MORPHOLOGICAL WAVELET TRANSFORM WITH ADAPTIVE DYADIC STRUCTURES

Zhen James Xiang and Peter J. Ramadge

Dept. of Electrical Engineering, Princeton University, Princeton NJ

ABSTRACT

We propose a two component method for denoising multidimensional signals, e.g. images. The first component uses a dynamic programing algorithm of complexity $\mathcal{O}(N \log N)$ to find an optimal dyadic tree representation of a given multidimensional signal of N samples. The second component takes a signal with given dyadic tree representation and formulates the denoising problem for this signal as a Second Order Cone Program of size $\mathcal{O}(N)$. To solve the overall denoising problem, we apply these two algorithms iteratively to search for a jointly optimal denoised signal and dyadic tree representation. Experiments on images confirm that the approach yields denoised signals with improved PSNR and edge preservation.

***Index Terms*—** Wavelet transforms, morphological operations, image enhancement, multidimensional signal processing, dynamic programming.

1. INTRODUCTION

The morphological wavelet transform [1, 2] is a non-linear wavelet transform that replaces the algebraic operations in the normal wavelet transform with $\max$ and $\min$ operators. This transform appears to better preserve morphological features (such as edges) in low resolution signals. We have recently shown, [3], that this property is closely connected to the accurate estimation of level sets using tree-based complexity regularizations [4]. Specifically, for a 1-D function the integral of the level set complexity measure of [4] over all levels yields the L_1 norm of the signal's Haar morphological wavelet coefficients. So L_1 regularization of the Haar morphological wavelet coefficients is equivalent to controlling the complexity of the function's level sets.

Here we explore the extension of this result to high dimensional signals and its application to image denoising. A major challenge in accomplishing this extension is that for multidimensional signals the choice of wavelet decomposition structure is not unique. Hence we first extend the result for any fixed dyadic tree signal representation. Then we pose and solve the problem of finding an optimal dyadic tree representation of a given multidimensional signal. Finally, we propose an iterative algorithm that jointly denoises a signal and finds the dyadic tree representation best adapted to the signal. The paper's contribution is to provide the theoretical justification for this approach and practical algorithms for its implementation.

The paper is organized as follows: After introducing required notation in §2, we establish the main theoretical and algorithmic results in §3. We then provide experiment comparison with other methods in §4 and conclude in §5.

2. PRELIMINARIES

The nodes of a binary tree are indexed by pairs of integers (k, l), $k, l \geq 0$. The root has index $(0, 0)$ and node (k, l) has two children with indices $(k + 1, 2l)$ and $(k + 1, 2l + 1)$.

Fix integers $d, m \geq 1$. Without loss of generality, we consider signals defined on the domain $[0, 1]^d$.

Definition 1. *A binary tree $\mathfrak{T}$ is called a "full dyadic tree" on $[0, 1]^d$ with resolution 2^m if it satisfies each of the following:*
(c1) The root node of $\mathfrak{T}$ is $J_{0,0} = [0, 1]^d$.
(c2) Any non-leaf node $J_{k,l}$ has color $c \in \{1, 2, \ldots, d\}$. Its children, $J_{k+1,2l}$ and $J_{k+1,2l+1}$, are congruent hyperrectangles obtained by cutting $J_{k,l}$ with the $d-1$ dimensional hyperplane through the center of $J_{k,l}$ and perpendicular to axis-c.
(c3) All leaf nodes are hypercubes of side length 2^{-m}.

By (c3), $\mathfrak{T}$ has height dm and by (c2), each hyperrectangle at depth k has volume $|J_{k,l}| = 2^{-k}$ $(0 \leq k \leq dm)$.

Let $\mathfrak{T}$ be a full dyadic tree on $[0, 1]^d$ with resolution 2^m. A signal $f : [0, 1]^d \mapsto \mathbb{R}$ has resolution 2^m if f is constant on every leaf node of $\mathfrak{T}$. For any threshold γ consider the level set $S_\gamma = \{x \in [0, 1]^d : f(x) \geq \gamma\}$. We can use $\mathfrak{T}$ as a decision tree to decide if $x \in S_\gamma$ since each leaf of $\mathfrak{T}$ is entirely inside or outside S_γ. By recursively merging siblings that are both inside or outside S_γ with their parent, we can prune $\mathfrak{T}$ to T_γ, the minimum decision tree required to decide if $x \in S_\gamma$. We let $\Phi(T_\gamma)$ denote a general form of the level set complexity regularizer proposed in [4]: $\Phi(T_\gamma) = \sum_{L \in \pi(T_\gamma)} \phi(|L|)$ where $\pi(T_\gamma)$ is the set of leaves of T_γ, and $\phi(|L|)$ only depends on the size (i.e., the depth) of the leaf L. To make this dependency explicit, let $\alpha_k = \phi(2^{-k})$ $(k = 0, 1, \ldots, dm)$. We assume $\alpha_0 = 0$ and $2\alpha_{k+1} > \alpha_k$ [3].

We can also compute a morphological Haar wavelet transform [2] of f in a bottom-up fashion on the dyadic structure $\mathfrak{T}$. We associate leaf nodes $J_{dm,l}$ with the highest resolution signal $\Psi_{dm}(l) = f(x)|_{x \in J_{dm,l}}$. Then we move up the tree and associate every non-leaf node $J_{k,l}$ with a value $\Psi_k(l)$ calculated from values of its children: $\Psi_{k+1}(2l)$ and $\Psi_{k+1}(2l+1)$ (see [1, 2, 3]). There are two versions of this transform, differentiated by superscripts $\vee$ (max-version) and $\wedge$ (min-version), yielding two sets of wavelet coefficients $W^{\vee}_{k,l}$ and $W^{\wedge}_{k,l}$.

3. THEORETICAL RESULTS

Our first result is an multidimensional extension of a 1-D result first proposed in [3].

Theorem 1. *Assume a fixed dyadic structure $\mathfrak{T}$ and a maximum resolution 2^m. Using the notation introduced above:*

$$\int \Phi(T_\gamma) d\gamma = \sum_{k=0}^{dm-1} \sum_{l=0}^{2^k-1} \alpha_{k+1} \left(|W^{\vee}_{k,l}| + |W^{\wedge}_{k,l}|\right). \quad (1)$$

Proof. Due to limited space we simply note that the result follows by transforming f to an 1-D sequence $\Psi_{dm}(l) = f(x)|_{x \in J_{dm,l}}$, then appealing to the main result in [3]. $\square$

The theorem connects the integral of the level set complexity measure and the L_1 norm of the nonlinear wavelet coefficients both obtained under a common dyadic structure. We now explore some of the ramifications of this connection. We begin with the obvious question: What is the best dyadic structure to use? Or equivalently: What is the best order in which to perform the morphological wavelet decomposition of a multidimensional signal? One could simply prescribe a fixed $\mathfrak{T}$. However, (1) suggests an interesting alternative: select $\mathfrak{T}$ that is best adapted to the signal f. To explore this idea, let $E_{m,\mathfrak{T}}(f) = \int \Phi(T_\gamma) d\gamma$. A good full dyadic representation of f should yield simple pruned trees T_γ that compactly represent the level sets of f. So smaller values of $E_{m,\mathfrak{T}}(f)$ indicate that $\mathfrak{T}$ is better adapted to the structure of f.

3.1. Optimal Dyadic Structure for Fixed f

The above discussion leads us to pose the following optimization problem:

$$\mathfrak{T}_{opt} \in \arg\min_{\mathfrak{T}} \left\{E_{m,\mathfrak{T}}(f)\right\}. \quad (2)$$

We will show how to use dynamic programming to solve (2) for a given function f. For hyperrectangle B in a full dyadic tree at depth k_B, define a "**sub-partition tree**" $\mathfrak{T}_B$ as any binary tree with root node B that satisfies conditions (c2) and (c3) in **Definition 1**. If k indexes the depth in the full dyadic tree, then $s(k) = k - k_B$ is the depth in the sub-partition tree $\mathfrak{T}_B$. The morphological Haar wavelet coefficients $W_{s(k),l}$ of f are calculated on $\mathfrak{T}_B$. Define the "**partial loss**" of $\mathfrak{T}_B$ as:

$$L(\mathfrak{T}_B) = \sum_{k=k_B}^{dm-1} \sum_{l=0}^{2^{s(k)}-1} \alpha_{k+1} \left(|W^{\vee}_{s(k),l}| + |W^{\wedge}_{s(k),l}|\right). \quad (3)$$

(3) is defined to be 0 for $k_B = dm$. Starting from leaves B (hypercubes of side length 2^{-m}), for which $L(\mathfrak{T}_B)$ is 0, we recursively solve for the optimal sub-partition tree $\mathfrak{T}_{B,opt} \in \arg\min_{\mathfrak{T}_B} L(\mathfrak{T}_B)$ and store the solutions in dictionary $\mathcal{D}_{k_B}$ as entry:

$$a(B) = \left(\max_{x\in B} f,\ \min_{x \in B} f,\ \mathfrak{T}_{B,opt},\ L(\mathfrak{T}_{B,opt})\right).$$

For non-leaf B, assume the color of $\mathfrak{T}_B$'s root node is c, this cuts B into hyperrectangles B^c_U and B^c_V. Connecting the sub-partition trees rooted at B^c_U and B^c_V with new root node B yields tree $\mathfrak{T}_B$. We visually represent this as $\mathfrak{T}_B = \{\mathfrak{T}_{B^c_U} \leftarrowtail B \rightarrowtail \mathfrak{T}_{B^c_V}\}$. By optimizing over c, the minimum value of $L(\mathfrak{T}_B)$ can be recursively computed as:

$$\min_c \left\{L(\mathfrak{T}_{B^c_U,opt}) + L(\mathfrak{T}_{B^c_V,opt}) + \alpha_{k_B+1}\left(|W^{\vee}_B| + |W^{\wedge}_B|\right)\right\}.$$

$W^{\vee}_B$ and $W^{\wedge}_B$ also depend on c. Nevertheless, the expression inside the curly bracket can be calculated using entries $a(B^c_U)$ and $a(B^c_V)$ in $\mathcal{D}_{k_B+1}$. Hence we denote this as $g(a(B^c_U), a(B^c_V))$. Once we have completed the dictionary $\mathcal{D}_{k_B+1}$, the optimal sub-partition tree for all hyperrectangles at depth k_B can be easily computed. Thus we can search for the best sub-partition tree using a bottom-up dynamic programming algorithm (displayed as Algorithm 1).

To analyze the complexity of this algorithm, we first note that $\sum_{k=0}^{dm} |\mathcal{D}_k| = (2^{m+1}-1)^d$. This holds because any hyperrectangle is the direct product of d 1-D intervals, each of which has $2^{m+1}-1$ different choices. We can implement dictionary $\mathcal{D}_k$ such that the search and insert operations are $\mathcal{O}(\log|\mathcal{D}_k|)$. The time complexity is thus:

$$\begin{aligned}&|\mathcal{D}_{dm}|\log|\mathcal{D}_{dm}| + d\textstyle\sum_{k=0}^{dm-1}|\mathcal{D}_{k+1}|C(\log|\mathcal{D}_k|+\log|\mathcal{D}_{k+1}|)\\ &< dC'(\textstyle\sum_{k=0}^{dm}|\mathcal{D}_k|)\log(\textstyle\sum_{k=0}^{dm}|\mathcal{D}_k|) = \mathcal{O}(md^2 2^d 2^{md}),\end{aligned}$$

and the space complexity is: $(\log(2^{m+1}-1)^d + \log d^{dm} + 3*b)\sum_{k=0}^{dm}|\mathcal{D}_k| < C(d(m+1) + dm\log d)(2^{m+1}-1)^d < \mathcal{O}(md\log d 2^d 2^{md})$, ($b$ is the number of bits to represent a float), because each entry of $\mathcal{D}_k$ takes $\log(2^{m+1}-1)^d + b + b + \log d^{dm} + b$ bits to store. For image signals of $N = 2^{md}$ pixels/voxels, the time and space complexity is $\mathcal{O}(N\log N)$ and $\mathcal{O}(N)$, respectively.

3.2. Signal Denoising with Fixed $\mathfrak{T}$

We now apply these ideas to signal denoising. First consider the case of fixed $\mathfrak{T}$. Given a noisy observation $\tilde{f}$ of f and a specific dyadic structure $\mathfrak{T}$ we can estimate f by finding:

$$\hat{f} = \arg\min_f \left\{||f - \tilde{f}||_2^2 + \lambda E_{m,\mathfrak{T}}(f)\right\}, \quad (4)$$

i.e., $\hat{f}$ should be close to $\tilde{f}$ but have low complexity w.r.t. $\mathfrak{T}$. This can be posed as a SOCP (Second Order Cone Programming) problem. To see this, first introduce some dummy

Algorithm 1 Algorithm to Solve Problem (2)

```
For all leafs $B$ store the following in the dictionary $\mathcal{D}_{dm}$:
    $a(B) = (f(x)|_{x\in B},\ f(x)|_{x\in B},\ \{B\},\ 0)$.
for $k = dm-1, \dots, 1, 0$ do
  Initialize $\mathcal{D}_k = \phi$.
  for all $a(U) \in \mathcal{D}_{k+1}$ do
    for $c = 1, 2, \dots, d$ do
      if The $c$-th dimensional side length of $U$ is 1 then
        continue
      end if
      $V \leftarrow$ The unique hyperrectangle that could be the
      sibling of $U$ under a parent node $B$ of color $c$.
      $B \leftarrow U \cup V$.
      Retrieve entries $a(U)$ and $a(V)$ from $\mathcal{D}_{k+1}$:
        $a(U) = (\max_U,\ \min_U,\ \mathfrak{T}_{U,opt},\ L(\mathfrak{T}_{U,opt}))$,
        $a(V) = (\max_V,\ \min_V,\ \mathfrak{T}_{V,opt},\ L(\mathfrak{T}_{V,opt}))$.
      $\max_B \leftarrow \max\{\max_U, \max_V\}$,
      $\min_B \leftarrow \min\{\min_U, \min_V\}$,
      $\mathfrak{T}_{B,opt} \leftarrow \{\mathfrak{T}_{U,opt} \leftarrow B \rightarrow \mathfrak{T}_{V,opt}\}$,
      $L(\mathfrak{T}_{B,opt}) \leftarrow g(a(U), a(V))$
      Retrieve the existing entry for $B$ from $\mathcal{D}_k$:
        $a'(B) = (\max'_B,\ \min'_B,\ \mathfrak{T}'_{B,opt},\ L(\mathfrak{T}'_{B,opt}))$.
      if $a'(B) = \phi$ or $L(\mathfrak{T}_{B,opt}) < L(\mathfrak{T}'_{B,opt})$ then
        Replace $a'(B)$ with the entry:
         $a(B) = (\max_B,\ \min_B,\ \mathfrak{T}_{B,opt},\ L(\mathfrak{T}_{B,opt}))$.
      end if
    end for
  end for
end for
Retrieve $a([0,1]^d)$ from $\mathcal{D}_0$ and output $\mathfrak{T}_{[0,1]^d,opt}$.
```

variables. For each leaf $J_{dm,l}$ of $\mathfrak{T}$, let $\varepsilon_l = (f - \tilde{f})|_{x\in J_{dm,l}}$ and let e denote the total L_2 loss. So e satisfies the constraint:

$$e \geq \sum_{l=0}^{2^{dm}-1} \varepsilon_l^2 \Longleftrightarrow (\tfrac{e+1}{2}, \tfrac{e-1}{2}, \varepsilon_0, \dots, \varepsilon_{2^{dm}-1}) \geq_{\mathcal{K}} 0 \quad (5)$$

where $\geq_{\mathcal{K}}$ denotes a second order cone (Lorentz cone) inequality. For every edge b of $\mathfrak{T}$ connecting parent J_p with child J_c, define non-negative dummy variables $\delta_b^{\vee} = (\max_{x\in J_p} f(x) - \max_{x\in J_c} f(x))$ and $\delta_b^{\wedge} = -(\min_{x\in J_p} f(x) - \min_{x\in J_c} f(x))$. These dummy variables and ε_l have the following linear dependencies. For any non-leaf node $J_{k,l}$, its "left path" $\mathcal{P}_{k,l}^U$ starts with $J_{k,l}$, always proceeds to *left* children, and finally reaches leaf node $\mathcal{L}_{k,l}^U = J_{dm,\theta_{k,l}^U}$. The "right path" $\mathcal{P}_{k,l}^V$, $\mathcal{L}_{k,l}^V$ and $\theta_{k,l}^V$ are defined likewise. For any pair (k,l) with $0 \leq k \leq dm-1$ and $0 \leq l \leq 2^k - 1$, the following equalities must hold:

$$\varepsilon_{\theta_{k,l}^U} + \tilde{f}|_{\mathcal{L}_{k,l}^U} + \sum_{b\in\mathcal{P}_{k,l}^U} \delta_b^{\vee} = \varepsilon_{\theta_{k,l}^V} + \tilde{f}|_{\mathcal{L}_{k,l}^V} + \sum_{b\in\mathcal{P}_{k,l}^V} \delta_b^{\vee}, \quad (6)$$

$$\varepsilon_{\theta_{k,l}^U} + \tilde{f}|_{\mathcal{L}_{k,l}^U} - \sum_{b\in\mathcal{P}_{k,l}^U} \delta_b^{\wedge} = \varepsilon_{\theta_{k,l}^V} + \tilde{f}|_{\mathcal{L}_{k,l}^V} - \sum_{b\in\mathcal{P}_{k,l}^V} \delta_b^{\wedge}, \quad (7)$$

since the equations in (6), (7) sum to $\max_{x\in J_{k,l}} f(x)$ and $\min_{x\in J_{k,l}} f(x)$ respectively.

We now formulate the SOCP problem:

$$\begin{aligned} &\text{minimize} && e + \lambda \sum_{k=0}^{dm-1} \sum_{\substack{\text{any edge } b \text{ that} \\ \text{starts at depth } k}} \alpha_{k+1} (\delta_b^{\vee} + \delta_b^{\wedge}) \\ &\text{subject to} && \text{constraints (5), (6), (7), and } \delta_b^{\vee}, \delta_b^{\wedge} \geq 0. \end{aligned} \quad (8)$$

Constraint (5) has to be equality to achieve optimality. So $e = ||f - \tilde{f}||_2^2$. If b_U and b_V are two edges at depth k that connect $J_{k,l}$ with its children, it is easy to verify that to achieve optimality, one of $\delta_{b_U}^{\vee}$ and $\delta_{b_V}^{\vee}$ has to be 0. Therefore $\delta_{b_U}^{\vee} + \delta_{b_V}^{\vee} = |\delta_{b_U}^{\vee} - \delta_{b_V}^{\vee}| = |W_{k,l}^{\vee}|$. Similarly $\delta_{b_U}^{\wedge} + \delta_{b_V}^{\wedge} = |W_{k,l}^{\wedge}|$. So (8) is equivalent to (4). As a SOCP of size $\mathcal{O}(2^{dm}) = \mathcal{O}(N)$, (8) can be solved efficiently with existing SOCP toolboxes.

3.3. Signal Denoising with Adaptive Dyadic Structure

Finally, we consider the problem of denoising while jointly determining the dyadic structure best adapted to the signal. Specifically, given a noisy observation $\tilde{f}$ of the signal f, we now consider the adaptive estimator:

$$(\hat{f}, \hat{\mathfrak{T}}) \in \arg\min_{(f,\mathfrak{T})} \left\{ ||f - \tilde{f}||_2^2 + \lambda E_{m,\mathfrak{T}}(f) \right\}, \quad (9)$$

where $\lambda > 0$ is a regularization coefficient. Notice that we add $E_{m,\mathfrak{T}}(f)$ as a regularizer to the L_2 loss, hence controlling the complexity of the denoised signal under *its best* representation. (9) is a complex problem due to the non-linearity of $E_{m,\mathfrak{T}}(f)$ and the combinatorial choices of $\mathfrak{T}$. To solve the problem, we propose an iterative scheme that alternates between optimizing $\hat{\mathfrak{T}}$ with $\hat{f}$ fixed and $\hat{f}$ with $\hat{\mathfrak{T}}$ fixed:

$$\text{T step: } \hat{\mathfrak{T}} \quad \leftarrow \quad \arg\min_{\mathfrak{T}} \left\{ ||\hat{f} - \tilde{f}||_2^2 + \lambda E_{m,\mathfrak{T}}(\hat{f}) \right\}. \quad (10)$$

$$\text{F step: } \hat{f} \quad \leftarrow \quad \arg\min_{f} \left\{ ||f - \tilde{f}||_2^2 + \lambda E_{m,\hat{\mathfrak{T}}}(f) \right\}. \quad (11)$$

We have presented algorithms to solve (10) and (11). Acting together they offer an approach to solving (9). This may only yield a local minima because (9) is not convex.

4. EXPERIMENTS

We tested the proposed joint denoising and optimal representation algorithm (10)-(11) on seven grayscale images (Figure 1) corrupted with additive white Gaussian noise and compared the results with: (a) non-adaptive denoising using fixed dyadic structure (solving (4) using prescribed $\mathfrak{T}$), (b) standard linear wavelet denoising using soft or hard thresholding and (c) total variation denoising [5]. For (a) and (b), we divided each image into 64×64 blocks, applied the denoising algorithms to each block, and incorporated partial cycle spinning

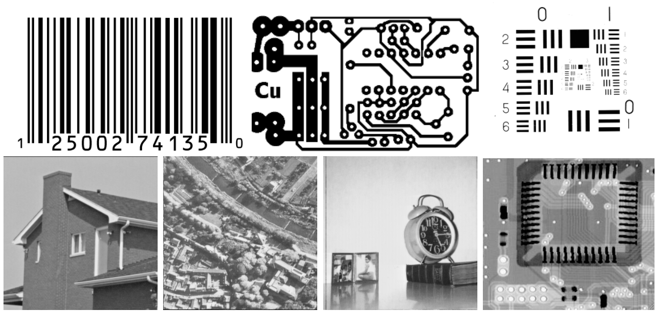

Fig. 1. **Test Images:** (1) Barcode, (2) PCB Board, (3) Resolution Chart, (4) House, (5) Aerial, (6) Clock, (7) PCB Board X-ray.

Image ID (Input PSNR)	Hard Thresholding	Non-adaptive Version	Total Variation	Proposed Method
1 (10 dB)	19.92 (Haar)	20.34 (Col.)	18.69	**24.68**
2 (10 dB)	16.76 (Haar)	16.34 (Mixed)	17.45	**18.25**
3 (10 dB)	20.38 (Haar)	19.17 (Mixed)	19.98	**21.92**
4 (20 dB)	30.08 (Haar)	29.12 (Mixed)	29.92	**30.35**
5 (20 dB)	24.80 (db3)	**25.36** (Mixed)	25.27	25.29
6 (20 dB)	29.23 (Haar)	28.21 (Mixed)	28.74	**29.42**
7 (30 dB)	33.31 (db3)	33.13 (Mixed)	33.31	**33.47**

Table 1. Best PSNR in dB for different methods. For hard thresholding we report the best wavelet type in parentheses; For Non-adaptive version we report the best dyadic structure in parentheses.

by averaging the results of 64 shifted, processed, and inverse shifts of each block (shifts: $(i, j), 0 \leq i, j \leq 7$ pixels). For our algorithm and for fixed $\mathfrak{T}$, we set $\{\alpha_k = 2^{-(1-s)k}\}$ and used cross-validation to find the best λ and s. For fixed $\mathfrak{T}$, we used cross-validation to find the best $\mathfrak{T}$ among: column-wise, row-wise or a fixed mixed decomposition structure. For linear wavelets, we used cross-validation to choose thresholds and the best wavelet among db1(Haar), db2-db5. In each case, the PSNR achieved is given in Table 1. Hard thresholding always outperformed soft thresholding. The proposed method slightly outperformed linear wavelet denoising methods and total variation method in all images (Table 1), and subjectively appears to better preserve edges (Figure 2). Using the adaptive dyadic structure is usually better than the prescribed choices of $\mathfrak{T}$, but may occasionally over-fit a noisy image.

5. CONCLUSIONS

We have given a multi-dimensional generalization of the equivalency between L_1 regularization of morphological wavelets and the level set estimation method of [4]. This led us to propose an adaptive denoising problem (9) which we solved using a dynamic programming algorithm of complexity $\mathcal{O}(N \log N)$ and a SOCP of size $\mathcal{O}(N)$, where N is the number of image pixels/voxels. Experimental results indicated that this denoising scheme improves estimation (higher PSNRs) and better preserves edge information. The latter may be important when edge information is critical.

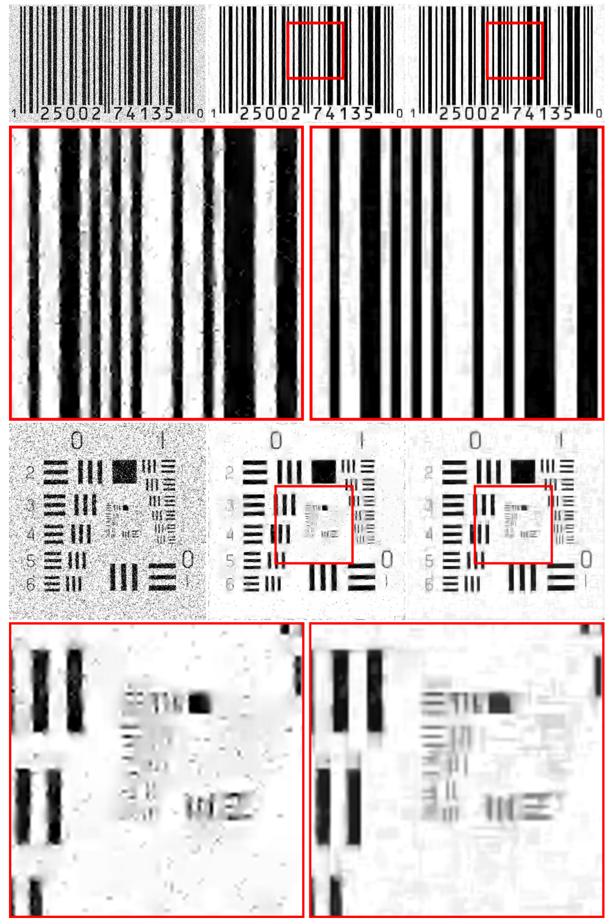

Fig. 2. Denoising results for images (1) and (3). 5 images/example: noisy image, denoising by linear wavelet with hard thresholding, denoising by proposed method, and zoom-in detail of the latter two.

6. REFERENCES

[1] J. Goutsias and H. Heijmans, "Nonlinear multiresolution signal decomposition schemes-Part I: Morphological pyramids," *IEEE Transactions on image processing*, vol. 9, no. 11, pp. 1862–1876, 2000.

[2] H. Heijmans and J. Goutsias, "Nonlinear multiresolution signal decomposition schemes-Part II: Morphological wavelets," *IEEE Transactions on Image Processing*, vol. 9, no. 11, pp. 1897–1913, 2000.

[3] Z.J. Xiang and P.J. Ramadge, "Morphological wavelets and the complexity of dyadic trees," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010.

[4] R.M. Willett and R.D. Nowak, "Minimax optimal level-set estimation," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2965–2979, 2007.

[5] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1-4, pp. 259–268, 1992.