries, are available. For example in the important case of the cone-beam geometry, the classical *Feldkamp algorithm* is widely used. For details on analytical methods see [10, 11].

A different approach is to see (1) as a system of integral equations and to discretize those directly. This could be done in a function as in Galerkin methods, but the standard technique

is to discretize the region of interest Ω into a *voxel* grid

$$\Omega = \bigcup_{j=1}^{N} V_j, \qquad V_j = [a_{j1}, b_{j1}] \times \cdots \times [a_{jd}, b_{jd}],$$

to assume the function f to have the constant value f_j on V_j and to rewrite (1) as

$$\frac{I_D}{I_S} =: y_L = \sum_{j \in J_L} \lambda_{L,j} f_j, \qquad J_L := \left\{ 1 \le j \le n : V_j \cap L \neq \emptyset \right\}. \tag{4}$$

The values λ_{Lj} is normally chosen as the length of the intersection $V_j \cap L$. Using a finite set of measurement rays, $L_1, \ldots, L_M$, which describe the scanning geometry, we end up with the linear system

$$y = \Lambda f, \qquad \Lambda = \left[\lambda_{L_j,k} : \begin{array}{c} j = 1, \ldots, M \\ k = 1, \ldots, N \end{array} \right], \qquad y = \begin{bmatrix} y_{L_1} \\ \vdots \\ y_{L_M} \end{bmatrix} \in \mathbb{R}^M, \; f = \begin{bmatrix} f_1 \\ \vdots \\ f_N \end{bmatrix} \in \mathbb{R}^N. \tag{5}$$

This is the simple concept of the *algebraic reconstruction technique* (ART), but solving the linear system (5) is far from trivial since it is usually huge. Nevertheless ART has some advantages:

1. The method works in any dimensionality and with any measurement geometry. Whether cone beam or parallel beam is used, only results in a different geometry matrix Λ.
2. The method works independently of dimension and the matrix is modestly sparse: if the voxels are arranged on an $n \times \cdots \times n$ grid, i.e., $N = n^d$, then any equation still involves only $O(n)$ variables.
3. A priori knowledge like obstructions or side conditions like positivity can be easily integrated into the approach as well as regularizers.

For more details see, for example [2, 6] and, of course, [10].

§3. Wavelet basics and definitions

Next, we briefly fix the notation for a wavelet *multiresolution analysis* (MRA). The starting point is a *refinable function* ϕ, i.e., a solution of the *refinement equation*

$$\phi = \sum_{k \in \mathbb{Z}} a_k \, \phi(2 \cdot - k), \qquad a \in \ell_0(\mathbb{Z}), \tag{6}$$

where $\ell_0(\mathbb{Z})$ stands for the space of all bi–infinite sequences with finite support:

$$\ell_0(\mathbb{Z}) = \{ c = (c_k : k \in \mathbb{Z}) : \|c\|_0 < \infty \}, \qquad \|c\|_0 := \#\{k : c_k \neq 0\}.$$

In addition, ϕ is called an *orthonormal scaling function* if its integer translates are mutually orthonormal, that is,

$$\delta_{k0} = \langle \phi, \phi(\cdot - k) \rangle = \int_{\mathbb{R}} \phi(x) \, \phi(x - k) \, dx, \qquad k \in \mathbb{Z}.$$

Substituting the refinement equation (6) into this requirement, it is easily seen that the sequence a, called *mask* in the subdivision literature [1], has to satisfy the *quadrature mirror filter equation*

$$\delta_{k0} = \frac{1}{2} \sum_{j \in \mathbb{Z}} a_{k-2j} a_j, \qquad k \in \mathbb{Z}. \tag{7}$$

A generic construction for finitely supported masks that satisfy (7) and give rise to L_2–solutions of (6) has first been given by Daubechies [4], see also [5]. This was the starting point for a multitude of different wavelet constructions, orthogonal as well as biorthogonal ones, and eventually the inclusion of wavelets into the JPEG2000 standard.

Based on an orthogonal scaling function, we define the MRA as the sequence of spaces

$$V_j := \operatorname{span} \left\{ \phi(2^j \cdot -k) : k \in \mathbb{Z} \right\}, \qquad j \in \mathbb{N}_0;$$

by (6), these spaces are *nested* in the sense that $V_0 \subset V_1 \subset \cdots$ and the associated *wavelet*, defined as

$$\psi = \sum_{k \in \mathbb{Z}} b_k \, \phi(2 \cdot -k), \qquad b_k := (-1)^k a_{1-k}, \quad k \in \mathbb{Z}, \tag{8}$$

belongs to V_1 and satisfies $\langle \phi, \psi(\cdot - k) \rangle = 0$, $k \in \mathbb{Z}$, so that

$$V_{j+1} = V_j \oplus W_j, \qquad W_j := \operatorname{span} \left\{ \psi \left(2^j \cdot -k \right) : k \in \mathbb{Z} \right\}, \qquad j \in \mathbb{N}_0. \tag{9}$$

Moreover, the integer shifts of the wavelet ϕ form an orthonormal basis of W_0 and, accordingly, the functions $\psi_{j,k} := 2^{j/2} \psi(2^j \cdot -k)$, $k \in \mathbb{Z}$, are an orthonormal basis of W_j. Hence, any function $f \in V_n$ can be written as

$$f = \sum_{k \in \mathbb{Z}} c_k^n(f) \, \phi \left(2^n \cdot -k \right) = \sum_{k \in \mathbb{Z}} c_k(f) \, \phi(\cdot - k) + \sum_{j=0}^{n-1} \sum_{k \in \mathbb{Z}} d_k^j(f) \, 2^{j/2} \psi \left(2^j \cdot -k \right), \tag{10}$$

where

$$d_k^j(f) = 2^{j/2} \int_{\mathbb{R}} f(x) \, \psi(2^j x - k) \, dx.$$

The main point in favor of wavelets, however, is that the conversion from c^n to $c, d^0, \ldots, d^{n-1}$, i.e., the transmission between the two representations of f in (10) can be performed very efficiently by means of discrete *filterbank* operations; this is Mallat's *discrete wavelet transform* [8, 9], see also [13].

Wavelets are naturally related to *subdivision schemes*. This is an immediate consequence of the refinement equation (6) which yields that for any $f \in V_0$ we have

$$f = \sum_{j \in \mathbb{Z}} c_j(f) \, \phi(\cdot - j) = \sum_{j,k \in \mathbb{Z}} c_j(f) \, a_k \, \phi(2 \cdot -2j - k) = \sum_{k \in \mathbb{Z}} \left(\sum_{j \in \mathbb{Z}} a_{k-2j} \, c_j(f) \right) \phi(2 \cdot -k),$$

or, in terms of (semidiscrete) convolutions,

$$f = c(f) * \phi = (S_a c(f)) * \phi(2 \cdot), \qquad (S_a c)_j := \sum_{k \in \mathbb{Z}} a_{j-2k} \, c_k, \tag{11}$$

and the subdivision operator S_a. We will use this connection later.

In s variables, usually $s = 2, 3$, we use the respective tensor product scaling functions

$$\phi(x) = \prod_{r=1}^{s} \phi(x_r),$$

so that, for $\alpha \in \mathbb{Z}^s$,

$$\int_{\mathbb{R}^s} \phi(x)\,\phi(x-\alpha)\,dx = \prod_{r=1}^{s} \int_{\mathbb{R}} \phi(x_r)\,\phi(x_r-\alpha_r)\,dx_r = \prod_{r=1}^{s} \delta_{\alpha_r,0} = \delta_{\alpha,0}.$$

To build wavelets, we set $\psi_0 = \phi$, $\psi_1 = \psi$ and define the $2^s - 1$ wavelet functions

$$\psi_\eta(x) := \prod_{r=1}^{s} \psi_{\eta_r}(x_r), \qquad \eta \in H := \{0,1\}^s \setminus \{0\}$$

and the refinement equation is given by

$$\psi_\eta = \sum_{\alpha\in\mathbb{Z}^s} b_{\eta,\alpha}\,\phi(2\cdot-\alpha), \qquad \eta \in H,$$

with

$$b_{\eta,\alpha} := \prod_{r=1}^{s} \left((1-\eta_r)a_{\alpha_r} + \eta_r b_{\alpha_r}\right), \qquad \eta \in H, \quad \alpha \in \mathbb{Z}^s.$$

These function satisfy the orthonormality condition

$$\int_{\mathbb{R}^s} \psi_\eta(x)\,\psi_{\eta'}(x-\alpha)\,dx = \delta_{\eta,\eta'}\,\delta_{\alpha,0}, \qquad \eta,\eta' \in \{0,1\}, \quad \alpha,\beta \in \mathbb{Z}^s.$$

Thus, with $V_j = \operatorname{span}\left\{\phi(2^j\cdot-\alpha) : \alpha \in \mathbb{Z}^s\right\}$, we again have the multiresolution analysis

$$V_{j+1} = V_j \oplus W_j, \qquad W_j := \operatorname{span}\left\{2^{js/2}\,\psi_\eta(2^j\cdot-\alpha) : \eta \in H,\ \alpha \in \mathbb{Z}^s\right\},$$

and the wavelet decomposition

$$f = \sum_{\alpha\in\mathbb{Z}^s} c^n_\alpha(f)\,\phi(2^n\cdot-\alpha) = \sum_{\alpha\in\mathbb{Z}^s} c_\alpha(f)\,\phi(\cdot-\alpha) + \sum_{j=0}^{n-1}\sum_{\eta\in H}\sum_{\alpha\in\mathbb{Z}^s} d^j_{\eta,\alpha}(f)\,2^{js/2}\psi_\eta\left(2^j\cdot-\alpha\right). \quad (12)$$

Remark 1. There exist orthogonal wavelet decompositions for arbitrary scaling matrices and even a generic tensor-product like approach to construct them, see [3, 7], but the classical dyadic tensor product construction offers an extremely efficient way of localizing the supports of scaling functions and wavelets which is useful in the implementation of a fast decomposition and reconstruction as described in [12].

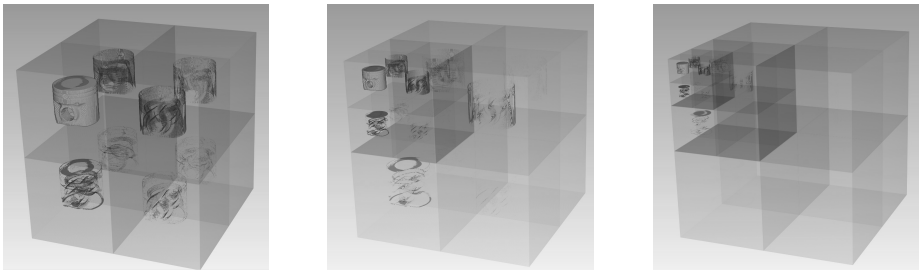

Figure 2: Different levels of wavelet decomposition of a typical industrial piece (engine piston). *(Left)* One level, the different types of edge and face dectections are clearly visible, *(middle)* even after two levels wavelet coefficients become almost invisible, *(right)* and after three iterations, they are mainly irrelevant.

§4. Wavelets and compression

The idea behind wavelet compression is rather simple: if a function $f : \mathbb{R}^s \to \mathbb{R}$ can be approximated well by V_k for some rather small k, then the wavelet coefficients $d^j_{\eta,\alpha}(f)$ for $j > k$ will be very small and can be discarded without significant loss of quality. This works for two reasons, cf. [5, 9]:

1. Any reasonable scaling function ϕ is compactly supported and (re)produces polynomials of a certain degree, but at least satisfies

$$\sum_{\alpha\in\mathbb{Z}^s} \phi(\cdot - \alpha) = 1.$$

   This means that the wavelet coefficients for *locally constant* functions vanish in the regions where the function is constant.

2. The spaces V_k usually have good approximation properties for smooth function which implies that in smooth regions the absolute values $\left|d^j_{\eta,\alpha}(f)\right|$ of the wavelet coefficients decay very fast with respect to j.

Starting from the high resolution $c^n(f)$ of

$$f \approx \sum_{\alpha\in\mathbb{Z}^s} c^n_\alpha(f)\, \phi(2^n \cdot -\alpha),$$

one computes the wavelet coefficients $d^{n-1}_\eta(f), d^{n-2}_\eta(f), \cdots, d^0_\eta(f)$ and the scaling coefficients $c^0(f)$ by means of a fast wavelet transform. Note that the index η also has an intuitive geometric meaning for the wavelet coefficient. Indeed, if $|\eta| = 1$, it detects faces parallel to the coordinate planes, if $|\eta| = 2$ the coefficients correspond to edges parallel to the axes and $\eta = (1, 1, 1)$ detects some "diagonal" feature, see Fig. 2.

While the originally sampled data $c^n(f)$ is usually dense, the wavelet transform is sparse if the underlying image is piecewise constant which is the case in most technical applications,

see again Fig. 2. In addition to deleting zero coefficients, lossy compression is obtained by thresholding the wavelet coefficients and replacing them by

$$\hat{d}^j_{\eta,\alpha} = t_\epsilon\left(d^j_{\eta,\alpha}\right), \qquad \eta \in H,\ \alpha \in \mathbb{Z}^s,\ j = 0,\dots,n-1,$$

by means of threshold function t_ε.

Definition 2. For a threshold level ε the *hard threshold* and the *soft threshold* use the functions

$$t^h_\varepsilon(x) = \begin{cases} 0, & |x| < \varepsilon, \\ x, & |x| \ge \varepsilon, \end{cases} \qquad t^s_\varepsilon(x) = \begin{cases} 0, & |x| < \varepsilon, \\ x - \varepsilon, & x \ge \varepsilon, \\ x + \varepsilon, & x \le -\varepsilon. \end{cases}$$

respectively.

While soft thresholding is known to perform a denoising operation, popular as *wavelet shrinkage*, hard threshold is more contrast and edge preserving. The choice of the threshold level can be made according to several strategies, for example

1. absolute choice of threshold level,
2. *best N–term approximation:* the threshold is chosen in such a way that only the N largest coefficients remain,
3. overall precision: ϵ is chosen such that

$$\sum_{j=0}^{n-1} \sum_{\eta\in H} \sum_{\alpha\in\mathbb{Z}^s} \left(d^j_{\eta,\alpha} - \hat{d}^j_{\eta,\alpha}\right)^2$$

   does not exceed a certain bound. Since these are the coefficients in an orthonormal expansion, this is also the norm of the L_2–error, hence a certain PSNR can be prescribed for the compression.

Recall also that after transformation and thresholding, the array of coefficients is encoded in an efficient way using a more or less standard entropy encoder, cf. [12] for details. We will not dwell on these issues here though they are of course important for the overall compression rates.

Definition 3. For a function f with a thresholded wavelet decomposition we define

$$N(f) := \#\{\alpha : \hat{c}_\alpha \ne 0\} + \sum_{j=0}^{n-1} \sum_{\eta\in H} \#\left\{\alpha : \hat{d}^j_{\eta,\alpha} \ne 0\right\}$$

as the number of nonzero coefficients in the representation.

§5. Wavelet correlation

The *correlation* between two functions f, g is defined as

$$f \star g(y) := \int_{\mathbb{R}^s} f(x)\, g(x+y)\, dx, \qquad y \in \mathbb{R}^s,$$

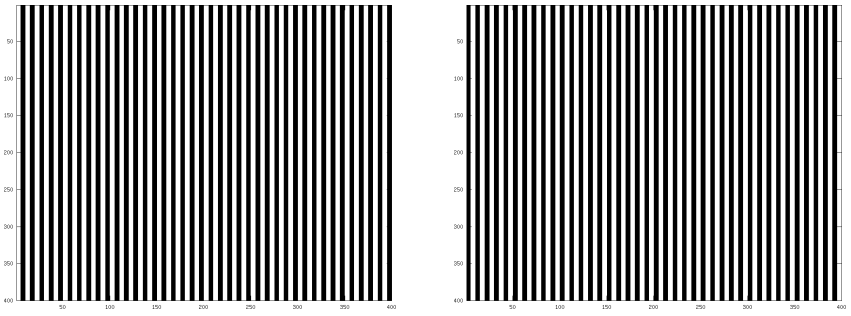

Figure 3: Two images with a PSNR value of 0, i.e., of maximal disparity, while $\sigma(f, g) = 1$.

and measures the best fit between f and a shifted version of g. It is useful for image comparison by using

$$\sigma(f, g) := \frac{1}{\|f\|_2 \, \|g\|_2} \max_{y \in \mathbb{R}^s} |f \star g(y)|$$

as a translation invariant similarity measure for images. An extremal example in this respect can be seen in Fig. 3. Moreover, local correlations are needed in order to stitch images together by finding a proper offset of one of the images such that the overlapping areas coincide as much as possible.

Taking into account that in many cases a complete reconstruction of a bigture is impossible due to memory limitations, we need an algorithm that computes the correlation entirely from the wavelet decomposition. We will develop such a method in this section. To that end, we set up some terminology first.

Definition 4. The *correlation* of two sequences $c, d \in \mathbb{Z}^s$ is defined as

$$(c \star d)_\alpha := \sum_{\beta \in \mathbb{Z}^s} c_{\alpha+\beta} \, d_\beta, \qquad \alpha \in \mathbb{Z}^s,$$

and the *translation operator* as

$$\left(\tau_\gamma c\right)_\alpha := c_{\alpha+\gamma}, \qquad \alpha \in \mathbb{Z}^s.$$

We start with $f, g \in V_n$, written as

$$f = \sum_{\alpha \in \mathbb{Z}^s} c_\alpha(f) \, \phi(\cdot - \alpha) + \sum_{j=0}^{n-1} 2^{js/2} \sum_{\eta \in H} \sum_{\alpha \in \mathbb{Z}^s} d^j_{\eta,\alpha}(f) \, \psi_\eta(2^j \cdot - \alpha),$$

$$g = \sum_{\alpha \in \mathbb{Z}^s} c_\alpha(g) \, \phi(\cdot - \alpha) + \sum_{j=0}^{n-1} 2^{js/2} \sum_{\eta \in H} \sum_{\alpha \in \mathbb{Z}^s} d^j_{\eta,\alpha}(g) \, \psi_\eta(2^j \cdot - \alpha),$$

and first observe that the integer correlations can be easily computed as shifted correlations of the discrete sequences.

Lemma 1. *For $\gamma \in \mathbb{Z}^s$ we have that*

$$f \star g(\gamma) = \tau_{-\gamma}\left(c(f) \star c(g)\right) + \sum_{j=0}^{n-1} \sum_{\eta \in H} \tau_{-2^j\gamma}\left(d_\eta^j(f) \star d_\eta^j(g)\right). \tag{13}$$

The number of arithmetic operations is bounded by $\min\{N(f), N(g)\}$ *and hence subadditive in the number of nonzero coefficients in the expansion of f or g, respectively.*

Proof. We compute

$$\begin{aligned}
&f \star g(\gamma)\\
&= \sum_{\alpha,\beta \in \mathbb{Z}^s} c_\alpha(f)\, c_\beta(g)\, \langle \phi(\cdot - \alpha), \phi(\cdot - \beta + \gamma)\rangle\\
&\quad + \sum_{j=0}^{n-1} 2^{-js/2} \sum_{\eta \in H} \sum_{\alpha,\beta \in \mathbb{Z}^s} \left(c_\alpha(f)\, d^j_{\eta,\beta}(g) + c_\alpha(g)\, d^j_{\eta,\beta}(f)\right) \left\langle \phi(\cdot - \alpha), \psi_\eta\left(2^j \cdot - \beta + 2^j\gamma\right)\right\rangle\\
&\quad + \sum_{j,k=0}^{n-1} 2^{-(j+k)s/2} \sum_{\eta,\eta' \in H} \sum_{\alpha,\beta \in \mathbb{Z}^s} d^j_{\eta,\alpha}(f)\, d^k_{\eta',\beta}(g) \left\langle \psi_\eta\left(2^j \cdot - \alpha\right), \psi_{\eta'}\left(2^k \cdot - \beta + 2^k\gamma\right)\right\rangle\\
&= \sum_{\alpha,\beta \in \mathbb{Z}^s} c_\alpha(f)\, c_\beta(g)\, \delta_{\alpha,\beta-\gamma} + \sum_{j,k=0}^{n-1} \sum_{\eta,\eta' \in H} \sum_{\alpha,\beta \in \mathbb{Z}^s} d^j_{\eta,\alpha}(f)\, d^k_{\eta,\beta}(g) \delta_{jk} \delta_{\alpha,\beta-2^k\gamma} \delta_{\eta,\eta'}\\
&= \sum_{\beta \in \mathbb{Z}^s} c_{\beta-\gamma}(f)\, c_\beta(g) + \sum_{j=0}^{n-1} \sum_{\eta \in H} \sum_{\beta \in \mathbb{Z}^s} d^j_{\eta,\beta-2^j\gamma}(f)\, d^j_{\eta,\beta}(g),
\end{aligned}$$

which gives (13). Since any contribution to the sum requests a nonzero coefficient of the expansion of f and g, the number of arithmetic operations is bounded by $\min\left(N(f), N(g)\right)$. □

For the general case, we define the bi–infinite matrix valued function

$$\Phi(y) := \left[\langle \phi(\cdot - \alpha), \phi(\cdot - \beta + y)\rangle : \alpha, \beta \in \mathbb{Z}^s\right], \qquad y \in \mathbb{R}^s, \tag{14}$$

which represents the correlation for $f, g \in V_0$ in the sense that $f \star g(y) = c(f)^T \Phi(y)\, c(g)$.

Lemma 2. *If $\phi \in L_2(\mathbb{R}^s)$ is a compactly supported orthonormal scaling function, then*

1. *the matrix $\Phi(y)$ is a banded* Toeplitz matrix*,*
2. *$\Phi(y)$ is almost* 1*-periodic: $\Phi(y + \gamma) = \Phi(y)\tau_\gamma = \tau_{-\gamma}\Phi(y)$, $\gamma \in \mathbb{Z}^s$,*
3. *$y \mapsto \Phi(y)$ is continuous with $\Phi(0) = I$, in the sense that the coefficients form a uniformly equicontinuous family.*

Proof. For 1) we note that

$$\Phi(y)_{\alpha,\beta} = \langle \phi(\cdot - \alpha), \phi(\cdot - \beta + y)\rangle = \langle \phi\left(\cdot - (\alpha - \beta)\right), \phi(\cdot + y)\rangle,$$

hence depends only on $\alpha - \beta$. Bandedness results from the compact support of ϕ: if $\|\alpha - \beta\|$ is large enough, then the supports of $\phi(\cdot - (\alpha - \beta))$ and $\phi(\cdot + y)$ are disjoint and the integral is zero. For 2) we consider

$$(\Phi(y+\gamma)c)_\alpha = \sum_{\beta\in\mathbb{Z}^s} \langle \phi(\cdot-\alpha), \phi(\cdot-(\beta-\gamma)+y)\rangle c_\beta = \sum_{\beta\in\mathbb{Z}^s} \langle \phi(\cdot-\alpha), \phi(\cdot-\beta+y)\rangle c_{\beta+\gamma}$$

and

$$\left(c^T\Phi(y+\gamma)\right)_\beta = \sum_{\alpha\in\mathbb{Z}^s} \langle \phi(\cdot-(\alpha+\gamma)), \phi(\cdot-\beta+y)\rangle c_\alpha = \left(\tau_{-\gamma}c^T\Phi(y)\right)_\beta.$$

To prove 3) we first note that due to 1) the matrix $\Phi(y)$ only contains finitely many nonzero entries of the form

$$\langle \phi(\cdot-\alpha), \phi(\cdot+y)\rangle, \qquad \alpha\in\mathbb{Z}^s.$$

For each $\alpha\in\mathbb{Z}^s$ and $\delta\in\mathbb{R}^s$ we then have that

$$|\langle \phi(\cdot-\alpha), \phi(\cdot+y+\delta)\rangle - \langle \phi(\cdot-\alpha), \phi(\cdot+y)\rangle| \le \|\phi\|_2 \ \|\phi(\cdot+\delta)-\phi\|_2$$

which tends to zero for $\|\delta\|\to 0$ uniformly in α and y, as $\phi\in L_2(\mathbb{R}^s)$. □

Example 1. In the simple case of Haar wavelets where $\phi = \chi = \chi_{[0,1]^s}$, the entries of $\Phi(y)$ can easily be computed explicitly, namely, for $y\in(0,1)^s$ as

$$\Phi(y)_{\alpha,\beta} = \int_{[0,1]^s} \chi(x+\alpha-\beta+y)\,dx = \prod_{r=1}^{s}\int_0^1 \chi(x+\alpha_r-\beta_r+y_r)\,dx.$$

Now, $[0,1]+\alpha_r-\beta_r+y_r\cap[0,1]\neq\emptyset$ only if $\beta_r=\alpha_r+1$ or $\beta_r=\alpha_r$, where

$$\int_0^1 \chi(x+\alpha_r-\beta_r+y_r)\,dx = \int_{-1}^0 \chi(x+y_r) = y_r$$

in the first case and

$$\int_0^1 \chi(x+\alpha_r-\beta_r+y_r)\,dx = \int_0^1 \chi(x+y_r) = 1-y_r$$

in the latter. Therefore, every row of $\Phi(y)$ contains exactly 2^s nonzero values, namely

$$\Phi(y)_{\alpha,\alpha+\eta} = \prod_{r=1}^{s}(\eta_r y_r + (1-\eta_r)(1-y_r)), \qquad \alpha\in\mathbb{Z}^s, \quad \eta\in\{0,1\}^s.$$

For arbitrary $y\in\mathbb{R}^s$, we apply Lemma 2. Note however that there is a shift in the matrices in this case:

$$\lim_{y\to(1,\dots,1)}\Phi(y) = \tau_{(1,\dots,1)} = \lim_{y\to 0}\Phi(y)\tau_{(1,\dots,1)}.$$

Since the finely sampled data $c^n(f)$ corresponds to evaluation of a function on the grid $2^{-n}\mathbb{Z}^s$, even if the sequence is indexed by integers, we have to be able to compute correlations at least for dyadic values $2^{-m}\gamma$, $\gamma\in\mathbb{Z}^s$, $0\le m\le n$. The next result shows that also in this case, correlations can be computed from wavelet coefficients by means of $\Phi(y)$.

Theorem 3. *For $\gamma \in \mathbb{Z}^s$ and $0 \le m \le n$ we have that*

$$\begin{aligned} f \star g\,(2^{-m}\gamma) = c(f)^T \Phi\,(2^{-m}\gamma)\, c(g) &+ 2^{-s} \sum_{j=0}^{m-1} \sum_{\eta,\eta' \in H} \left(S_{b_\eta} d_\eta^j(f)\right)^T \Phi\left(2^{j+1-m}\gamma\right) \left(S_{b_{\eta'}} d_{\eta'}^j(g)\right) \\ &+ \sum_{0 \le j < k \le m-1} 2^{-\left(1+\frac{k-j}{2}\right)s} \sum_{\eta,\eta' \in H} \left(S_a^{k-j} S_{b_\eta} d_\eta^j(f)\right)^T \Phi\,(2^{-m}\gamma) \left(S_{b_{\eta'}} d_{\eta'}^k(g)\right) \\ &+ \sum_{0 \le k < j \le m-1} 2^{-\left(1+\frac{j-k}{2}\right)s} \sum_{\eta,\eta' \in H} \left(S_{b_\eta} d_\eta^j(f)\right)^T \Phi\,(2^{-m}\gamma) \left(S_a^{j-k} S_{b_{\eta'}} d_{\eta'}^k(g)\right) \\ &+ \sum_{j=m}^{n-1} \sum_{\eta \in H} \tau_{-2^{j-m}\gamma} \left(d_\eta^j(f) \star d_\eta^j(g)\right). \end{aligned} \tag{15}$$

Proof. As all levels at least as fine as m are still orthogonal, we obtain

$$\begin{aligned} & f \star g(2^{-m}\gamma) \\ &= \sum_{\alpha,\beta \in \mathbb{Z}^s} c_\alpha(f)\, c_\beta(g) \left\langle \phi(\cdot - \alpha), \phi(\cdot - \beta + 2^{-m}\gamma) \right\rangle \\ &\quad + \sum_{j,k=0}^{m-1} 2^{(j+k)s/2} \sum_{\eta,\eta' \in H} \sum_{\alpha,\beta \in \mathbb{Z}^s} d_{\eta,\alpha}^j(f)\, d_{\eta',\beta}^k(g) \left\langle \psi_\eta\left(2^j \cdot - \alpha\right), \psi_{\eta'}\left(2^k \cdot - \beta + 2^{k-m}\gamma\right) \right\rangle \\ &\quad + \sum_{j=m}^{n-1} \sum_{\eta \in H} \sum_{\beta \in \mathbb{Z}^s} d_{\eta,\beta-2^{j-m}\gamma}^j(f)\, d_{\eta,\beta}^j(g) \\ &= c(f)^T \Phi\,(2^{-m}\gamma)\, c(g) + \sum_{j=m}^{n-1} \sum_{\eta \in H} \sum_{\beta \in \mathbb{Z}^s} d_{\eta,\beta-2^{j-m}\gamma}^j(f)\, d_{\eta,\beta}^j(g) \\ &\quad + \sum_{j,k=0}^{m-1} 2^{-(j+k)s/2} \sum_{\eta,\eta' \in H} \sum_{\alpha,\beta \in \mathbb{Z}^s} d_{\eta,\alpha}^j(f)\, d_{\eta',\beta}^k(g) \left\langle 2^{js} \psi_\eta\left(2^j \cdot - \alpha\right), 2^{ks} \psi_{\eta'}\left(2^k \cdot - \beta + 2^{k-m}\gamma\right) \right\rangle, \end{aligned} \tag{16}$$

which already gives the first and the last term in the right hand side of (15). We are left with calculating the correlations of all levels from zero to $m-1$.

Recalling that

$$\psi_\eta = \sum_{\alpha \in \mathbb{Z}^s} b_{\eta,\alpha}\, \phi(2 \cdot - \alpha), \qquad \eta \in H,$$

we obtain for $j \le n-1$ that

$$\begin{aligned} \sum_{\alpha \in \mathbb{Z}^s} d_{\eta,\alpha}^j(f)\, \psi_\eta\left(2^j \cdot - \alpha\right) &= \sum_{\alpha \in \mathbb{Z}^s} d_{\eta,\alpha}^j(f) \sum_{\beta \in \mathbb{Z}^s} b_{\eta,\beta} \phi\left(2^{j+1} \cdot - 2\alpha - \beta\right) \\ &= \sum_{\beta \in \mathbb{Z}^s} \left(\sum_{\alpha \in \mathbb{Z}^s} b_{\eta,\beta-2\alpha}\, d_\alpha^j(f) \right) \phi\left(2^{j+1} \cdot - \beta\right) = \sum_{\beta \in \mathbb{Z}^s} \left(S_{b_\eta} d_\eta^j(f)\right)_\beta \phi\left(2^{j+1} \cdot - \beta\right), \end{aligned}$$

and the refinement equation (6) for ϕ yields in the same way for $k > j$ that

$$\sum_{\alpha \in \mathbb{Z}^s} d_{\eta,\alpha}^j(f)\, \psi_\eta\left(2^j \cdot - \alpha\right) = \sum_{\beta \in \mathbb{Z}^s} \left(S_a^{k-j} S_{b_\eta} d_\eta^j(f)\right)_\beta \phi\left(2^{k+1} \cdot - \beta\right).$$

With these formulas at hand, we split the sum from (16) into its diagonal

$$\begin{aligned}
&\sum_{j=0}^{m-1} 2^{js} \sum_{\eta,\eta'\in H} \sum_{\alpha,\beta\in\mathbb{Z}^s} d^j_{\eta,\alpha}(f)\, d^j_{\eta',\beta}(g) \Big\langle \psi_\eta\left(2^j\cdot-\alpha\right), \psi_{\eta'}\left(2^j(\cdot+2^{-m}\gamma)-\beta\right)\Big\rangle \\
&= \sum_{j=0}^{m-1} 2^{js} \sum_{\eta,\eta'\in H} \sum_{\alpha,\beta\in\mathbb{Z}^s} \left(S_{b_\eta} d^j_{\eta,}(f)\right)_\alpha \left(S_{b_{\eta'}} d^j_{\eta'}(g)\right)_\beta \left\langle \phi\left(2^{j+1}\cdot-\alpha\right), \phi\left(2^{j+1}(\cdot+2^{-m}\gamma)-\beta\right)\right\rangle \\
&= 2^{-s} \sum_{j=0}^{m-1} \sum_{\eta,\eta'\in H} \left(S_{b_\eta} d^j_\eta(f)\right) \Phi\left(2^{j+1-m}\gamma\right) \left(S_{b_{\eta'}} d^j_{\eta'}(g)\right),
\end{aligned}$$

the lower triangle

$$\begin{aligned}
&\sum_{0\le j<k\le m-1} 2^{(j+k)s/2} \sum_{\eta,\eta'\in H} \sum_{\alpha,\beta\in\mathbb{Z}^s} d^j_{\eta,\alpha}(f)\, d^k_{\eta',\beta}(g) \Big\langle \psi_\eta\left(2^j\cdot-\alpha\right), \psi_{\eta'}\left(2^k(\cdot+2^{-m}\gamma)-\beta\right)\Big\rangle \\
&\qquad = \sum_{0\le j<k\le m-1} 2^{(j+k)s/2} \sum_{\eta,\eta'\in H} \sum_{\alpha,\beta\in\mathbb{Z}^s} \left(S_{b_\eta} d^j_{\eta,}(f)\right)_\alpha \left(S_{b_{\eta'}} d^j_{\eta'}(g)\right)_\beta \\
&\qquad\qquad \cdot \left\langle \phi\left(2^{j+1}\cdot-\alpha\right), \phi\left(2^{k+1}(\cdot+2^{-m}\gamma)-\beta\right)\right\rangle \\
&\qquad = \sum_{0\le j<k\le m-1} 2^{(j+k)s/2} \sum_{\eta,\eta'\in H} \sum_{\alpha,\beta\in\mathbb{Z}^s} \left(S_a^{k-j} S_{b_\eta} d^j_{\eta,}(f)\right)_\alpha \left(S_{b_{\eta'}} d^j_{\eta'}(g)\right)_\beta \\
&\qquad\qquad \cdot \left\langle \phi\left(2^{k+1}\cdot-\alpha\right), \phi\left(2^{k+1}(\cdot+2^{-m}\gamma)-\beta\right)\right\rangle \\
&\qquad = \sum_{0\le j<k\le m-1} 2^{\left(\frac{j-k}{2}-1\right)s} \sum_{\eta,\eta'\in H} \left(S_a^{k-j} S_{b_\eta} d^j_{\eta,}(f)\right)^T \Phi\left(2^{k+1-m}\gamma\right) \left(S_{b_{\eta'}} d^j_{\eta'}(g)\right)
\end{aligned}$$

and the upper triangle

$$\sum_{m-1\ge j>k\ge 0} \cdots = \sum_{m-1\ge j>k\ge 0} 2^{\left(\frac{k-j}{2}-1\right)s} \sum_{\eta,\eta'\in H} \left(S_{b_\eta} d^j_{\eta,}(f)\right)^T \Phi\left(2^{j+1-m}\gamma\right) \left(S_a^{j-k} S_{b_{\eta'}} d^j_{\eta'}(g)\right)$$

that is obtained in the same way. □

Correlation is a standard tool for matching objects whose mutual displacement is due to shifts. This goal can often be achieved by aligning calibrated reference objects which are first used to compensate rotational effects. After that, maximizing the correlation means finding the best alignment between the two objects f and g. For this purpose we propose the following hierarchical algorithm:

1. Determine the best integer shift γ^0 by computing the correlations with the formula (13) from Lemma 1.
2. For $m = 1, 2, \ldots, n$ determine the best dyadic shift $2^{-m}\gamma^m$ among the $3^s - 1$ neighbors $2^{1-m}\gamma^{m-1} + 2^{-m}\kappa$, $\kappa \in \{-1, 1, 1\}^s \setminus \{0\}$ of $2^{1-m}\gamma^{m-1}$.

Note that for $m = 1$ the iteration only requires to apply the subdivision operator S_b to $d^0_\eta(f)$ and $d^1_\eta(g)$ and the matrices $\Phi(2^{-1}\gamma)$ and $\Phi(\gamma) = I\,\tau_\gamma$ to compute

$$c(f)^T \Phi\left(2^{-1}\gamma\right) c(g) + 2^{-s} \sum_{j=0}^{m-1} \sum_{\eta,\eta'\in H} \tau_{-\gamma}\left(S_{b_\eta} d^j_\eta(f) \star S_{b_{\eta'}} d^j_{\eta'}(g)\right) + \sum_{j=1}^{n-1} \sum_{\eta\in H} \tau_{-2^{j-m}\gamma}\left(d^j_\eta(f) \star d^j_\eta(g)\right).$$

If we store the level correlations seperately for $j = 0, \dots, n$, we only need to recompute the first two sums of the above expression. This technique can in general be extended to later iterations, but the two middle sums in (15) also necessitate the application of the subdivision S_a to the low level wavelet coefficient vectors. Of course, the finer the resolution becomes, the effort increases, but keep in mind that also then, due to the hierarchical procedure, only $3^s - 1$ correlations have to be computed which even for $s = 3$ is still the relatively moderate value of 26.

§6. Conclusions

Tomography is much more than a medical diagnosis tool, and the technology available for industrial tomography enables us to generate spectacular measurements with very high resolution or of very large objects. The resulting amount of data, however, is no more tractable on even well-equipped computer systems without switching to sparse representations. This provides a lot of mathematical challenges starting from an efficient creation of such sparse representations, requiring the development of compression and storage strategies that allow fast access to full resolution data in certain regions of interest and leading to a redefinition of standard operations in terms of the sparse representation – the sparse perspective only makes sense if it is respected in all steps of computations and manipulations.

The example of correlation shows that this is doable, even efficiently, but requires some non-straightforward mathematical operations.

Acknowledgements

The work in this paper has been funded by the Bavarian Ministry of Economy by means of the research project “Big Picture”. We also thank Thomas Lang (FORWISS, University of Passau) for providing the examples on segmentation in Fig. 1.

References

[1] Cavaretta, A., Dahmen, W., and Micchelli, C. *Stationary Subdivision*. American Mathematical Society: Memoirs of the American Mathematical Society. American Mathematical Society, 1991.

[2] Censor, Y. Row-action methods for huge and sparse systems and their applications. *SIAM Review 23* (1981), 444–466.

[3] Cotronei, M., Rossini, M., Sauer, T., and Volontè, E. Filters for anisotropic wavelet decompositions. *Journal of Computational and Applied Mathematics 349* (2019), 316 – 330.

[4] Daubechies, I. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics 41*, 7 (1988), 909–996.

[5] Daubechies, I. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992.

[6] Gordon, R., Bender, R., and Herman, G. T. Algebraic reconstruction techniques (art) for three-dimensional electron microscopy and x-ray photography. *Journal of Theoretical Biology 29*, 3 (1970), 471 – 481.

[7] Han, B. Compactly supported tight wavelet frames and orthonormal wavelets of exponential decay with a general dilation matrix. *Journal of Computational and Applied Mathematics 155* (2003), 43 – 67.

[8] Mallat, S. Multiresolution approximations and wavelet orthonormal bases of l2(r). *Transactions of the American Mathematical Society 315*, 1 (1989), 69–87.

[9] Mallat, S. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, Inc., 2008.

[10] Natterer, F. *The Mathematics of Computerized Tomography*. John Wiley & Sons, 1986.

[11] Natterer, F., and Wübbeling, F. *Mathematical Methods in Image Reconstruction*. Society for Industrial and Applied Mathematics, 2001.

[12] Stock, A. M., Herl, G., Sauer, T., and Hiller, J. Edge preserving compression of CT scans using wavelet. *International Symposium on Structural Health Monitoring and Nondestructive Testing* (2018).

[13] Vetterli, M., and Kovačevic, J. *Wavelets and Subband Coding*. Prentice-Hall, Inc., 1995.

B. Diederichs
University of Passau & Fraunhofer IIS Research Group “Knowledge Based Image Processing”
Innstr. 43, D–94032 Passau, Germany
benedikt.diederichs@iis.fraunhofer.de

T. Sauer
Lehrstuhl für Digitale Bildverabeitung & FORWISS, University of Passau
Fraunhofer IIS Research Group “Knowledge Based Image Processing”
Innstr. 43, D–94032 Passau, Germany
tomas.sauer@uni-passau.de

A. M. Stock
FORWISS, University of Passau
Innstr. 43, D–94032 Passau, Germany
stock@forwiss.uni-passau.de