

IMAGE INPAINTING WITH A WAVELET DOMAIN HIDDEN MARKOV TREE MODEL

George Papandreou¹, Petros Maragos¹, and Anil Kokaram²

¹Sch. of Electrical and Computer Engineering, National Technical Univ. of Athens, Greece

²Dept. of Electronic and Electrical Engineering, Trinity College Dublin, Ireland

gpapan@cs.ntua.gr maragos@cs.ntua.gr anil.kokaram@tcd.ie

ABSTRACT

We present a novel technique for image inpainting, the problem of filling-in missing image parts. Image inpainting is ill-posed and we adopt a probabilistic model-based approach to regularize it. The main elements of our image model are, first, an over-complete complex-wavelet image representation, which ensures good shift invariance and directional selectivity and, second, a discrete-state/continuous-observation Hidden Markov Tree model for the wavelet coefficients, which captures key statistical properties of natural image wavelet responses, such as heavy-tailed histograms and persistence of large wavelet coefficients across scales. We show how these ideas can be integrated into a multi-scale generative process for natural images and present alternative deterministic and Markov chain Monte Carlo algorithms for image inpainting under this model. We demonstrate the effectiveness of the method in digitally restoring images of ancient wall-paintings.

Index Terms— Wavelet transforms, Image restoration, Monte Carlo methods.

1. INTRODUCTION

Image inpainting, a term coined in the influential paper [1], refers to the problem of filling-in missing image parts. Image inpainting, and the related technique of texture synthesis, have many practical applications, such as degraded film restoration, movie post-production, photograph restoration and retouching, and damaged paintings restoration. Apart from its practical applications, image inpainting draws research interest as a model problem on which the effectiveness of diverse image modeling approaches can be tested.

There is an abundance of image inpainting and texture synthesis techniques in the literature; we can broadly divide existing methods into exemplar-based and model-based. The first class of techniques utilizes non-parametric statistical image models [2–6]. While exemplar-based techniques often achieve the most visually pleasing results by exploiting information already present in the image, they are data intensive and do not supply a compact image representation for further image manipulation/description. The second class of techniques uses either partial differential equation (PDE) processes [1, 7, 8] designed to propagate image content into holes in a way that favors good edge continuation, or parametric probabilistic models [9, 10] which capture and reproduce key statistical properties of images of interest; the proposed method belongs to the latter class of parametric probabilistic techniques.

This work was supported by grant ΠΕΝΕΔ-2003-ΕΔ865 [co-financed by E.U.-European Social Fund (75%) and the Greek Ministry of Development-GSRT (25%)], by the European Community projects MUSCLE and ASPI, and by an Adobe grant to the Sigmedia Laboratory at Trinity College Dublin. We thank Prof. N. Kingsbury for sharing his DT-CWT code and F. Georma, Dr. A. Vlachopoulos, and Prof. C. Doumas of Akrotiri Excavation, Thera, for their invaluable help with the Akrotiri wall-paintings.

Among the parametric probabilistic models, those based on the principle of sparse image representation in linear bases are recently gaining significant attention, supported by studies on the statistics of wavelet coefficients of natural images [11, 12]. Image wavelet coefficient sparsity is also closely related to soft thresholding for image denoising and utilized in data analysis techniques such as Matching Pursuit or Basis Pursuit [13].

A number of recent inpainting methods build on the sparse image representation principle, demonstrating promising results [14–17]. However, a potential drawback of these techniques is that they utilize sparsity promoting priors which treat each wavelet coefficient independently. Various authors have argued that there are substantial benefits in introducing additional structure by modeling the inter-scale dependencies between wavelet coefficients; see *e.g.* [18] and the references therein. Previous studies have mostly demonstrated the effectiveness of structured models, such as the Hidden Markov Tree Model (HMT), in the denoising problem [19, 20]. We argue that, although imposing structured inter-scale Markov priors is more challenging for the inpainting problem due to the missing image areas, the benefits of such a modeling approach can be substantial. Intuitively, smaller part of the support of coarser scale filters falls in image holes and thus coarser scale wavelet coefficients are less affected by missing image parts and can lead and improve estimation of fine scale coefficients through the interscale dependency mechanism. Our paper, after reviewing the wavelet HMT model in Sec. 2, shows in Sec. 3 how it can be efficiently applied for inpainting, and then demonstrates in Sec. 4 its performance in filling-in ancient wall-paintings with missing parts and other artificially occluded images.

2. IMAGE MODEL: HIDDEN MARKOV TREE ON COMPLEX WAVELET COEFFICIENTS

We build our inpainting model upon a linear multiscale image representation. Critically-sampled discrete wavelet transforms (DWT) are neither shift-invariant nor directionally selective. Redundant, over-complete representations overcome these limitations and are thus better suited for our task [21]. In our work we use the *Dual-Tree Complex Wavelet Transform* (DT-CWT) [22, 23], which yields a four-times overcomplete, near shift-invariant, Gabor-like frame image representation consisting of six directionally selective complex-valued subbands per scale at four-times the computational cost of the critically-sampled DWT. An alternative representation with similar properties could be the complex steerable pyramid [10].

The wavelet transformation, due to its good decorrelation effect, is well-suited for statistical image modeling. However, independent gaussian modeling of wavelet coefficients is inadequate due to the following key wavelet coefficient properties of natural images [12]:

1. Heavy-tailed marginal histograms: Wavelet marginal his-

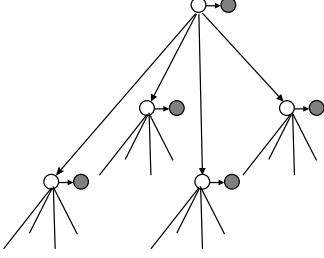


Fig. 1. Quadtree dependence structure of the DT-CWT/HMT model: each white node denotes a discrete state q_t and the adjacent gray node corresponds to the complex wavelet coefficient y_t .

tograms have a strong zero peak due to the weak wavelet response at smooth image parts spanning most of the image, and heavy tails due to the strong wavelet response at edges.

2. Persistence across scales: Edges cause strong responses to wavelet coefficients in their vicinity across multiple scales; this manifests itself in bow-tie-like joint histograms of scale-adjacent wavelet coefficients.

The heavy-tailed nature of marginal wavelet histograms can be modeled by sparsity-promoting priors. An effective way to model the inter-scale wavelet coefficient dependency is by means of scale-recursive Markov models [18].

In this work we use the Hidden Markov Tree (HMT) model, first introduced in the context of signal denoising in [19]. The HMT model captures both the heavy-tailed property by modeling the continuous wavelet observations as mixtures of Gaussians, and the inter-scale persistence by imposing a Markov scale dependence structure on the discrete hidden variables, as shown in Fig. 1. In our particular case of the DT-CWT, we follow the approach of [24], modeling each complex coefficient as a mixture of spherical 2-D Gaussians and utilizing an independent tree for each of the six directional subbands. Let $\theta = \{\pi, \{a_{ji}\}, \mu, \Sigma\}$ be the set of HMT parameters, corresponding to prior root node state probabilities and, for each scale, inter-scale state transition probabilities and mixture means/covariances, respectively. Also let t index the $T + 1$ nodes of the tree, where $t = 0$ is the root node, $t = T$ is the last finest scale node, and $p_t = \lfloor (t-1)/4 \rfloor$ is t 's parent. Then the a priori probability of a hidden state path $\mathbf{q} = (q_0, \dots, q_T)^T$ under the model is

$$P(\mathbf{q}) = \pi_{q_0} \prod_{t=1}^T a_{q_{p_t} q_t}, \quad (1)$$

while the probability of the wavelet coefficients vector $\mathbf{y} = (y_0, \dots, y_T)^T$, with $y_t = (y_{t,r}, y_{t,i})^T$ gathering the real and imaginary parts of each complex wavelet coefficient, given the path \mathbf{q} is

$$P(\mathbf{y}|\mathbf{q}) = N(\mathbf{y}; \mu_{\mathbf{q}}, \Sigma_{\mathbf{q}}) = \prod_{t=0}^T N(y_t; \mu_{q_t}, \Sigma_{q_t}), \quad (2)$$

where $N(\mathbf{x}; \mu, \Sigma)$ denotes a multi-variate Gaussian p.d.f. on \mathbf{x} with mean μ and covariance Σ .

Similarly to Hidden Markov Models (HMMs), the HMT is amenable to efficient calculations using scale-recursive $\mathcal{O}(T)$ algorithms; in particular, given the values of wavelet coefficients \mathbf{y} , we can perform estimation, *i.e.* calculation of $\gamma_t(i) = P(q_t = i|\mathbf{y})$, and decoding, *i.e.* calculation of the most probable path $\hat{\mathbf{q}} = \arg\max_{\mathbf{q}} P(\mathbf{q}|\mathbf{y})$, using the upward-downward and Viterbi algorithms, respectively; both just need an upward and then a downward pass of the quadtree [19]. Model training, specifically maximum-likelihood estimation of θ given training image data, is done by EM, utilizing the upward-downward estimation algorithm as sub-routine.

In the case of the classic denoising problem [19], one observes white Gaussian noise corrupted image intensity values throughout the image plane. Due to the orthogonal nature of the DWT we end up with white Gaussian noise corrupted wavelet coefficients \mathbf{y}_n , whose probability given the path \mathbf{q} can be shown to be

$$P(\mathbf{y}_n|\mathbf{q}) = N(\mathbf{y}_n; \mu_{\mathbf{q}}, \Sigma_{\mathbf{q}} + \Sigma_n) = \prod_{t=0}^T N(y_{n,t}; \mu_{q_t}, \sigma_{q_t}^2 + \sigma_n^2), \quad (3)$$

where $\Sigma_n = \sigma_n^2 I$ is the noise covariance. The last equation is very similar to the non-noisy case of Eq. (2), and thus hidden state estimation can be performed efficiently using the upward-downward algorithm; subsequent denoising just amounts to independent manipulation of each wavelet coefficient [19]. We see that noise independence in the wavelet domain is crucial in making exact inference and denoising with the HMT model tractable.

3. IMAGE INPAINTING WITH THE HMT MODEL

For inpainting the situation is more intricate than for the denoising problem. In particular, the missing value structure of the inpainting problem breaks down the orthogonal property of the wavelet transform and introduces dependencies between wavelet coefficients. More specifically, let us write $\mathbf{x} = (\mathbf{x}_o^T, \mathbf{x}_h^T)^T$ by splitting the M -pixel image intensity vector \mathbf{x} into its M_o -pixel observed part \mathbf{x}_o and its M_m -pixel ($M_m = M - M_o$) missing part \mathbf{x}_m ; then assuming white Gaussian spatial domain noise yields the observation model

$$P(\mathbf{x}_o|\mathbf{y}) = N(\mathbf{x}_o; A\mathbf{y}, \Sigma_n), \text{ with } A = WF_s, \Sigma_n = \sigma_n^2 I, \quad (4)$$

where $W = [I_{M_o \times M_o} \ O_{M_o \times M_m}]$ is the $M_o \times M$ inpainting projection matrix and F_s denotes the inverse (synthesis) DT-CWT. Simply put, we only measure the wavelet coefficients \mathbf{y} indirectly through a map A , consisting of three steps: (1) inverse wavelet transform F_s , (2) occlusion at the inpainting mask W , and (3) addition of noise with covariance Σ_n . The non-orthogonal and non-invertible nature of A , even if we used an orthogonal wavelet representation, precludes treating each wavelet coefficient separately as in Eq. (3).

We are interested in estimating the distribution of the wavelet coefficients \mathbf{y} , given the observations:

$$P(\mathbf{y}|\mathbf{x}_o) \propto P(\mathbf{x}_o|\mathbf{y})P(\mathbf{y}) = P(\mathbf{x}_o|\mathbf{y}) \sum_{\mathbf{q} \in Q} P(\mathbf{y}|\mathbf{q})P(\mathbf{q}), \quad (5)$$

where Q is the set of all possible paths \mathbf{q} . Unfortunately, since the wavelet coefficients \mathbf{y} are coupled, we cannot estimate \mathbf{y} with the efficient upward-downward algorithm applicable in the denoising scenario, while explicitly summing over all paths $\mathbf{q} \in Q$ in Eq. (5) has exponential complexity. We discuss next two approaches, one deterministic and one stochastic, to efficiently estimate \mathbf{y} . Both rely on the fact that estimating \mathbf{y} is straightforward when \mathbf{q} is known and vice versa. In particular,

$$P(\mathbf{y}|\mathbf{q}, \mathbf{x}_o) \propto P(\mathbf{x}_o|\mathbf{y})P(\mathbf{y}|\mathbf{q}) = N(\mathbf{y}; \mu_{\mathbf{y}|\mathbf{q}}, \Sigma_{\mathbf{y}|\mathbf{q}}), \quad (6)$$

with $\mu_{\mathbf{y}|\mathbf{q}} = \Sigma_{\mathbf{y}|\mathbf{q}}(\Sigma_{\mathbf{q}}^{-1} \mu_{\mathbf{q}} + A^T \Sigma_n^{-1} \mathbf{x}_o)$ and $\Sigma_{\mathbf{y}|\mathbf{q}}^{-1} = \Sigma_{\mathbf{q}}^{-1} + A^T \Sigma_n^{-1} A$. To efficiently obtain $\mu_{\mathbf{y}|\mathbf{q}}$ we solve the system $\Sigma_{\mathbf{y}|\mathbf{q}}^{-1} \mu_{\mathbf{y}|\mathbf{q}} = \Sigma_{\mathbf{q}}^{-1} \mu_{\mathbf{q}} + A^T \Sigma_n^{-1} \mathbf{x}_o$ iteratively with the conjugate gradients (CG) method, using $\Sigma_{\mathbf{q}}^{-1}$ as diagonal preconditioner. The CG method avoids explicitly forming the system matrix $\Sigma_{\mathbf{y}|\mathbf{q}}^{-1}$, only requiring repeated computation of the matrix-vector product $\Sigma_{\mathbf{y}|\mathbf{q}}^{-1} \mathbf{y}$; this involves computing $F_s \mathbf{y}$ and $F_s^T \mathbf{x}$, which can be done efficiently with the backward (using the primal filter set) and forward (using the dual filter set), respectively, discrete filterbank implementation of the bi-orthogonal DT-CWT [13, 22].

3.1. Alternating Optimization Approach

In the spirit of EM-like approaches [16, 17], we can cycle in an ICM-inspired [25] fashion between hidden path and wavelet coefficient optimization until convergence to a self-consistent solution. Specifically, we can alternate until convergence the following two steps: (1) Given a guess \mathbf{y}_{i-1} for the wavelet coefficients, compute the most probable path $\mathbf{q}_i = \operatorname{argmax} P(\mathbf{q}|\mathbf{y}_{i-1})$ using the Viterbi algorithm outlined in Sec. 2. (2) Given \mathbf{q}_i , compute \mathbf{y}_i as the mode $\boldsymbol{\mu}_{\mathbf{y}|\mathbf{q}_i}$ of the gaussian $P(\mathbf{y}|\mathbf{q}_i, \mathbf{x}_o)$ given in Eq. (6). Although this method cannot yield a complete description of the wavelet coefficient posterior (5), it can lead to a satisfactory inpainting result in practice.

3.2. Handling Missing Data by Efficient Block Gibbs Sampling

A more principled approach to inpainting using the HMT model is based on random sampling from the joint distribution of the wavelet coefficients and the hidden state path

$$P(\mathbf{y}, \mathbf{q}|\mathbf{x}_o) \propto P(\mathbf{x}_o|\mathbf{y}, \mathbf{q})P(\mathbf{y}, \mathbf{q}) = P(\mathbf{x}_o|\mathbf{y})P(\mathbf{y}|\mathbf{q})P(\mathbf{q}). \quad (7)$$

Drawing samples directly from Eq. (7) is infeasible. However, as we detail next, we can design a block Gibbs sampler which alternates between sampling from $\mathbf{y}|\mathbf{q}, \mathbf{x}_o$ and $\mathbf{q}|\mathbf{y}, \mathbf{x}_o$, yielding a practical MCMC method which efficiently explores the joint distribution (7).

3.2.1. Wavelet sampling conditional on hidden state: $\mathbf{y}|\mathbf{q}, \mathbf{x}_o$

When the hidden state \mathbf{q} is given, it can be readily verified that we can draw a sample from the gaussian $P(\mathbf{y}|\mathbf{q}, \mathbf{x}_o)$ given in Eq. (6) by means of the following procedure: (1) Draw sample $\mathbf{y}_{u,s} \sim N(\boldsymbol{\mu}_{\mathbf{q}}, \Sigma_{\mathbf{q}})$; (2) Draw sample $\mathbf{x}_{o,s} \sim N(\mathbf{x}_o, \Sigma_n)$; (3) Solve $\mathbf{y}_s = \Sigma_{\mathbf{y}|\mathbf{q}}(\Sigma_{\mathbf{q}}^{-1}\mathbf{y}_{u,s} + A^T\Sigma_n^{-1}\mathbf{x}_{o,s})$.

3.2.2. Hidden state path sampling conditional on wavelet coefficients: $\mathbf{q}|\mathbf{y}, \mathbf{x}_o$

Given an estimate of the wavelet coefficients, the distribution of hidden paths is independent of \mathbf{x}_o :

$$P(\mathbf{q}|\mathbf{y}, \mathbf{x}_o) = P(\mathbf{q}|\mathbf{y}). \quad (8)$$

Drawing a random sample \mathbf{q} from Eq. (8) is distinct from Viterbi decoding, which yields the most probable path. The conventional approach to sampling from Eq. (8) (see *e.g.* [20]) is to update each hidden markov node independently, *i.e.* sequentially draw samples from $P(q_t|\mathbf{y}, \mathbf{q}_{-t})$, where \mathbf{q}_{-t} denotes all but the t -th discrete hidden state. However this single-site update strategy can lead to slow mixing; a much better alternative is to draw an independent sample of the whole hidden path \mathbf{q} at once. This can be achieved by an efficient recursive algorithm whose HMM version was first developed in [26]. For the case of the HMT model the algorithm amounts to the usual upward estimation pass on the tree, which computes the quantity $\beta_t(i) = P(q_t = i|\mathbf{y}_{\mathcal{T}_t})$ ($\mathbf{y}_{\mathcal{T}_t}$ denotes the \mathbf{y} values in the subtree with root t) [19], followed by a stochastic Viterbi-like downward back-tracking pass in which: (1) we assign the root node's state with probability $P(q_0 = i|\mathbf{y}) \propto \pi_i\beta_0(i)$ and (2) recursively, given the parent's assigned state, assign the t node's state with probability $P(q_t = i|q_{p_t} = j, \mathbf{y}) \propto a_{ji}\beta_t(i)$.

3.2.3. Summarizing Gibbs samples through Rao-Blackwellization

Alternating the steps outlined in Secs. 3.2.1 and 3.2.2, and possibly discarding the first few 'burn-in' samples, yields a sequence $\{\{\mathbf{q}_1, \mathbf{y}_1\}, \{\mathbf{q}_2, \mathbf{y}_2\}, \dots, \{\mathbf{q}_S, \mathbf{y}_S\}\}$ of Gibbs samples from the

joint distribution (7), which we typically wish to summarize into a single wavelet coefficient estimate $\hat{\mathbf{y}}$ that will produce the final image inpainting result $\hat{\mathbf{x}} = F_s\hat{\mathbf{y}}$. There are many plausible criteria for selecting $\hat{\mathbf{y}}$, the MMSE being the most commonly used.

The conventional estimator of the MMSE wavelet coefficients is the sample mean $\hat{\mathbf{y}}_S = \frac{1}{S}\sum_{s=1}^S \mathbf{y}_s$. However we opt here for the *Rao-Blackwellized* (RB) estimator $\hat{\mathbf{y}}_{RB} = \frac{1}{S}\sum_{s=1}^S E[\mathbf{y}|\mathbf{q}_s, \mathbf{x}_o]$ which significantly reduces the variance of the estimate [27]; intuitively, this is achieved because for each hidden path sample \mathbf{q}_s the RB estimator utilizes the exact MMSE estimate $\boldsymbol{\mu}_{\mathbf{y}|\mathbf{q}_s} = E[\mathbf{y}|\mathbf{q}_s, \mathbf{x}_o]$ instead of relying on the single draw $\mathbf{y}_s \sim \mathbf{y}|\mathbf{q}_s, \mathbf{x}_o$. Thanks to the reduced variance of the RB estimator and the rapid chain mixing facilitated by the efficient block Gibbs sampling steps described above, we can compute satisfactory image inpainting results after only very few Gibbs sampling steps S .

3.3. PDE-based Initialization

Both the alternating optimization and the Gibbs sampling algorithms for inpainting under the HMT model are iterative and thus benefit from a good initial guess. We get this guess by solving the Laplace PDE equation $\nabla^2 u = 0$ in image holes, with Dirichlet boundary conditions at the known image boundaries. This can be considered as a simple and efficient form of PDE-based inpainting which has been proven adequate in our experiments.

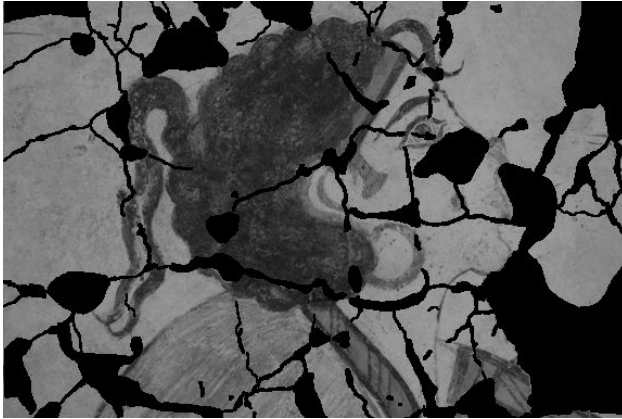
4. INPAINTING EXPERIMENTS

Our interest in image inpainting is motivated by the problem of digital restoration of the ancient wall-paintings discovered in the prehistoric settlement of Akrotiri, Thera, which are famous for their thematic gamut, their artistic value, and the abundant and diversified information they yield about the Aegean world more than 3500 years ago [28]. Restoring these unique 17th century B.C. wall-paintings is particularly challenging; damages include smaller or bigger missing parts, heavy or finer cracks. In Fig. 2 we show the result of applying our HMT inpainting technique (solved with Gibbs sampling) on a representative Theran wall-painting image on which we have overlaid the manually-specified inpainting mask. The results are quite satisfactory, especially in smooth areas; however further improvements would be desirable, *e.g.* sharper edge continuation at big gaps.

To further validate the proposed approach, we have performed image inpainting experiments on artificially occluded images with known ground truth. Figure 3 compares inpainting results on the Lena image with the Laplace PDE method and the HMT model, solved with either the alternating optimization or the Gibbs sampling technique. Our results were computed with 10 alternating optimization/Gibbs sampler steps, utilizing 3 HMT hidden states and 6 analysis levels, and running 50 conjugate gradient iterations for each sample. HMT model training was done with EM [19] on the Laplace PDE image estimate. We observe that the proposed HMT inpainting model produces qualitative better, sharper results than the Laplace method, while also having better PSNR values (discrepancy between inpainting result and ground-truth in the occluded part).

5. DISCUSSION AND FUTURE WORK

The main contribution of our work is to incorporate structured inter-scale dependencies of wavelet coefficients into a sound probabilistic treatment of the image inpainting problem in the framework of sparse image representations. There are many promising directions



(a) Thera wall painting with missing parts annotated



(b) Inpainting result with proposed HMT model

Fig. 2. Inpainting on ancient Thera wall-painting.

for future work; for example, finding ways to efficiently integrate intra-scale wavelet dependencies in the model can further improve the performance of the model, especially in sharply propagating linear image structures in large holes; another promising approach is to integrate into the model additional families of wavelet dictionaries, which has been shown to result in substantial performance improvements [15, 17, 20]. Regarding the practical applicability of inpainting approaches in general, it is important to pursue automatic missing part detection techniques, that will require little manual user effort.

6. REFERENCES

- [1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proceedings of SIGGRAPH*, 2000, pp. 417–424.
- [2] J. S. De Bonet, "Multiresolution sampling procedure for analysis and synthesis of texture images," in *Proc. of SIGGRAPH*, 1997, pp. 361–368.
- [3] A. A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling," in *Proc. ICCV*, 1999, vol. 2.
- [4] A. Criminisi, P. Pérez, and K. Toyama, "Object removal by exemplar-based inpainting," in *Proc. CVPR*, 2003, pp. 417–424.
- [5] J. Jia and C.-K. Tang, "Inference of segmented color and texture description by tensor voting," *IEEE Trans. on PAMI*, vol. 26, no. 6, pp. 771–786, 2004.
- [6] J. Sun, L. Yuan, J. Jia, and H.-Y. Shum, "Image completion with structure propagation," in *Proc. of SIGGRAPH*, 2005, pp. 861–868.
- [7] S. Masnou and J.-M. Morel, "Level lines based disocclusion," in *Proc. ICIP*, 1998, vol. 3, pp. 259–263.
- [8] T. F. Chan and J. Shen, "Mathematical models for local nontexture inpaintings," *SIAM J. Applied Math.*, vol. 62, no. 3, pp. 1019–1043, 2001.
- [9] S.C. Zhu and D. Mumford, "Prior learning and gibbs reaction-diffusion," *IEEE Trans. on PAMI*, vol. 19, no. 11, pp. 1236–1250, 1997.



(a) original occluded

(b) Laplace PDE (PSNR: 20.62)



(c) HMT / Alt. Opt. (PSNR: 21.95) (d) HMT / Gibbs (PSNR: 22.00)

Fig. 3. Inpainting results on artificially occluded Lena image.

- [10] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *IJCV*, vol. 40, no. 1, pp. 49–71, 2000.
- [11] D.J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *JOSA (A)*, vol. 4, no. 12, pp. 2379–2394, 1987.
- [12] E. P. Simoncelli, "Modeling the joint statistics of images in the wavelet domain," in *Proc. of SPIE*, 1999, vol. 3813, pp. 188–195.
- [13] S. Mallat, *A Wavelet Tour of Signal Processing*. Acad. Press, 2 edition, 1999.
- [14] M. Elad, J.-L. Starck, P. Querre, and D. L. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis," *App. and Comp. Harm. Anal.*, vol. 19, pp. 340–358, 2005.
- [15] O. G. Guleryuz, "Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising-part I: theory," *IEEE Tr. on Image Proc.*, vol. 15, no. 3, pp. 539–554, Mar. 2006.
- [16] T. Lee and X.-L. Meng, "A self-consistent wavelet method for denoising images with missing pixels," in *Proc. ICASSP*, 2005, vol. II, pp. 41–44.
- [17] M. J. Fadili and J.-L. Starck, "EM algorithm for sparse representation-based image inpainting," in *Proc. ICIP*, 2005.
- [18] A. Willsky, "Multiresolution markov models for signal and image processing," *Proc. of the IEEE*, vol. 90, no. 8, pp. 1396–1458, 2002.
- [19] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk, "Wavelet-based statistical signal processing using hidden markov models," *IEEE Trans. on Signal Processing*, vol. 46, no. 4, pp. 886–902, 1998.
- [20] C. Févotte, B. Torrèsani, L. Daudet, and S. J. Godsill, "Sparse linear regression with structured priors and application to denoising of musical audio," *IEEE Trans. on Aud., Speech, and Lang. Proc.*, vol. 16, no. 1, pp. 174–185, 2008.
- [21] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable multiscale transforms," *IEEE Trans. on Inform. Theory*, vol. 38, no. 2, pp. 587–607, 1992.
- [22] N. Kingsbury, "Complex wavelets for shift invariant analysis and filtering of signals," *App. and Comp. Harm. Anal.*, vol. 10, no. 3, pp. 234–253, 2001.
- [23] I. W. Selesnick, R. G. Baraniuk, and N. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Proc. Mag.*, vol. 22, no. 6, pp. 123–151, 2005.
- [24] H. Choi, J. K. Romberg, R. G. Baraniuk, and N. Kingsbury, "Hidden markov tree modeling of complex wavelet transforms," in *Proc. ICASSP*, 2000, pp. 133–136.
- [25] J. Besag, "On the statistical analysis of dirty pictures," *J. R. Stat. Soc., Ser. B.*, vol. 48, no. 3, pp. 259–302, 1986.
- [26] S. Chib, "Calculating posterior distributions and modal estimates in markov mixture models," *J. of Econometrics*, vol. 75, pp. 79–97, 1996.
- [27] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer, 2005.
- [28] C. Doumas, *The Wall-paintings of Thera*. The Thera Foundation, 1992.