Joint Multiresolution Magnetic Particle Imaging and System Matrix Compression

Marco Maass\textsuperscript{a,}\textsuperscript{*}, Christian Mink\textsuperscript{a}, Alfred Mertins\textsuperscript{a}

\textsuperscript{a}Institute for Signal Processing, University of Lübeck, Lübeck, Germany

\textsuperscript{*}Corresponding author, email: maass@isip.uni-luebeck.de

Received 27 March 2018; Accepted 02 August 2018; Published online 14 November 2018

© 2018 Maass; licensee Infinite Science Publishing GmbH
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

It is well known that the discrete cosine transform is a good transform to compress system matrices from magnetic particle imaging scanners that have a field-free point traveling along a Lissajous trajectory. With help of the compressed system matrix, the reconstruction of the particle distribution can be accelerated. However, the discrete cosine transform is a global transform along the spatial dimension of the system matrix. This has the disadvantage that no multiresolution analysis can be performed due to the loss of the spatial resolution. A typical strategy to implement a multiresolution analysis is to perform a discrete wavelet transform. Unfortunately, the discrete wavelet transform is not as sparsifying as the cosine transform for magnetic particle imaging system matrices. By combing the advantages of both transforms, we develop a strategy that gives rise to a multiresolution analysis with a similar performance as the discrete cosine transform at the full-resolution level, but with further resolution levels. We shortly discuss that the developed representation is a good starting point for advanced particle distribution reconstruction techniques that benefit from a multiresolution analysis.

I. Introduction

Magnetic particle imaging (MPI) is a tracer-based medical imaging method that is based on the nonlinear magnetization characteristics of super-paramagnetic iron oxide nanoparticles (SPIOs) \cite{1}. With different accelerated and static magnetic fields, the MPI-scanner generates a small area in which the magnetic fields neutralize each other. This area is called the field free point (FFP). The FFP is normally periodically moved on a pre-defined trajectory over the whole field of view (FOV). The change of magnetization leads to an induced voltage in a receive coil, where, due to the nonlinear magnetization characteristics of the SPIOs, only SPIOs from the vicinity of the FFP contribute significantly to the measured signal. For MPI-scanners with a Lissajous FFP-trajectory, the system response normally has to be saved. For this, a probe of SPIOs material is placed on different spatial positions, and the responses are saved in a so-called system matrix. With help of the system matrix, the inverse problem of estimating the SPIOs’ distribution from the voltage signal can be solved. Unfortunately, the system matrix can be very dense and huge in size. For a dense matrix, the reconstruction process can be very slow. In \cite{2}, it was observed that the system matrix of MPI-scanners with a FFP traveling along a Lissajous-trajectory can be highly compressed by the discrete cosine transform (DCT) followed by thresholding. Recently, a work for matrix compression was published on a non-Euclidean grid, where the Chebyshev transform becomes orthogonal and the compression performance is even improved \cite{3}. In this work, we develop a multiresolution representation \cite{4} for the system matrix. In particular, we use a combination of the DCT-II and the discrete wavelet transform (DWT) for the joint system-matrix compression and multiresolution reconstruction. We show that the multiresolu-
The idea is to choose $T$ in such a way that the matrix $S_T c_T = ST T^{-1} c = f$, (2)

where $T \in \mathbb{R}^{N \times N}$ describes an invertible transform, $S_T$ is the system matrix in the transform domain, and $c_T$ denotes the particle distribution in the transform domain. The idea is to choose $T$ in such a way that the matrix $S_T = ST$ has many small components that become zero after thresholding. It is well known that the DCT-I and -II are such kinds of transforms for MPI system matrices. Unfortunately, the DCT is a global transform on the spatial domain and offers no strategy for a multiresolution analysis (MRA). One transform to apply a MRA is the DWT [4], which can also be represented as a linear transform $T$ as well.

**III. Discrete wavelet transform**

Firstly, we describe how the one-dimensional discrete wavelet transform is related to a multiresolution analysis. Then we give a short overview of boundary handling for signals with finite support. Finally, we formulate the DWT as a matrix transform and explain how multidimensional signals can be processed.

**III.I. From multiresolution analysis to discrete wavelet transform**

In a multiresolution analysis, an arbitrary signal $x(t) \in L_2(\mathbb{R})$ can be approximated from lower to finer resolution by projecting it onto nested subspaces $V_{m+1} \subset V_m \subset L_2(\mathbb{R})$. A formal MRA definition can be found in [8].

In wavelet MRA, a space $V_{m-1}$ can be written as the direct sum of $V_m$ and a so-called wavelet space $W_m$: $V_{m-1} = V_m \oplus W_m$. (3)

The intersection between $V_m \cap W_m = \{0\}$ only includes the zero vector. Roughly speaking, the wavelet space $W_m$ contains the resolution lost when projecting a signal from $V_{m-1}$ onto the coarser resolution space $V_m$.

The space $V_m$ is spanned by scaled and time shifted versions of a function $\phi(t)$ according to $V_m = \text{span}\{\phi_{mn}(t) = \phi(2^{-m} t - n) | n \in \mathbb{Z}\}$. (4)

Additionally, the wavelet space $W_m$ is spanned by scaled and time shifted versions of a mother wavelet $\psi(t)$: $W_m = \text{span}\{\psi_{mn}(t) = \psi(2^{-m} t - n) | n \in \mathbb{Z}\}$. (5)

We can express the approximation by $V_m = V_{m+1} \oplus W_{m+1}$ and, equivalently, the signal $x_m(t) \in V_m$ as sum of the coarser approximated signal $x_{m+1}(t) \in V_{m+1}$ and the detail signal $y_{m+1}(t) \in W_{m+1}$.

Knowing $x_0(t) \in V_0$, one can compute the corresponding discrete coefficients $c_0(n)$, and from these, it is possible to calculate all discrete wavelet coefficients $d_m(n)$ recursively. Most times $c_0(n) \approx x_c(n) = x(nT)$ is assumed. More formally, for a given discrete sequence $c_m(n)$, the sequences $c_{m+1}(l)$ and $d_{m+1}(l)$ with $m, n, l \in \mathbb{Z}$ can be calculated by a two-channel filterbank with the discrete wavelet transform.

**Figure 1:** Discrete wavelet transform analysis filterbank with lowpass filter $h_0(n)$ and highpass filter $h_1(n)$.

**Figure 2:** Discrete wavelet transform synthesis filterbank with lowpass filter $g_0(n)$ and highpass filter $g_1(n)$.
analysis filter \( h_0(n) \) and \( h_1(n) \), where \( h_0(n) \) is a lowpass filter and \( h_1(n) \) a highpass filter (see Figure 1):

\[
c_{m+1}(\ell) = \sum_{n \in \mathbb{Z}} c_m(n)h_0(2\ell - n) \\
d_{m+1}(\ell) = \sum_{n \in \mathbb{Z}} d_m(n)h_1(2\ell - n).
\]

(6)

The discrete synthesis procedure of \( c_m(n) \) from \( c_{m+1}(\ell) \) and \( d_{m+1}(\ell) \), which both have the half sampling rate of \( c_m(n) \), is given by

\[
c_m(n) = \sum_{\ell \in \mathbb{Z}} c_{m+1}(\ell)g_0(n - 2\ell) + \sum_{\ell \in \mathbb{Z}} d_{m+1}(\ell)g_1(n - 2\ell).
\]

(7)

This corresponds to a discrete two-channel synthesis filterbank with the discrete synthesis filter coefficients \( g_0(n) \) and \( g_1(n) \) (see Figure 2). The impulse response \( g_0(n) \) belongs to a lowpass filter and \( g_1(n) \) belongs to a highpass filter.

How such filter coefficients \( h_0(n) \), \( h_1(n) \), \( g_0(n) \), and \( g_1(n) \) are designed is out of the scope of this paper. Most likely the filters are designed as finite impulse response (FIR) filters. In particular, in this work, we use the 9/7-biorthogonal wavelets, which can be found in [4].

### III.II. Boundary conditions

The previously introduced DWT is defined for infinitely long signals \( x(n) \) with \( n \in \mathbb{Z} \). However, many signals like images have finite support. A signal of length \( N \) can, for example, be defined as

\[
x = [x(0), x(1), \ldots, x(N-1)]^T.
\]

(8)

To extend the signal to infinite support, one naive boundary extension is the zero padding solution

\[
x_{zp}(n) = \begin{cases} x(n) & \text{for } n = 0, 1, \ldots, N - 1 \\ 0 & \text{elsewhere.} \end{cases}
\]

(9)

However, processing \( x_{zp}(n) \) will result in an expansive transform.

The second option is the periodic extension

\[
x_{pe}(n) = x(\text{mod}(n, N)).
\]

(10)

This processing allows nonexpansive processing for arbitrary FIR-filter coefficients \( h_0(n) \) and \( h_1(n) \) as long as \( N \) is an even number. However, the drawback is that large intensity jumps can be introduced when \( x(0) \) and \( x(N-1) \) differ significantly.

If \( h_0(n) \) and \( h_1(n) \) are linear-phase FIR-filters, a nonexpansive processing can also be achieved by a method known as symmetric reflection. Compared to the periodic extension, no artificial jumps are introduced, leading to better compression. It depends on the symmetries of \( h_0(n) \) and \( h_1(n) \) and the signal length (even or odd) how the reflection of the signal \( x(n) \) has to be performed in detail. Such an approach is used in this work. More information on the treatment of the boundaries can be found in [9–11]. In this work, the discrete wavelet transform was implemented using biorthogonal wavelets in a nonexpansive form for arbitrary signal length in lifting structure [12, 13].

### III.III. Mathematical expression of the DWT

Mathematically, the whole processing of one stage of the filterbank can be expressed with help of a matrix \( W = [W_{h_0}^T, W_{h_1}^T]^T \) with \( W \in \mathbb{R}^{N \times N} \), \( W_{h_0} \in \mathbb{R}^{(N/2) \times N} \), \( W_{h_1} \in \mathbb{R}^{N \times (N/2)} \), and

\[
x_{w} = Wx = \begin{bmatrix} x_{h_0} \\ x_{h_1} \end{bmatrix} = \begin{bmatrix} W_{h_0} \\ W_{h_1} \end{bmatrix} x,
\]

(11)

where \( x_{h_0} = W_{h_0}x \) and \( x_{h_1} = W_{h_1}x \) correspond to the lowpass and highpass filtered coefficients of the signal \( x(n) \), respectively. It is not surprising that also the processing of more than one stage of the one-dimensional DWT can be expressed as multiplication with a matrix.

### III.IV. The multidimensional DWT

For the multidimensional extension of the DWT, typically one level of the one-dimensional DWT is separately performed along each dimension of the multidimensional signal. The signal that was filtered with lowpass \( h_0(n) \) along all dimensions and downsampled will then form the basis for the next stage of the \( n \)-dimensional DWT decomposition, and so on. It should be mentioned that besides the separable wavelet transform considered here,
also non-separable versions of the wavelet transform exist, having their own pros and cons, like Curvelets [14] and Shearlets [15]. However, we exclude such types of transforms here.

IV. A joint MRA and matrix-compression approach

Although the DWT has a quite good compression performance for natural images, it is not as sparsifying for MPI system matrices as the DCT, which was used in [2, 16, 17] with good success. The advantage of the DWT, however, is that the lowpass filtered and downsampled coefficients of the system matrix can be interpreted as a coarser version of the system matrix at the finer level. This is nothing but the MRA condition of the DWT.

The idea of the method proposed in this article is now to combine the advantages of the DWT and DCT to develop a method that gives rise to an MRA for both the MPI system matrices and the particle distribution reconstruction. We decompose the system matrix level-wise by the d-dimensional DWT with respect to the dimensionality $d \in \{1, 2, 3\}$, followed by a DCT of the lowpass filtered coefficients of the system matrix in each level. The calculated system matrix wavelet coefficients then describe the transfer from the coarse level to a finer level.

In the publications [16], [18], and [19] it was observed that the system-matrix compression performance depends on the spatial symmetries of the frequency components. Furthermore, in [3] it was observed that compression can even be improved when the FOV is directly limited to the size of the FFP-trajectory. Our symmetric boundary conditions for the DWT help us to perversive the symmetries needed to get a maximum compression performance from the DCT, which implicitly assumes symmetric boundary conditions for spatial structure of the frequency components.

For demonstration purposes, the two-dimensional decomposition can be defined recursively by

$$
\begin{align*}
    S^\ell_{LL} &= S^{\ell-1}_{LL} W_{h_0 h_0}^T, \\
    S^\ell_{HL} &= S^{\ell-1}_{HL} W_{h_0 h_1}^T, \\
    S^\ell_{LH} &= S^{\ell-1}_{LH} W_{h_1 h_0}^T, \\
    S^\ell_{HH} &= S^{\ell-1}_{HH} W_{h_1 h_1}^T, \\
    \hat{S}^\ell &= S^{\ell-1}_{LL} C_{2D}^T,
\end{align*}
$$

where $S^{\ell}_{LL}$ is the lowpass filtered and downsampled sub-matrix, which will also be transformed by the two-dimensional DCT $C_{2D}$. The matrices $S^\ell_{HL}$, $S^\ell_{LH}$, and $S^\ell_{HH}$ include the bandpass filtered wavelet components of our system matrix. The transformation matrices $W_{h_0 h_0}$, $W_{h_0 h_1}$, $W_{h_1 h_0}$, and $W_{h_1 h_1}$ describe the filtering with the (low/low)-pass, (low/high)-pass, (high/low) and the (high/high)-pass of the filterbank, followed by downsampling with factor two in each dimension.

As an example, a two-dimensional decomposition is shown in Figure 3. For each stage of the two-dimensional DWT, the system matrix will be decomposed into four submatrices. The lowpass filtered version of the system matrix will then be transformed with the DCT to the compressive domain, and it will also form the basis for the next decomposition level of the spatial pyramid.

V. Level-wise reconstruction

The previously introduced formulation of the system matrix is independent of the used reconstruction method, so that the introduced formulation can be included in different solvers. We consider the Tikhonov regularized least-squares reconstruction problem

$$
\hat{c}^\ell = \arg \min_{c \in \mathbb{R}^{K^\ell}} \| S^\ell T^{-1} c - f \|_2^2 + \lambda^2 \| c \|_2^2
$$

with $S^\ell \in \mathbb{C}^{M \times K^\ell}$ being the compressed system matrix on decomposition stage $\ell$, $\lambda > 0$, and $T^{-1} f \in \mathbb{R}^{K_{\ell-1}}$ representing the DWT+DCT transform of level $\ell$, where $K_0 = N = N_x N_y N_z$ denotes the number of voxels of the particle distribution. The number of voxels of each subspace problem in (13) is $K^\ell = \left\lceil \frac{N}{2^\ell} \right\rceil \left\lceil \frac{N}{2^\ell} \right\rceil$ where $\lceil \cdot \rceil$ means the ceiling operator. We start the reconstruction on the coarse level $I_{\text{max}}$ and end on the finest level 0. The reconstructed SPIOs distribution from the coarser resolution is used as input for the next finer resolution stage by inserting it into the low-resolution components of the finer stage. The unknown high-resolution components were initialized with zeros for the iterative reconstruction. The reconstruction problem is solved by a variation of the fast iterative shrinkage thresholding (FISTA) [20]. Therefore, the problem (13) is reformulated as

$$
\hat{c}^\ell = \arg \min_{c \in \mathbb{R}^{K^\ell}} \| S^\ell T^{-1} c - f \|_2^2 + \lambda^2 \| c \|_2^2 + \lambda \text{I}_{\mathbb{R}^+_{K^\ell}}(c)
$$

where

$$
\text{I}_{\mathbb{R}^+_{K^\ell}}(c) = \begin{cases} 
0 & \text{for } c \in \mathbb{R}^+_{K^\ell} \\
\infty & \text{otherwise}
\end{cases}
$$

and

$$
\text{I}_{\mathbb{R}^+_{K^\ell}}(c) = \begin{cases} 
0 & \text{for } c \in \mathbb{R}^+_{K^\ell} \\
\infty & \text{otherwise}
\end{cases}
$$

denotes the indicator function for the non-negative subspace in $\mathbb{R}^+_{K^\ell}$. It should be mentioned that FISTA is more commonly known as solver for $\ell_1$-minimization problems, where $h(c) = \lambda |c|_1$ and the corresponding proximity operator is the soft-thresholding operator. However, in [20] the algorithm was developed in a more general
way, for arbitrary closed convex functions \( h(e) \). By making use of this fact, only an new proximity operator for \( h(e) \) in (15) has to be derived. It reads

\[
\text{prox}_h(e) = \arg \min_{u \in \mathbb{R}^K} h(u) + \frac{1}{2} \| u - c \|^2
\]

\[
= \left( \max \left\{ 0, \frac{1}{1 + \lambda^2 c^2} \right\} \right) \sum_{l=1}^K.
\]  

(16)

By replacing the soft-thresholding operator from the typical \( \ell_1 \)-minimization problem with the proximity operator in (16), the variation of FISTA for the non-negative least squares problem in (13) can be solved easily. The resulting algorithm can be seen as a projected gradient descent with a special adaptive choice of the step length. However, in coarse resolution levels, exact non-negative least squares solvers like the one in [21] can be used as well. For large-scale matrices \( S^l_\ell \), however, exact solvers can become quite slow.

VI. Dataset

We tested the approach on the simulated Lissajous-trajectory MPI system matrix dataset which was also used in [17]. For the simulated dataset, the first 62 frequency components were deleted. To speed up the reconstruction, the system matrix was globally thresholded, so that every sub-level system matrix retained a given percentage of its energy. In particular, we used the biorthogonal 9/7-filters to implement our wavelet transform. We reconstructed the SPIOs distribution with and without energy normalization of the rows of the subsystem matrix. The parameters were \( \lambda = 0.35 \) for the reconstruction with and \( \lambda = 10^{-4} \) for the reconstruction without normalization. Both regularization parameters were selected by hand. Regarding the chosen regularization parameter \( \lambda \), it should be mentioned that the DWT filter coefficients were used with the same scaling as in [13]. With this scaling, the biorthogonal transform becomes nearly energy preserving. Consequently, the parameter \( \lambda \) was chosen equally on all levels. More sophisticated strategies for choosing the regularization parameter may be possible, but are out of the scope of this publication. As stopping criterion, an upper limit of \( 10^{-4} \) for the relative change of the objective function and for the maximum number of iterations 3000 was set. In practice, the maximum number of iterations was seldom reached. Our processor was an Intel Core i7-3370K with 3.5 GHz and 8 logical kernels. We used Matlab R 2017b and did not employ parallel programming.

VII. Results

VII.I. Duration time for reconstruction

In a first test, the level-wise reconstruction was tested with regard to time consumption. Therefore, we repeated the experiment ten times and calculated the average execution time in seconds to reduce the influence of the different workloads on the processor. We measured the reconstruction time for the full 250 \( \times \) 250 resolution phantom from the coarsest to the finest resolution. For comparison reasons, also the reconstruction speed for the whole DCT compressed system matrix was tested. The results are shown in Figure 4. It can be observed that with higher compression (i.e., lower retained energy inside the system matrices), the reconstruction time was reduced. In Figure 4(a), only two decomposition steps were performed, whereas in Figure 4(b), four decomposition steps were performed.
Due to the noise inside the system matrix, a small compression helps to improve the reconstruction, because it helps to filter out noise from the system matrix. The energy normalized reconstruction shows, again, better reconstruction than the unnormalized one in Figure 5(a). In Figure 5(b), around the optimum, the performance of the MRA variants are nearly the same. Overall, the proposed algorithm improves the reconstruction performance. It should be noted that the reconstruction of the particle distribution with 98.5% retained energy for this system matrix is already significantly disturbed.

VII.III. Resolution phantom on different coarse levels

We exemplary show an example of SPIOs-distribution reconstruction on different levels in Figure 6. It can be recognized that the structure of the 250 × 250-pixel distribution is also identifiable inside the low-resolution reconstructions. It can be observed that the 63 × 63-pixel reconstruction still has a good separation between the particle spots and that the shapes are quite similar to the 250 × 250 high-resolution version. The 32 × 32-pixel reconstruction, however, has significant artifacts and deformations of the particle spots.

VIII. Discussion

The presented method uses a combination of two matrix compression techniques and gives rise to an MRA. It has the ability to first reconstruct the particle distribution on a coarse level, and, if more computational power is at hand, a high-resolution reconstruction can be performed. The compression is not a necessary part of this method. The method can also be used for a fast reconstruction inside a multiresolution analysis without compression of the system matrix. The MRA was able to speed up the reconstruction significantly when only a small number of coefficients was retained. Overall the MRA has the same quality level in terms of mean squared error as the DCT-based method. While only symmetric boundary conditions were used, also other ones, such as optimized boundary filters [22, 23], may be useful as well. A possible application of the developed method could be to find the support of the SPIOs distribution earlier on the finer grid, where the linear system’s condition is typically worse. In addition, an interesting scenario for the approach with its joint local and global transformations could be the use inside a compressed-sensing framework for the reconstruction of SPIO distributions. Since the DWT and DCT offer compression for both the system matrix and the image to be reconstructed, our developed transform offers us the best of two worlds: A good compressive transform for the system matrix and a

**Figure 5**: The mean squared error vs. the percentage of retained energy inside the system matrix for the combination of the DWT and DCT as well as for using the DCT only. The term Eng Norm means that an energy normalization is additionally performed for each frequency component. In (a) only two stages of the MRA decomposition were performed, whereas in (b) four stages were used.
good compressive transform for the particle distribution, which is highly promising for this purpose.

IX. Conclusions

We developed an MRA formulation for MPI based on the DWT and the DCT-II. We were able to show the efficiency of our approach, which offers the possibility to proceed step-wise from a coarse level to a high-resolution reconstruction of the SPIOs distribution. Our future research will be directed toward developing a compressed-sensing based reconstruction of the particle distributions using our MRA formulation of the system matrix.

Acknowledgments

This work was supported by the German Research Foundation under grant number ME 1170/7-1.

References

[10] J. Bradley, C. Brislawn, and V. Faber, Reflected boundary conditions for multirate filter banks, in Proceedings of the IEEE-SP Inter-

Figure 6: The SPIOs reconstruction of the MRA from left to right: 250 × 250, 63 × 63, and 32 × 32 pixels. The upper row was reconstructed with energy normalization, and the lower without. The images were clipped to [0, 2] level for visualization.


