An adaptive hierarchical sensing scheme for sparse signals

Henry Schütze, Erhardt Barth, Thomas Martinetz

Institute for Neuro- and Bioinformatics
University of Lübeck
Ratzeburger Alle 160
23562 Lübeck, Germany

ABSTRACT

In this paper, we present Adaptive Hierarchical Sensing (AHS), a novel adaptive hierarchical sensing algorithm for sparse signals. For a given but unknown signal with a sparse representation in an orthogonal basis, the sensing task is to identify its non-zero transform coefficients by performing only few measurements. A measurement is simply the inner product of the signal and a particular measurement vector. During sensing, AHS partially traverses a binary tree and performs one measurement per visited node. AHS is adaptive in the sense that after each measurement a decision is made whether the entire subtree of the current node is either further traversed or omitted depending on the measurement value. In order to acquire an N-dimensional signal that is K-sparse, AHS performs $\mathcal{O}(K \log N/K)$ measurements. With AHS, the signal is easily reconstructed by a basis transform without the need to solve an optimization problem. When sensing full-size images, AHS can compete with a state-of-the-art compressed sensing approach in terms of reconstruction performance versus number of measurements. Additionally, we simulate the sensing of image patches by AHS and investigate the impact of the choice of the sparse coding basis as well as the impact of the tree composition.

Keywords: Compressed sensing, compressive imaging, adaptive sampling, sparse signal recovery, gist

1. INTRODUCTION

Natural images can be encoded sparsely, meaning that they can be represented with rather few coefficients in an appropriate basis.[1,2] This is a prerequisite that facilitates efficient sampling of visual scenes by compressed sensing (CS), a signal acquisition principle[3–5] whose theory has rapidly emerged during the last decade. With CS, it is possible to simultaneously compress and acquire an unknown signal of interest. Given the unknown signal, a certain number of measurements is recorded. A measurement, which can be considered as a filter operation, is the inner product of the signal and a predefined measurement vector. The objective is to approximate the signal accurately by performing only few informative measurements. If the signal is K-sparse in a specific orthogonal basis and the matrix of measurement vectors satisfies certain incoherence conditions, e.g. the restricted isometry property,[3,6] then the number of measurements required for a faithful signal reconstruction is of order $\mathcal{O}(K \log N/K)$. The signal is typically reconstructed from the collected measurement values by solving an optimization problem, namely by seeking for the sparsest solution of an underdetermined system of linear equations.

Here, we address a similiar sensing task, for which the sparse coding basis is assumed to be known in advance. This is reasonable e.g. for the class of natural images or other signal classes, for which sparse coding bases are already known or can be learned[7,8] from training data. Let $\vec{x} \in \mathbb{R}^N$ denote the given but unknown signal that we wish to acquire. We assume that $\vec{x}$ has a sparse representation $\vec{a} = \mathbf{\Psi}^T \vec{x}$ in the orthonormal basis $\mathbf{\Psi} \in \mathbb{R}^{N \times N}$ ($\mathbf{\Psi}^T \mathbf{\Psi} = \mathbf{I}_N$), meaning that the signal energy is spread over only few coefficients. We denote $\vec{x}$ to be K-sparse in $\mathbf{\Psi}$ if exactly K entries of $\vec{a}$ are non-zero ($\|\vec{a}\|_0 = K$). The sensing task is to identify the non-zero signal coefficients in the basis $\mathbf{\Psi}$ by performing preferably few measurements $y_i = \vec{\phi}_i^T \vec{x}$, where we denote $\vec{\phi}_i \in \mathbb{R}^N$ as a measurement vector.

Further author information: (Send correspondence to Henry Schütze via schuetze@inb.uni-luebeck.de)

Published in Proceedings SPIE Electronic Imaging, Vol. 9014, Human Vision and Electronic Imaging XIX, 2014.

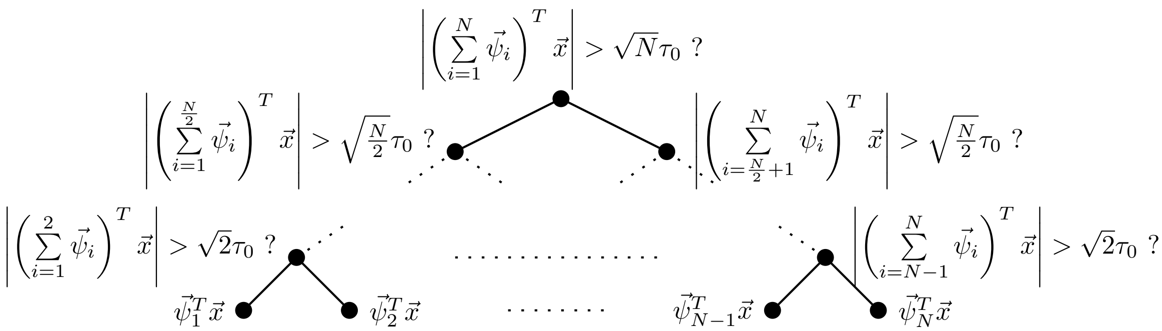


Figure 1: Schematic illustration of an AHS tree. During sensing, AHS partially traverses the tree. For each visited node, one measurement, i.e., the inner product of the unknown signal $\vec{x}$ with a particular measurement vector, is conducted.

We propose the novel sensing algorithm Adaptive Hierarchical Sensing (AHS). From the given sparse coding basis, a binary tree is constructed that assigns a measurement vector to each node. AHS partially traverses the tree and for each visited node the corresponding measurement value with the signal is recorded. Based on the measurement value at each node a decision is made, whether the subtree of the current node should be further traversed or not.

In contrast to CS, the new algorithm is highly adaptive because the selection of the next measurement vector depends on the previous measurement values. The tree gives a hierarchy that deterministically defines the measurement vectors and avoids random measurements. Moreover, solving an optimization problem for the reconstruction is unnecessary, because the signal approximation is obtained simply by a basis transform with $\mathbf{\Psi}$.

Deutsch et al. addressed a similar sensing task but their approach is limited to sparse representations within wavelet bases. They proposed the algorithm Adaptive Direct Sampling (ADS) that directly samples relevant transform coefficients of an image in a specific wavelet domain.[9] The approach uses a heuristic based on the Lipschitz exponent. ADS starts to sample all coefficients in all subbands assigned to the coarsest scales. For each spatial location and subband, the corresponding coefficients of the next finer scale are sampled depending on the slope of spatially corresponding coefficients across already sampled (coarser) scales.

Note that an alternative hierarchical sensing scheme is presented by Burciu et al.[10] in this volume.

Finally, a further motivation for our work is due to the fact that human vision seems to involve related sensing schemes - see section 4.

2. ADAPTIVE HIERARCHICAL SENSING (AHS)

2.1 The AHS algorithm

In the following, we assume that the signal dimensionality N is a power of 2. The fundamental AHS data structure is a binary tree with N leaves which we denote as AHS tree in the following. Given a particular basis $\mathbf{\Psi}$ that provides sparse signal representations, each leaf i of the AHS tree ($i \in \mathcal{N} = \{1, ..., N\}$) represents one particular basis vector $\vec{\psi}_i \in \mathbb{R}^N$ as a measurement vector. Any other node of the AHS tree represents a measurement vector which is composed by the sum of measurement vectors that belong to its two descendant nodes. Figure 1 illustrates such an AHS tree. If N is not a power of 2, a balanced AHS tree can nevertheless be constructed and used for sensing.

AHS works in the following fashion: to acquire a given but unknown signal $\vec{x}$, the AHS tree is traversed starting from the root. At each visited node, one measurement is performed, namely the inner product of $\vec{x}$ with the measurement vector associated with that node. If the magnitude of the measurement value is larger than a threshold, the traversion is continued for the subtree of that node. Otherwise, the entire subtree of the node remains unvisited. Each visited leaf provides one transform coefficient of $\vec{x}$ because the corresponding measurement vector is a basis vector of $\mathbf{\Psi}$. Finally, these measurement values recorded at the leaves can be used to reconstruct the signal. For a pseudocode[*] formulation of AHS, see Algorithm 1.

Algorithm 1 Adaptive Hierarchical Sensing (AHS)

Input: unknown signal $\vec{x} \in \mathbb{R}^N$,
  binary sensing (sub)tree represented by leaf index vector $\vec{l} = (l_1, ..., l_L)^T$
  orthogonal basis $\mathbf{\Psi} \in \mathbb{R}^{N \times N}$,
  canonical threshold $\tau_0 \geq 0$
Output: approximate coefficients $\vec{\hat{a}} \in \mathbb{R}^L$ according to $\vec{l}$
Create measurement vector $\vec{\phi} \leftarrow \sum_{i=1}^{L} \psi_{l_i}$
Perform one particular measurement $y \leftarrow \vec{\phi}^T \vec{x}$
if $L > 1$ **then**
  if $|y| > \sqrt{L}\tau_0$ **then**
    $\vec{\hat{a}}_I \leftarrow \text{AHS}(\vec{x}, (l_1, ..., l_{\lfloor L/2 \rfloor})^T, \mathbf{\Psi}, \tau_0)$
    $\vec{\hat{a}}_{II} \leftarrow \text{AHS}(\vec{x}, (l_{\lfloor L/2 \rfloor + 1}, ..., l_L)^T, \mathbf{\Psi}, \tau_0)$
    $\vec{\hat{a}} \leftarrow \left(\vec{\hat{a}}_I^T, \vec{\hat{a}}_{II}^T\right)^T$
  else
    $\vec{\hat{a}} \leftarrow \vec{0}_L$
  end if
else
  $\vec{\hat{a}} \leftarrow y$
end if

If the transform coefficients of an AHS subtree with a leaf set $\mathcal{J} \subset \mathcal{N}$ do not carry any signal energy, i.e. $\forall j \in \mathcal{J},\ \vec{\psi}_j^T \vec{x} = 0$, then the measurement at the root of this subtree gives $y = \vec{\phi}^T \vec{x} = (\sum_{j \in \mathcal{J}} \vec{\psi}_j)^T \vec{x} = 0$. Even if the transform coefficients are contaminated with small amplitude noise, the magnitude of the measurement value, $|y|$, is relatively small. Please note that the converse is not necessarily true, meaning that a small measurement value does not necessarily imply exclusively unrelevant transform coefficients. In unfavorable cases, the negative and positive terms of the sum may cancel each other out or yield a low absolute measurement below the threshold, although relevant coefficients are present. Such a case would result in the wrong decision to omit a subtree and hence in a loss of information.

The threshold is slightly adapted depending on the position of the current node in the AHS tree. A multiplication of the canonical threshold τ_0 by $\sqrt{L}$ attempts to capture the signal energy relative to the number of leaves, i.e. relative to the number of coefficients in consideration.

2.2 Signal reconstruction

When the traversion of the AHS tree is completed, each visited leaf i has provided one transform coefficient a_i of the signal, i.e. one entry of the true coefficient vector $\vec{a} = \mathbf{\Psi}^T \vec{x}$. In order to reconstruct the signal $\vec{x}$, an approximate coefficient vector $\vec{\hat{a}} \in \mathbb{R}^N$ is created, containing entries $\hat{a}_i = \vec{\psi}_i^T \vec{x}$ for each visited leaf i and zero for each unvisited leaf. Then, the signal approximation $\vec{\hat{x}}$ is merely obtained by the basis transform $\vec{\hat{x}} = \mathbf{\Psi}\vec{\hat{a}}$. Note that, in contrast to CS, solving an inverse optimization problem is unnecessary. If $\vec{x}$ is K-sparse in $\mathbf{\Psi}$, it can be perfectly reconstructed, if every leaf that corresponds to a non-zero coefficient is visited. If some small coefficients are missed, one obtains an approximate reconstruction.

2.3 Sensing performance

The sensing performance is considered to be high, if merely few measurements are performed and the approximation error is low. How many sensing actions AHS eventually conducts depends on several conditions.

If the signal is not strictly sparse (e.g. if noise is present) the applied threshold is crucial. A rather low threshold tends to increase the total number of sensing actions. It leads to higher sensitivity because subtrees

---

[*]The AHS pseudocode contains minor corrections of the original version, i.e. the indexing of the subtree representation as well as the order of the two conditional statements.

are traversed whose leaves do not significantly contribute to the signal energy. A rather high threshold, on the other hand, tends to decrease the number of sensing actions. It is more likely to lead to wrong decisions by which subtrees are omitted whose leaves would provide relevant non-zero coefficients and hence the approximation error would increase. In the ultimate worst case, e.g. the signal is not sparse or the threshold is all too small, each node of the AHS tree is visited and consequently $2N-1$ measurements would be performed.

For K-sparse signals, the canonical threshold $\tau_0 = 0$ can be applied. Under mild conditions, e.g. if the non-zero coefficients stem from a probability density function, the signal will (almost surely) be perfectly sensed by AHS, and we can easily determine upper and lower bounds on the number of measurements. For $K = 1$, AHS needs $2\log_2 N + 1$ measurements in order to track down the only non-zero leaf. For $K > 1$ we have two limiting cases which yield the lower bound and the upper bound respectively.

The lowest number of AHS measurements arises, if all K non-zero leaves are maximally clustered within one subtree. Such a subtree has at least K leaves, and within this subtree, at least $2K-1$ nodes need to be visited. In addition we have to visit the nodes on the way from the root of the AHS tree to the root of the subtree. Note that the latter need to be counted twice, because one additional measurement per node is required in order to decide to omit all other subtrees. Hence, the *lower bound* on the required number of measurements for perfectly sensing a K-sparse signal is $2\log_2(N/K) + 2K - 1$.

The highest number of measurements is necessary, if the K non-zero leaves are uniformly distributed. This leads to K disjoint subtrees of equal size, each carrying one non-zero leaf. The number of leaves of each of these K subtrees is at most N/K. Consequently, the number of measurements within each subtree is at most $2\log_2(N/K) + 1$. Starting from the root of the AHS tree, $K-1$ measurements are required to reach the roots of these subtrees. Hence, the *upper bound* on the number of measurements is $2K\log_2(N/K) + 2K - 1$.

3. RESULTS

3.1 Results on full-size images

We applied AHS to three uncompressed gray scale images of size 512×512 pixels with a gray level depth of 8 bit (*Lena*, *Cameraman*, *Pirate*). We chose non-standard 2D Haar wavelets as sparse coding basis $\mathbf{\Psi}$. The images were normalized to unit energy before sensing and were inversly scaled after reconstruction. In order to evaluate the approximation performance as a function of the number of performed measurements, we applied for each image 200 different canonical thresholds ranging from $\tau_0 = 5 \cdot 10^{-6}$ to $\tau_0 = 10^{-4}$.

For comparison, the same images are also sensed with conventional compressed sensing (CS). In order to evaluate approximation performances for various numbers of measurements, multiple measurement matrices $\mathbf{\Phi}$ were generated by row-wise selecting basis vectors from the (real valued) noiselet basis uniformly at random. The measurement ratios $m = \frac{M}{N}$ were increased from 0.1 to 0.9 in steps of 0.1. As for AHS, non-standard 2D Haar wavelets were used as sparse coding basis $\mathbf{\Psi}$ for the reconstructions. Both ensembles are mutually incoherent[11] and this significantly reduces the required number of measurements. The CS reconstructions were obtained by solving the following ℓ_1-optimization problem:

$$\vec{x}^* = \underset{\vec{x}\in\mathbb{R}^N}{\arg\min} \|\mathbf{\Psi}^T\vec{x}\|_1 \text{ , s.t. } \|\mathbf{\Phi}\vec{x} - \vec{y}\|_2 \leq \varepsilon \ . \tag{1}$$

In order to solve (1), we used the log-barrier algorithm `l1qc_logbarrier.m` from the L1-MAGIC package.[12] For the reconstructions, we set $\varepsilon = 10^{-4}\|\vec{y}\|_2$.

In Figure 2, the approximation performances of the reconstructed images are plotted in terms of peak signal-to-noise ratio (PSNR) against measurement ratio m. With the image *Lena*, the approximation performance with AHS is around 2 dB better than with conventional CS if $m \leq 0.8$. For higher measurement ratios, i.e. $m \geq 0.9$, CS reconstructs the image *Lena* slightly better than AHS. With the image *Cameraman*, we also obtain superior sensing performance with AHS for measurement ratios $m < 0.65$. For $0.65 \leq m \leq 0.8$ both sensing aprroaches perform approximately equally well. For higher values of m, CS reconstructs the *Cameraman* image better than AHS. With the image *Pirate*, the approximation performance with AHS is around 1 dB better than with CS, for measurement ratios $m \leq 0.4$. For the investigated thresholds, no AHS reconstructions were available for

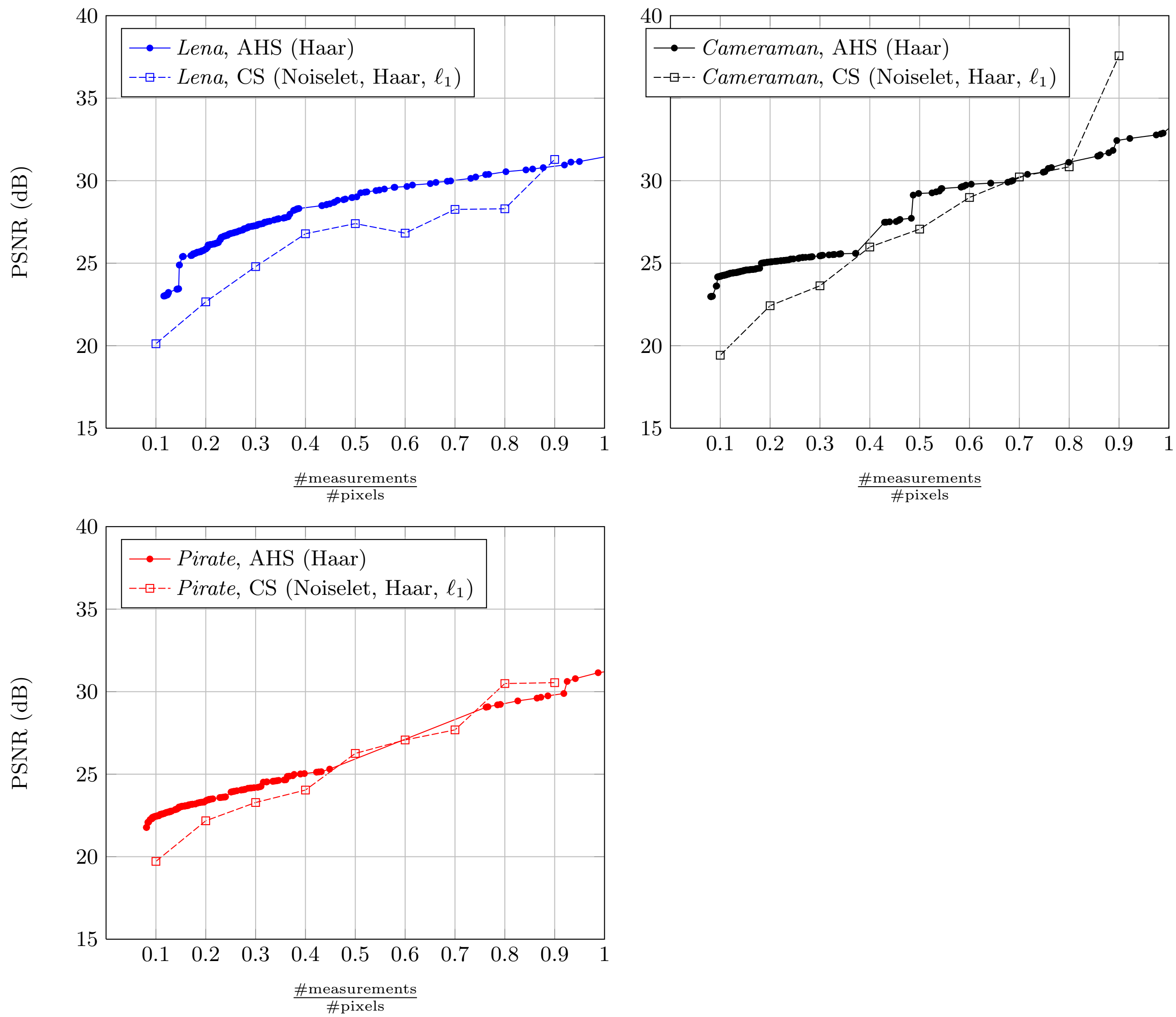


Figure 2: Sensing performance from a simulation with uncompressed images (*Lena*, *Cameraman*, *Pirate*) of size 512×512 pixels. AHS is compared to a conventional compressed sensing approach (CS). The peak signal-to-noise ratio (PSNR) is plotted against the measurement ratio.

(a) Original image (b) AHS reconstruction (29.79 dB) (c) CS reconstruction (28.98 dB)

Figure 3: Reconstructions of the test image *Cameraman* with a measurement ratio $\frac{\#\text{measurements}}{\#\text{pixels}} = 0.6$.

$0.45 \leq m \leq 0.75$. For measurement ratios larger than 0.8, CS has better approximation performance than AHS. Note that performance at high values of m is not really relevant.

Figure 3 depicts the original *Cameraman* image and exemplarily its reconstructions with AHS and CS at the measurement ratio $m = 0.6$. Note that the overall image quality is slightly better with AHS than with CS. On the other hand, edges and contours as well as facial details seem to be more accurate for the CS reconstruction at the cost of coarser block artifacts in the background. Block artifacts, however, appear in reconstructions of both methods, due to discontinuity properties of the Haar basis. These artifacts increase as m decreases. In order to reduce artifacts and to improve AHS performance, other orthogonal sparse coding bases might be more suitable.

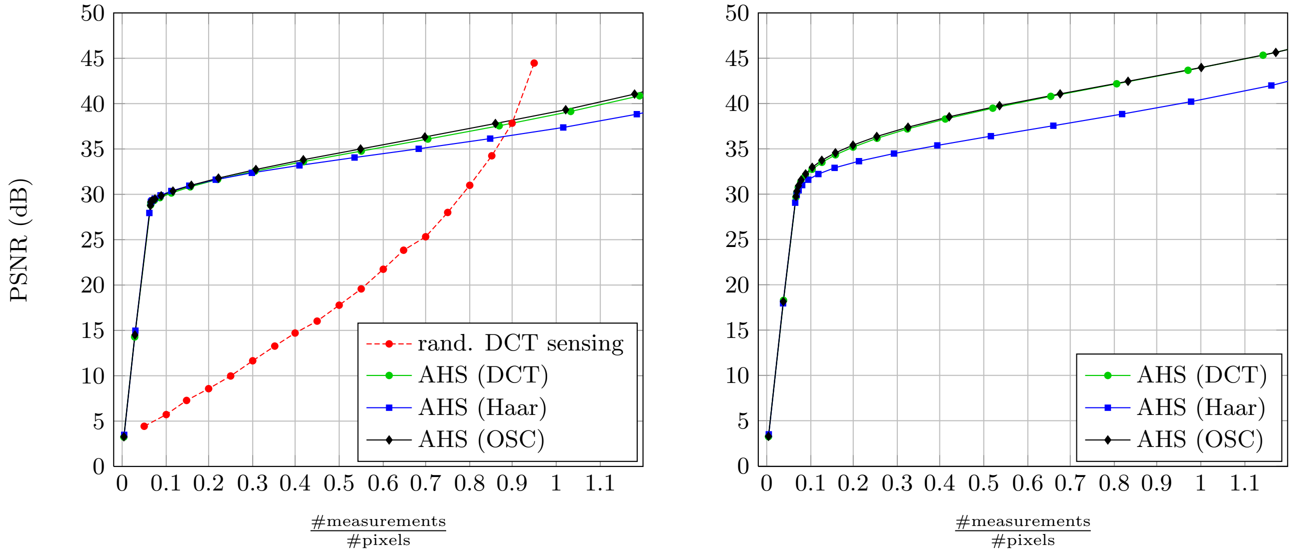
3.2 Results on natural image patches

We investigated the impact of (*i*) the choice of $\mathbf{\Psi}$, and (*ii*) the arrangement of leaf nodes within the AHS tree regarding uncompressed natural image patches of size 16×16 pixels. We extracted the image patches randomly from set one of the Nature Scene Collection,[13] i.e. from images of nature scenes containing no man made objects or people. In our analysis, we compared the 2D Discrete Cosine Transform (DCT) basis, non-standard 2D Haar wavelets, and a basis learned by Orthogonal Sparse Coding (OSC)[7,8] from training data. In the first scenario, we randomly assigned the basis vectors to the leaves of the AHS tree. In the second scenario, an ordered assignment was done as follows. We merged all basis vectors into disjoint sets of size two such that the correlation of the squared coefficients within the sets is maximal with respect to a training data set. This procedure is iteratively continued by merging these sets into disjoint sets of size four, eight, etc.. This scheme provides a natural composition rule for the AHS tree, such that commonly active coefficients share a common subtree more likely.

In Figure 4, the average AHS performance of a test set ($6 \cdot 10^3$ samples) is plotted against the measurement ratio m. Each marker corresponds to the mean of the scatter plot derived with the test data set for a particular canonical threshold. The findings indicate, that the Haar wavelet basis is not the optimal choice to sense natural images by AHS. The 2D DCT basis and the OSC basis yield equally good sensing performances, both significantly superior to the 2D Haar basis. Furthermore, the comparison of Figure 4a and Figure 4b demonstrates that sensing performance increases if the leaf nodes of the AHS tree are properly arranged. Figure 5 illustrates the measurement vectors, i.e. the filters, of a structured AHS subtree based on the OSC basis.

4. DISCUSSION

We introduced Adaptive Hierarchical Sensing (AHS), a novel sensing algorithm that measures the relevant transform coefficients of a sparse signal by performing fewer measurements than the number of signal dimensions depending on the degree of sparseness. AHS performs $\mathcal{O}(K \log N/K)$ sensing actions for K-sparse signals. We simulated the sensing of natural visual scenes with AHS and a conventional compressed sensing approach (CS). With AHS, we obtained sensing performances that can compete with those of CS. With AHS, higher PSNR values



(a) *Unstructured* AHS tree: leaves randomly arranged. The dashed curve gives a naive baseline obtained by random DCT measurements and direct reconstruction.

(b) *Structured* AHS tree: leaves arranged, s.t. commonly active coefficients share common subtrees more likely.

Figure 4: Average AHS performance for different bases from a simulation with uncompressed natural image patches of size 16×16 pixels. The peak signal-to-noise ratio (PSNR) is plotted against the measurement ratio.

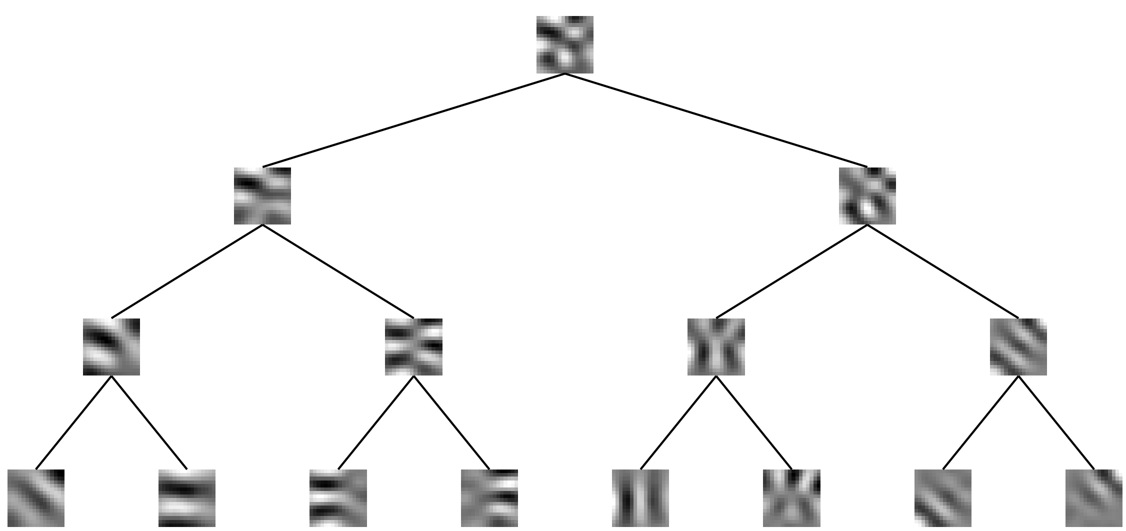


Figure 5: Measurement vectors (filters) of an AHS subtree. The leaves correspond to a subset of a sparse coding basis $\boldsymbol{\Psi}$ that was learned with OSC on natural image patches. For display purposes, the measurement vectors are scaled.

are obtained, especially for the more relevant cases of fewer measurements. Furthermore, we demonstrated on natural image patches that AHS performance can be improved if an appropriate sparse coding basis $\mathbf{\Psi}$ is chosen and the AHS tree is properly arranged. AHS performs deterministic measurements in an adaptive fashion. New sensing actions depend on previous measurements. Signals, measured with AHS, can be easily reconstructed and do not require to solve inverse optimization problems.

Our hierarchical and adaptive sensing procedure is related to known phenomena in human vision. With AHS, measurements at the higher layers of the AHS tree are used to gather rough information about the scene. This corresponds to a 'gist' of a scene, since these coarse measurements give rise to actions which lead to more refine sampling, i.e., further samples are acquired in some of the branches and leaves, and an increasingly more detailed representation of the scene is thus obtained. A further analogy to human vision is due to eye movements: the periphery provides a coarse, but not simply blurred,[14] view of the scene, based on which further more detailed measurements are made in the fovea.

Several improvements and extensions of AHS are possible. Currently, the decision making is based on relatively simple comparisons of the measurement values with a threshold. We intend to improve the decision strategy by employing more sophisticated rules, for instance Bayesian reasoning. In general, for making a decision at a particular node, it might be helpful to additionally involve the values measured at the ancestor nodes.

ACKNOWLEDGMENTS

The research is funded by the DFG Priority Programme SPP 1527, grant number MA 2401/2-1.

REFERENCES

[1] Olshausen, B. and Field, D., "Sparse coding of natural images produces localized, oriented, bandpass receptive fields," Technical Report CCN-110-95, Department of Psychology, Cornell University, Ithaca, New York 14853 (1995).

[2] Olshausen, B. A. and Field, D. J., "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature* (381), 607–609 (1996).

[3] Candès, E. J., Romberg, J., and Tao, T., "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on* **52**(2), 489–509 (2006).

[4] Romberg, J., "Imaging via compressive sampling," *Signal Processing Magazine, IEEE* **25**(2), 14–20 (2008).

[5] Wakin, M. B., Laska, J. N., Duarte, M. F., Baron, D., Sarvotham, S., Takhar, D., Kelly, K. F., and Baraniuk, R. G., "An architecture for compressive imaging.," in [*ICIP*], 1273–1276, IEEE (2006).

[6] Candes, E. J. and Tao, T., "Decoding by linear programming," *Information Theory, IEEE Transactions on* **51**, 4203–4215 (Dec. 2005).

[7] Schütze, H., Barth, E., and Martinetz, T., "Learning orthogonal bases for k-sparse representations," in [*Workshop New Challenges in Neural Computation 2013*], 119–120 (2013).

[8] Schütze, H., Barth, E., and Martinetz, T., "Learning efficient data representations with orthogonal sparse coding," submitted (2014).

[9] Deutsch, S., Averbuch, A., and Dekel, S., "Adaptive compressed image sensing based on wavelet modeling and direct sampling," in [*International Conference on Sampling Theory and Applications (SAMPTA'09)*], Tel Aviv University (2009).

[10] Burciu, I., Ion-Mărgineanu, A., Martinetz, T., and Barth, E., "Visual manifold sensing," in [*Human Vision and Electronic Imaging XIX*], Rogowitz, B. E. and de Ridder, T. N. P. H., eds., *Proc. of SPIE Electronic Imaging* **this volume** (2014).

[11] Tuma, T. and Hurley, P., "On the incoherence of noiselet and haar bases," (2009).

[12] Candes, E., "L1-magic: Recovery of sparse signals," *http://www.acm.caltech.edu/l1magic/* .

[13] Geisler, W. S. and Perry, J. S., "Statistics for optimal point prediction in natural images," *Journal of Vision* **11** (Oct. 2011).

[14] Hocke, J., Dorr, M., and Barth, E., "A compressed sensing model of peripheral vision," in [*Human Vision and Electronic Imaging XVII*], Rogowitz, B. E., Pappas, T. N., and de Ridder, H., eds., **8291**, Proceedings of SPIE (2012).