

Unwrapping of MR Phase Images Using a Markov Random Field Model

Lei Ying*, *Member, IEEE*, Zhi-Pei Liang, *Fellow, IEEE*, David C. Munson, Jr., *Fellow, IEEE*, Ralf Koetter, *Member, IEEE*, and Brendan J. Frey, *Member, IEEE*

Abstract—Phase unwrapping is an important problem in many magnetic resonance imaging applications, such as field mapping and flow imaging. The challenge in two-dimensional phase unwrapping lies in distinguishing jumps due to phase wrapping from those due to noise and/or abrupt variations in the actual function. This paper addresses this problem using a Markov random field to model the true phase function, whose parameters are determined by maximizing the *a posteriori* probability. To reduce the computational complexity of the optimization procedure, an efficient algorithm is also proposed for parameter estimation using a series of dynamic programming connected by the iterated conditional modes. The proposed method has been tested with both simulated and experimental data, yielding better results than some of the state-of-the-art method (e.g., the popular least-squares method) in handling noisy phase images with rapid phase variations.

Index Terms—Bayesian estimation, field mapping, magnetic resonance imaging, Markov random field, phase unwrapping.

I. INTRODUCTION

A. Background

Magnetic resonance (MR) phase images often contain useful information, such as spatial inhomogeneities in the polarizing magnetic field and the velocity of blood flow [1], [2]. Extracting the phase image ϕ from its measured complex MR image $I = |I| \exp(\phi)$ is nontrivial because ϕ is uniquely defined only in the principal value range of $(-\pi, \pi]$. Any value outside this interval will be folded back into the principal value range to produce the so-called wrapped phase ψ , which differs from ϕ by an unknown integer multiple of 2π . Formally, the phase unwrapping problem is defined as: given the wrapped phase $\psi \in (-\pi, \pi]$, find the “true” phase ϕ , which is related to ψ by

$$\psi = \mathcal{W}(\phi) = \phi - 2\pi \left\langle \frac{\phi}{2\pi} \right\rangle \quad (1)$$

Manuscript received June 21, 2005; revised October 6, 2005. This work was supported in part by the National Science Foundation (NSF) under Grant CCR01-05719. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was J. Pauly. *Asterisk indicates corresponding author.*

*L. Ying is with the Department of Electrical Engineering and Computer Science, University of Wisconsin, Milwaukee, WI 53201 USA (e-mail: leiying@uwm.edu).

Z.-P. Liang and R. Koetter are with the Department of Electrical and Computer Engineering, University of Illinois, Urbana, IL 61801 USA.

D. C. Munson, Jr., is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA.

B. J. Frey is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada.

Digital Object Identifier 10.1109/TMI.2005.861021

where \mathcal{W} is called the wrapping operator, and $\langle \cdot \rangle$ rounds its argument to the closest integer. The phase unwrapping problem is mathematically ill-posed because \mathcal{W} is a surjective mapping. A common approach to addressing the problem is to use spatial and/or temporal constraints (e.g., the continuity/smoothness constraint).

In the one-dimensional (1-D) case, for example, phase unwrapping can be done relatively easily if phase continuity can be assumed. In 1982, Itoh showed that the wrapped phase gradient modulo 2π is the same as the corresponding true phase gradient if the latter is less than π everywhere [3], as summarized in the following lemma with respect to sampled phase values.

Lemma 1.1: Let

$$\Delta\phi_n = \phi_n - \phi_{n-1}, \quad \Delta\psi_n = \psi_n - \psi_{n-1} \quad (2)$$

where ϕ_n and ψ_n represent the true and wrapped phase value at location n , respectively. If the “smoothness condition”

$$|\Delta\phi_n| \leq \pi \quad (3)$$

is satisfied, then

$$\Delta\phi_n = \mathcal{W}(\Delta\psi_n). \quad (4)$$

The above Lemma suggests a simple phase unwrapping method, that is, integrating the wrapped phase gradient.

Lemma 1.1 can also be extended to higher dimensions. For example, for two-dimensional (2-D) phase functions, we have the following result.

Lemma 1.2: Let $\phi_{m,n}$ and $\psi_{m,n}$ represent the true and wrapped phase values at pixel (m, n) , and

$$\begin{aligned} \Delta_x\phi_{m,n} &= \phi_{m,n} - \phi_{m-1,n} \\ \Delta_y\phi_{m,n} &= \phi_{m,n} - \phi_{m,n-1} \\ \Delta_x\psi_{m,n} &= \psi_{m,n} - \psi_{m-1,n} \\ \Delta_y\psi_{m,n} &= \psi_{m,n} - \psi_{m,n-1}. \end{aligned} \quad (5)$$

If the following inequalities hold:

$$|\Delta_x\phi_{m,n}| \leq \pi \quad \text{and} \quad |\Delta_y\phi_{m,n}| \leq \pi \quad (6)$$

then

$$\begin{aligned} \Delta_x\phi_{m,n} &= \mathcal{W}(\Delta_x\psi_{m,n}) \\ \Delta_y\phi_{m,n} &= \mathcal{W}(\Delta_y\psi_{m,n}). \end{aligned} \quad (7)$$

Lemma 2 suggests that phase unwrapping in two dimensions may also be accomplished by phase integration, provided that the conditions in (6) are satisfied, which is rarely the case in practice due to measurement noise and/or rapid phase variations. As a result, integration over the wrapped phase gradient field can produce different results depending on the integration

path taken [4]. This path-dependent problem, unique to multi-dimensional phase unwrapping, presents many challenges [5]. Nonetheless, multidimensional data do provide additional information that can make phase unwrapping more robust than 1-D data. For example, in the 1-D case, a phase unwrapping error at a single point will propagate to the remainder of the unwrapped phase function along the integration path, while this error may be avoidable in multidimensional spaces because of the existence of multiple integration paths [6]–[8].

B. State-of-the-Art Algorithms

Phase unwrapping is a classic problem in signal processing and much work has been done. This section provides a brief summary of several key phase-unwrapping algorithms relevant to the proposed work. For a comprehensive review of phase unwrapping, the reader is referred to [9].

There are two key components in a phase unwrapping method: 1) a model for the true phase function, and 2) an algorithm for phase recovery based on the model. Phase modeling is undoubtedly the key to successful phase unwrapping; it allows for prior information to be incorporated, making the phase unwrapping problem solvable. There are deterministic and statistical models. Existing deterministic models are usually heuristic and can be viewed as a direct extension of the model in Lemma 1.2. When (7) is not satisfied, an approximate model is often constructed by minimizing the difference between the left- and right-hand sides of (7) while ensuring path-independence of the resulting phase gradients. A general objective function used to measure phase difference is in the form of the L_p -norm [9]

$$E = \left(\sum_{m,n} [\Delta_x \phi_{m,n} - \mathcal{W}(\Delta_x \psi_{m,n})]^p + \sum_{m,n} [\Delta_y \phi_{m,n} - \mathcal{W}(\Delta_y \psi_{m,n})]^p \right)^{\frac{1}{p}}. \quad (8)$$

The minimum-norm solution, $\arg \min E$, yields a smooth phase map but may have large phase error in the presence of noise and phase discontinuities [10]. In addition, the unwrapped phase usually fails the congruence test, which requires that rewrapping the unwrapped result reproduce the measured phase. Improvements over the minimum-norm solution have been made by incorporation of pixel-by-pixel weightings in (8) to reduce the effects of background noise [2]. Models have also been proposed to include phase discontinuities [6], [11]–[13]. Recently, statistical phase models have been proposed for synthetic aperture radar applications [14], [15]. A desirable feature about these models is their ability to take into account both phase discontinuities and noise.

Determining the unwrapped phase values under a given phase model entails solving an optimization problem, which is often nontrivial. A popular algorithm to find the minimum-norm (or weighted minimum-norm) solution is by solving a differential equation. For $p = 2$ (the least-squares case), the Poisson equation results, which can be solved efficiently using fast Fourier transform-based [16], discrete cosine transform-based [17], or multigrid methods [18]. When $p \neq 2$, the partial differential equation is usually solved iteratively, which is often much slower than the least-squares method. Another class of

optimization methods, known as path-following methods, is designed for models that include discontinuities. They integrate the phase derivatives along some “optimal” paths that allow for discontinuities [11] in the unwrapped phase function. In search for an optimal integration path, several algorithms have been proposed, which include the branch-cut algorithm [6], [19], network-flow algorithms [5], [12], [13], and minimum and maximum spanning tree algorithms [20]–[23]. Most of these algorithms can be used with a statistical phase model when the model is a simple function of the gradient field $\Delta\Phi$ [15], [24], although they cannot handle more complex phase models.

This paper proposes a new method to unwrap 2-D MR phase images. Two novel features of the proposed method are: 1) use of a Markov random field (MRF) to model the true phase image, and 2) an efficient algorithm for phase estimation. The proposed method is detailed in Section II. Experimental results to demonstrate the performance of the proposed method are shown in Section III, followed by some discussions in Section IV. The paper is concluded in Section V.

II. PROPOSED METHOD

A. Phase Modeling

The proposed method is based on a statistical model, in which both the true phase $\phi_{m,n}$ and the wrapped phase $\psi_{m,n}$ at each pixel are treated as random variables, and their corresponding images as random fields denoted as $\Phi = \{\phi_{m,n}, (m,n) \in \mathcal{S}\}$ and $\Psi = \{\psi_{m,n}, (m,n) \in \mathcal{S}\}$, where \mathcal{S} is the index set of the image pixels. In this setting, phase unwrapping becomes an estimation problem. Specifically, similar to other methods [15], [25], [26], the phase is estimated using a maximum *a posteriori* (MAP) estimator, which solves for the true phase image by

$$\begin{aligned} \hat{\Phi} &= \arg \max_{\Phi} p(\Phi|\Psi) \\ &\equiv \arg \max_{\Phi} p(\Psi|\Phi)p(\Phi) \end{aligned} \quad (9)$$

where $p(\Phi|\Psi)$ and $p(\Phi)$ denote the posterior and prior probabilities, $p(\Psi|\Phi)$ is the likelihood function, and $\arg \max_{\Phi}$ denotes the maximizing value of Φ within its domain. Equation (9) generalizes the MAP estimator used by Chen and Zebker [15], in which $p(\Delta\Phi|\Delta\Psi)$ is defined on a gradient field. To solve the above MAP problem, we need to specify the likelihood function and the prior probability, which are discussed below.

1) *Likelihood Function*: We define the likelihood function, based on the congruence constraint, as

$$p(\Psi|\Phi) = \delta[\Psi - \mathcal{W}(\Phi)] \quad (10)$$

where δ is the Dirac delta function. Given the above likelihood function, the maximum likelihood (ML) estimate of the true phase is not unique. Additional constraints on the desired true phase are incorporated in the prior probability.

2) *Prior Probability*: We propose to use a Markov random field (MRF) to model the desired “true” phase image. Specifically, based on the standard MRF theory, the elements of Φ are assumed to satisfy the following two conditions:

$$p(\phi_{m,n}) > 0 \quad (11)$$

$$p(\phi_{m,n} | \phi_{\mathcal{S} \setminus (m,n)}) = p(\phi_{m,n} | \phi_{\mathcal{N}_{m,n}}) \quad (12)$$

where $\mathcal{S} \setminus (m,n)$ is the entire set of pixels excluding pixel (m,n) , and $\mathcal{N}_{m,n}$ represents the set of (m,n) 's neighboring pixels. This neighborhood system \mathcal{N} has the following properties.

- A pixel is not a neighbor of itself: $(m, n) \notin \mathcal{N}_{m,n}$.
- The neighboring relationship is mutual: $(m, n) \in \mathcal{N}_{m',n'} \iff (m', n') \in \mathcal{N}_{m,n}$.

The MRF model is selected based on three considerations. First, because each individual pixel itself provides no information for phase unwrapping, there is a need for incorporating spatial constraints to specify how pixels or groups of pixels interact [9]. A random field is an effective way to model this interaction and the uncertainties within this interaction [27]. Second, the Markovian assumption is based on the fact that the contextual dependence in a natural image is primarily local [28], and the true phase image has a close relationship to a natural image. In fact, the “locality” assumption has been used implicitly in other phase models (e.g., the minimum-norm model), which require the difference between adjacent pixels to be small. Compared with these models, the MRF model provides a more general mechanism for modeling local contextual dependence, with the flexibility to handle effectively both smooth and nonsmooth features [29], [30]. Third, the MRF model is computationally appealing because the local nature of the associated energy functions results in computationally efficient algorithms, as demonstrated in this paper.

According to the Hammersley-Clifford theorem [28], an MRF has an equivalent Gibbs distribution given by

$$p(\Phi) \propto \exp\left(-\frac{\sum_{c \in C} [V_c(\phi_c)]}{T}\right) \quad (13)$$

where T is a temperature parameter chosen to be unity in the paper. The argument of the exponential function includes a sum of clique potentials $V_c(\phi_c)$ over all the possible cliques C , with ϕ_c denoting a vector composed of the set of phase values within a clique c . A clique c defines a subset of pixels in which every pair of distinct pixels is neighbors, except for single-pixel cliques. For example, in the first-order neighborhood system shown in Fig. 1, where the phase at the center pixel is only dependent on its four adjacent neighbors, a clique is a pair of two adjacent pixels that is either $(\phi_{m,n}, \phi_{m,n-1})$ or $(\phi_{m,n}, \phi_{m-1,n})$. For a given clique c , the nonnegative definite potential $V_c(\phi_c)$ defines how the phase of the neighboring pixels in the clique interact. MR phase images usually contain two components: a spatially smooth component due to inhomogeneities in the static B_0 field, and a nonsmooth component due to sudden magnetic susceptibility changes from one material to another (e.g., tissue/bone/air interface). A wide range of potential functions have been studied in the image restoration literature, which can handle effectively this type of mixed image characteristics [28]. Focusing on the first-order neighborhood system, we have

$$p(\Phi) \propto \exp\left(-\sum_{(m,n) \in S} V(\phi_{m,n}, \phi_{m,n-1}) - \sum_{(m,n) \in S} V(\phi_{m,n}, \phi_{m-1,n})\right). \quad (14)$$

Taking the potential as the square of adjacent differences leads to a Gauss-Markov prior

$$V(\phi_{m,n}, \phi_{m,n-1}) = (\phi_{m,n} - \phi_{m,n-1})^2 \quad (15)$$

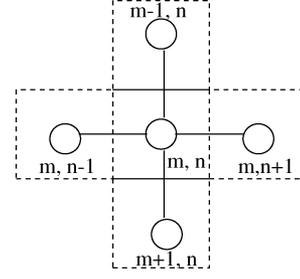


Fig. 1. The structure of the first-order neighborhood system. Pixel (m, n) , represented by a circle at the center, is conditioned only on its four adjacent pixels.

which produces smooth images with very low probability of sharp transitions in intensity [31]. To accommodate an increased probability of sharp transitions, other convex potential functions may be chosen, such as the generalized p -Gaussian model where $V(\phi_{m,n}, \phi_{m,n-1}) = |\phi_{m,n} - \phi_{m,n-1}|^p$, $1 < p < 2$ [29], the Huber prior in which the potential function switches from quadratic to linear at user-specified transition locations [32], [33], and $V(\phi_{m,n}, \phi_{m,n-1}) = \log \cosh[c(\phi_{m,n} - \phi_{m,n-1})]$ where c is a user-specified parameter [34]. Non-convex potential functions, such as $V(\phi_{m,n}, \phi_{m,n-1}) = (\phi_{m,n} - \phi_{m,n-1})^2 / ((\phi_{m,n} - \phi_{m,n-1})^2 + \eta^2)$ [35] with a constant η , have been employed to allow for even sharper intensity transitions. Higher-order neighborhood systems have also been used and proven to be able to capture correlation structures having a greater spatial complexity [36].

All the above functions are applicable to phase unwrapping, although one may be preferred over another in a specific application. It is also worth mentioning that the true phase of an MR image is not necessarily smooth, so the unwrapped results based on the above smooth models are not necessarily the same as the original true phase. However, the unsmooth portion of the phase cannot be recovered since the information is already lost in phase wrapping. Therefore, the smoothness is a practical optimality criterion, not a physical property of MRI, and is used in all existing phase unwrapping methods. The difference here is that it offers a more general mathematical definition for smoothness by the use of an MRF model. In the following section, we employ only the first-order MRF as an example to demonstrate the proposed method. Generalization of the proposed unwrapping algorithm to higher-order MRF's will be discussed in Section IV.

Given the above specific likelihood and prior probabilities, the MAP estimator in (9) becomes

$$\hat{\Phi} = \arg \max_{\Phi} \left\{ \prod_{m,n} \delta[\psi_{m,n} - \mathcal{W}(\phi_{m,n})] \times \exp\left(-\sum_{m,n} [V(\phi_{m,n}, \phi_{m,n-1}) + V(\phi_{m,n}, \phi_{m-1,n})]\right) \right\}. \quad (16)$$

The likelihood function that multiplies the exponential requires that the unwrapped phase differ from the wrapped phase at each pixel by an integer multiple of 2π , where the integer is termed the “wrap count” and is denoted by $k_{m,n}$ for pixel (m, n) , and \mathbf{k} for the matrix. As a result, solving (16) is equivalent to searching for a set of integers $\hat{\mathbf{k}}$ such that $\hat{\Phi} = \Psi + 2\hat{\mathbf{k}}\pi$ satisfies the

MRF requirement in the exponential of (16). Thereby, (16) can be further reduced to minimizing an overall potential $U(\mathbf{k}|\Psi)$ over \mathbf{k} given the wrapped phase Ψ

$$\begin{aligned} \hat{\mathbf{k}} &= \arg \min_{\mathbf{k}} U(\mathbf{k}|\Psi) \\ &= \arg \min_{\mathbf{k}} \\ &\quad \times \sum_{m,n} [V(\psi_{m,n} + 2\pi k_{m,n}, \psi_{m-1,n} + 2\pi k_{m-1,n}) \\ &\quad + V(\psi_{m,n} + 2\pi k_{m,n}, \psi_{m,n-1} + 2\pi k_{m,n-1})] \end{aligned} \quad (17)$$

where $\arg \min_{\mathbf{k}}$ denotes the minimizing value of \mathbf{k} in its domain of integer matrices.

B. Optimization Algorithm

In solving the optimization problem in (17), enumeration of all possible solutions is computationally infeasible for realistic image sizes. Although simulated annealing [37] is widely used, it still is computationally very intensive [38]. We propose a novel, efficient approach to the optimization in (17). The approach breaks a phase image into a number of blocks, which, when following a certain raster scan, span the entire image. As a result, the 2-D optimization problem is simplified to a series of 1-D ones with each being solved by dynamic programming [39]. The approach also employs an iterated conditional modes (ICMs) algorithm [40] to guarantee the sequential block-wise optimizations result in a global one over the entire image.

We show first how the dynamic programming can be applied to find the minimizing wrap counts for a certain block given the criterion in (17). Given an image block, dynamic programming searches for an optimal sequence, with each element of the sequence consisting of the wrap counts for a column of the phase image. For an M by N block, there are N elements in the sequence, denoted by column vectors \underline{k}_n , $n = 1, \dots, N$. Each vector is M -dimensional, corresponding to the wrap counts of M pixels in a single column, i.e., $\underline{k}_n = [k_{1,n}, k_{2,n}, \dots, k_{M,n}]^T$. The optimum of the sequence is defined by the overall potential $U(\mathbf{k}|\Psi)$, defined in (17). To minimize this potential, we decompose $U(\mathbf{k}|\Psi)$ into a summation of N potentials

$$U_1(\underline{k}_1) = \sum_{m=2}^{m=M} V(\psi_{m,1} + 2\pi k_{m,1}, \psi_{m-1,1} + 2\pi k_{m-1,1})$$

and

$$\begin{aligned} U_n(\underline{k}_{n-1}, \underline{k}_n) &= \sum_{m=2}^{m=M} V(\psi_{m,n} + 2\pi k_{m,n}, \psi_{m-1,n} + 2\pi k_{m-1,n}) \\ &\quad + \sum_{m=1}^{m=M} V(\psi_{m,n} + 2\pi k_{m,n}, \psi_{m,n-1} + 2\pi k_{m,n-1}) \end{aligned} \quad (18)$$

for $n = 2, \dots, N$. Note that the first potential depends only on column one, and the n th depends only on columns $n - 1$ and n . Dynamic programming takes advantage of this local dependence and searches for the best wrap counts for one column at a time. Specifically, the algorithm first searches for the wrap-count vector of the first column \underline{k}_1 that minimizes the accumulated potential $U_1 + U_2$. Only these minimized values (a function of \underline{k}_2) are passed to the next step and added to U_3 to update the accumulated potential; the accumulated potential is then used as the objects for minimization with respect to \underline{k}_2 . This procedure is repeated until the last column of the image is reached, at

which point the final accumulated potential is minimized. The procedure can be summarized mathematically as follows:

$$\begin{aligned} D_1(\underline{k}_2) &= \min_{\underline{k}_1} [U_1(\underline{k}_1) + U_2(\underline{k}_1, \underline{k}_2)] \\ D_2(\underline{k}_3) &= \min_{\underline{k}_2} [D_1(\underline{k}_2) + U_3(\underline{k}_2, \underline{k}_3)] \\ &\quad \vdots \\ D_{N-1}(\underline{k}_N) &= \min_{\underline{k}_{N-1}} [D_{N-2}(\underline{k}_{N-1}) + U_N(\underline{k}_{N-1}, \underline{k}_N)] \\ D_N &= \min_{\underline{k}_N} D_{N-1}(\underline{k}_N). \end{aligned} \quad (19)$$

At each step, the optimal \underline{k}_n values are determined and recorded as

$$\begin{aligned} \hat{\underline{k}}_1(\underline{k}_2) &= \arg \min_{\underline{k}_1} [U_1(\underline{k}_1) + U_2(\underline{k}_1, \underline{k}_2)] \\ \hat{\underline{k}}_2(\underline{k}_3) &= \arg \min_{\underline{k}_2} [D_1(\underline{k}_2) + U_3(\underline{k}_2, \underline{k}_3)] \\ &\quad \vdots \\ \hat{\underline{k}}_{N-1}(\underline{k}_N) &= \arg \min_{\underline{k}_{N-1}} [D_{N-2}(\underline{k}_{N-1}) + U_N(\underline{k}_{N-1}, \underline{k}_N)] \\ \hat{\underline{k}}_N &= \arg \min_{\underline{k}_N} D_{N-1}(\underline{k}_N). \end{aligned} \quad (20)$$

These minimizing wrap counts comprise the surviving history H , which keeps track of the best sequence at each stage. Specifically, the surviving history is initialized by setting $\{H(\underline{k}_1)\}$ to be a list of all possible realizations of \underline{k}_1 , and then updated by

$$\{H(\underline{k}_n)\} = \left\{ H(\hat{\underline{k}}_{n-1}(\underline{k}_n)), \underline{k}_n \right\}, \quad n = 2, \dots, N \quad (21)$$

for all possible realizations of \underline{k}_n . For the last column, the optimal sequence is obtained by tracing back the surviving history of $\{H(\hat{\underline{k}}_N)\}$. Upon conclusion of the minimization step, the wrapped phase is adjusted appropriately using the corresponding wrap counts to obtain the desired unwrapped phase.

The proposed dynamic programming procedure is illustrated in Fig. 2, where each unit in the n th column of the figure represents a possible \underline{k}_n for the n th column of the image. Each transition from one unit ($n - 1$) to another (n) represents an increment of the accumulated potential $D_{n-2}(\underline{k}_{n-1}) + U_n(\underline{k}_{n-1}, \underline{k}_n)$. Among all the possible combinations of transitions interconnecting units, only one is chosen, which minimizes (19). The highlighted path is obtained by tracing back the surviving history, and the units along the path give the optimal wrap counts.

Compared with exhaustive search, this dynamic programming method reduces the computational complexity from $O(K^{MN})$ down to $O(K^M N)$, where K is the maximum wrap count. Block size is an important factor for the algorithm. Clearly, in the extreme case that no block decomposition is done, the final phase image is obtained after the above optimization procedure. However, this procedure is computationally expensive for a realistic image size (e.g., 256×256). On the other hand, if an image is divided into very small blocks (e.g., block size is $2 \times N$), the optimization becomes easy and efficient within each block, but can be challenging to achieve a unified optimum when the blocks are combined to yield the final image. To address this issue, we employ the idea of ICMs [40]. The ICM algorithm uses a ‘‘greedy,’’ iterative strategy in optimization. To maximize the posterior probability $p(\Phi|\Psi)$,

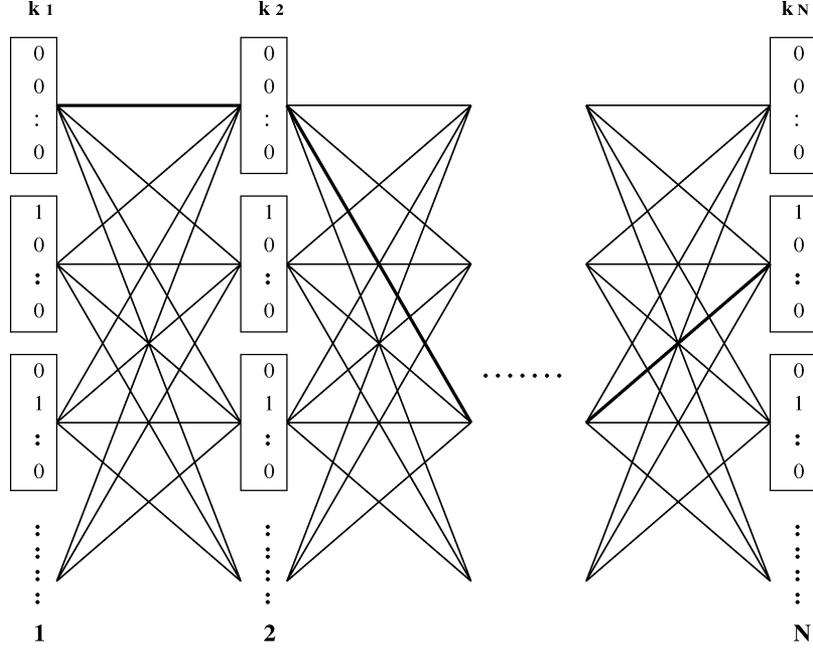


Fig. 2. Illustration of dynamic programming which searches for an optimal sequence of k_1, k_2, \dots, k_N . Each unit in the n th column denotes a realization of k_n . The connections between units represent accumulations of potentials. The minimum of these accumulated potentials gives the optimal sequence, shown as the highlighted path of connections.

the algorithm sequentially updates each block $\hat{\Phi}_{\text{block}}^{(i)}$ into $\hat{\Phi}_{\text{block}}^{(i+1)}$ by maximizing the conditional probability, given the wrapped phase and the provisional estimate at the rest blocks $\hat{\Phi}_{S \setminus \text{block}}$. Specifically

$$\hat{\Phi}_{\text{block}}^{(i+1)} = \max_{\Phi_{\text{block}}} p\left(\Phi_{\text{block}} | \Psi, \hat{\Phi}_{S \setminus \text{block}}^{(i)}\right) \quad (22)$$

where $S \setminus \text{block}$ denotes the set of pixels outside the block under evaluation. Invoking the assumption that the likelihood functions are independently and identically distributed (i.i.d) pixel-wise and the prior probability is Markovian, we can rewrite the above conditional probability as

$$p\left(\Phi_{\text{block}} | \Psi, \hat{\Phi}_{S \setminus \text{block}}^{(i)}\right) \propto p(\Psi | \Phi_{\text{block}}) p\left(\Phi_{\text{block}} | \hat{\Phi}_{\mathcal{N}_{\text{block}}}^{(i)}\right) \quad (23)$$

where $\mathcal{N}_{\text{block}}$ denotes the neighboring blocks. Following the same derivation from (16) to (17), maximizing the right-hand side of (23), which is an exponential function defined by the MRF model, is equivalent to minimizing the corresponding potential using

$$\hat{\mathbf{k}}_{\text{block}}^{(i+1)} = \min_{\mathbf{k}_{\text{block}}} U\left(\mathbf{k}_{\text{block}} | \Psi, \hat{\mathbf{k}}_{\mathcal{N}_{\text{block}}}^{(i)}\right) \quad (24)$$

where $U(\mathbf{k}_{\text{block}} | \Psi, \hat{\mathbf{k}}_{\mathcal{N}_{\text{block}}}^{(i)})$ is the accumulated potentials of the target block. Compared with the objective function $U(\mathbf{k} | \Psi)$ in (17), the only difference in (24) is the additional boundary constraints introduced by the neighboring blocks. Therefore, the same dynamic programming procedure described above can be used to solve the minimization problem in (24) to obtain the wrap counts within each block. For example, if a block composed of two complete rows is chosen, the potential to be minimized in (17) becomes $U(\mathbf{k}_{\text{row } m \ \& \ m+1} | \Psi, \hat{\mathbf{k}}_{\text{row } m-1 \ \& \ m+2})$

in (24), which can be decomposed into summation of partial potentials that are

$$\begin{aligned} U_n \left(\begin{matrix} k_{m,n-1} & k_{m,n} & \hat{k}_{m-1,n} \\ k_{m+1,n-1} & k_{m+1,n} & \hat{k}_{m+2,n} \end{matrix} \right) \\ = V(\psi_{m,n} + 2\pi k_{m,n}, \psi_{m,n-1} + 2\pi k_{m,n-1}) \\ + V(\psi_{m+1,n} + 2\pi k_{m+1,n}, \psi_{m,n} + 2\pi k_{m,n}) \\ + V(\psi_{m+1,n} + 2\pi k_{m+1,n}, \psi_{m+1,n-1} + 2\pi k_{m+1,n-1}) \\ + V(\psi_{m,n} + 2\pi k_{m,n}, \psi_{m-1,n} + 2\pi \hat{k}_{m-1,n}) \\ + V(\psi_{m+2,n} + 2\pi \hat{k}_{m+2,n}, \psi_{m+1,n} + 2\pi k_{m+1,n}) \end{aligned} \quad (25)$$

where “ \hat{k} ” denotes the provisional estimates assumed to be fixed in the current iteration. The above decomposition can be readily generalized to different block shapes (e.g., two columns) and block sizes (e.g., three or four rows or columns). When applied to each block in turn, the above procedure is repeated according to a well-defined raster scan scheme (e.g., top-to-bottom, left-to-right), which defines an updating cycle of ICM. The iteration continues until convergence. The convergence is guaranteed for the serial updating and is rapid [40]. We term this optimization approach “structured iterated conditional modes” (SICMs). We have heuristically found that alternation of the raster scan scheme from one iteration to another results in convergence to an extreme closer to the true value. In addition, the result obtained by ICM depends on the initial provisional estimate [28], and it is currently not known how to set the initialization properly to obtain a good solution. In our implementation, we used the same block-by-block optimization in (24), except that the conditional probability depends only on the blocks that are already estimated.

The computational complexity of this approach is the power of the smaller dimension of the block; for example, if a block of two complete rows is chosen, the complexity becomes $O(K^2MN)$ per iteration, which is linear to the image size.

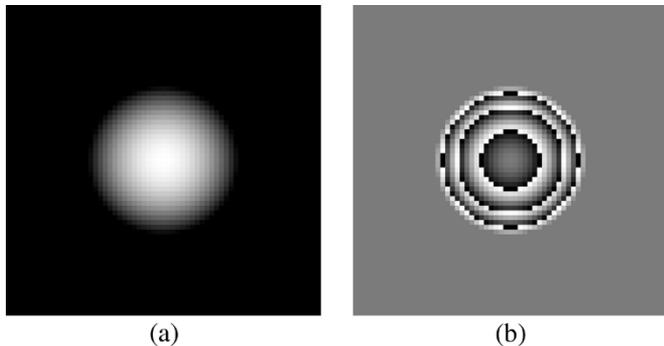


Fig. 3. (a) Simulated true phase image, which is a parabola in two dimensions and (b) its corresponding wrapped phase.

III. RESULTS

We have tested our proposed algorithm on various MRF models using various simulated and experimental data having different noise levels. In this section, we present some results based on a first-order Gaussian MRF statistical model. No significant differences have been observed in the results computed using other MRF models. For the first-order MRF, it is sufficient to use a 2 by N (or M by 2) block in the optimization step. All examples shown in this section were calculated using a row-by-row and top-to-bottom raster scan (first iteration), followed sequentially by a column-to-column/left-to-right scan, a row-by-row/bottom-to-top scan, then a column-to-column/right-to-left scan. This scheme is then repeated until convergence.

A. Simulation Studies

We have tested the performance of our algorithm on synthetic data, an example of which is shown in Fig. 3(a), where the true phase function, Φ_{clean} , is a parabola located at the center of the image. We compared the proposed algorithm with the least squares (LS) algorithm—a representative of minimum-norm methods—and the algorithm of Chen and Zebker [5], which is a path-following approach. In the noise-free case, all algorithms can perfectly recover the true phase from the wrapped phase shown in Fig. 3(b).

In the presence of additive noise, the wrapped phase becomes $\mathcal{W}(\Phi_{\text{clean}} + \xi)$, where ξ denotes the additive random noise. In this case, phase unwrapping algorithms attempt to obtain $\hat{\Phi} = \Phi_{\text{clean}} + \xi$, without any attempt to remove noise. The resulting $\hat{\Phi}$ may deviate considerably from the noise-free phase values as noise increases, leading to errors much larger than the noise level. To compare the robustness of our algorithms in the presence of noise, we plot the mean-squared-error (MSE) as a function of the signal-to-noise ratio (SNR) in Fig. 4, where the MSE is computed using the noise-free phase as a reference. Both the proposed algorithm and the algorithm of Chen and Zebker have similar performance, while the MSE of the LS method increases at a faster rate than the others, as noise increases. In particular, we show in Fig. 5(a) and (b) the error images ($\hat{\Phi} - \Phi_{\text{clean}}$) for both the LS and the proposed algorithms

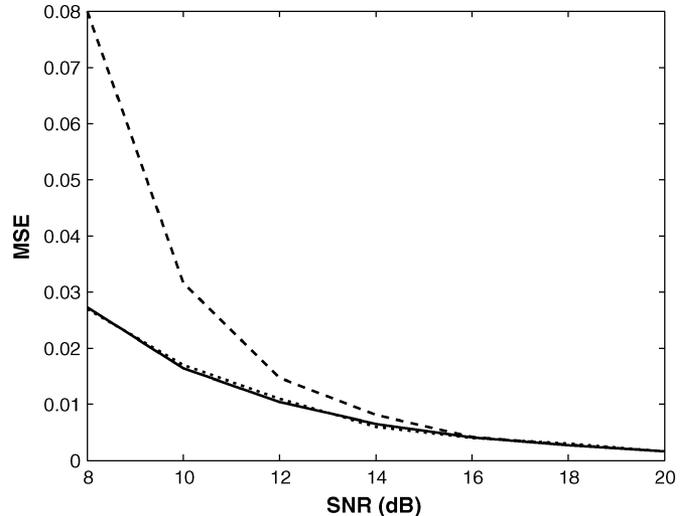


Fig. 4. MSE of the unwrapped results as a function of SNR. Solid line, dotted line, and dashed line represent, respectively, the proposed method, the algorithm of Chen and Zebker, and the LS method.

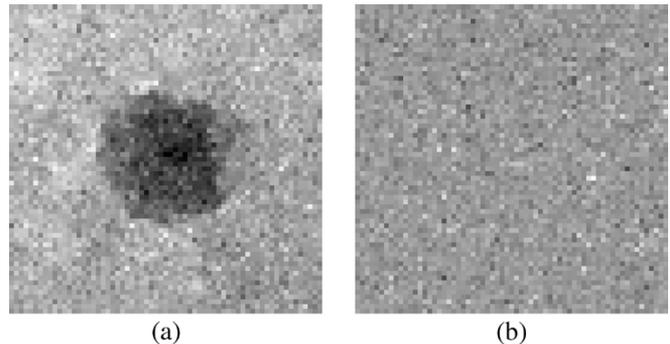


Fig. 5. Difference images between the true phase and the unwrapped phase based on (a) the LS method and (b) the proposed method at SNR = 12 dB.

when SNR = 12 dB. The result using the algorithm of Chen and Zebker is not shown as it is quite close to Fig. 5(b). The results suggest that even with low SNR, the proposed algorithm is able to accurately estimate the wrap counts so that the error appears only as random noise. The LS solutions usually underestimate the local-average phase gradient when noise is large. This behavior is analogous to least-squares curve fitting, in which the fitted curve is always smoother than the measurements. Compared with the LS method, the proposed method is significantly more robust.

B. Experimental Studies

The algorithms have been tested using several MR data sets. A set of representative results from a phantom experiment is shown in Fig. 6. The phantom image was acquired on a 1.5T MRI using a fat imaging with steady-state precession sequence, with $T_R = 45$ ms and $T_E = 12$ ms. The magnitude image was segmented to generate a mask having unity value inside the object and zero for the background. The wrapped phase image, weighted by the mask, was then unwrapped by the three algorithms. Fig. 6(a)–(d) shows the wrapped phase image, the

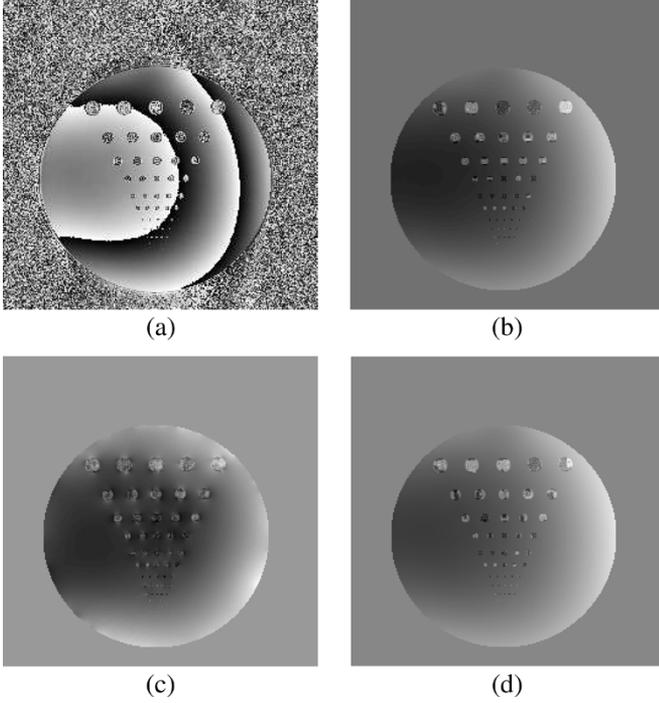


Fig. 6. MR phase images of a phantom: (a) wrapped phase image, and phase images unwrapped by (b) the proposed method, (c) the LS method, and (d) the algorithm of Chen and Zebker.

unwrapped phase image by the proposed algorithm, the LS method, and the algorithm of Chen and Zebker, respectively. All methods yield a slowly varying unwrapped phase within the big circle. The noise within the small circles is seen to be large and is problematic. The error due to the noise in the LS method is not limited to the boundaries of the small circles, but in fact propagates and affects other regions in the larger circle. This introduces additional error in the unwrapped phase. Compared with Fig. 6(c), the errors of Fig. 6(b) and (d) are more localized to the small circles and do not affect the larger region as significantly. The proposed method and the algorithm of Chen and Zebker, thus, have better performance than does the LS method.

Another set of representative results for MR head data is shown in Fig. 7. The data set was obtained *in vivo* from a 1.5T MRI using a gradient-echo sequence with $T_E = 10$ ms and $T_R = 200$ ms. Fig. 7(a) shows the wrapped phase image. It can be seen that the phase variation inside the brain is slow, but becomes rapid at the nasal, ear, and oral regions due to magnetic susceptibility changes. In unwrapping the phase, similar to the procedure for the phantom, a mask generated by segmenting the magnitude image is applied prior to unwrapping. To compare their performance, we show the unwrapped phase computed by the proposed method, the LS method, and the algorithm of Chen and Zebker in Fig. 7(b)-(d), respectively. As can be seen, the proposed algorithm performed quite well in removing the 2π phase jumps of Fig. 7(a). The phase image produced by the LS method appears more blurred, while artifacts appear in the phase

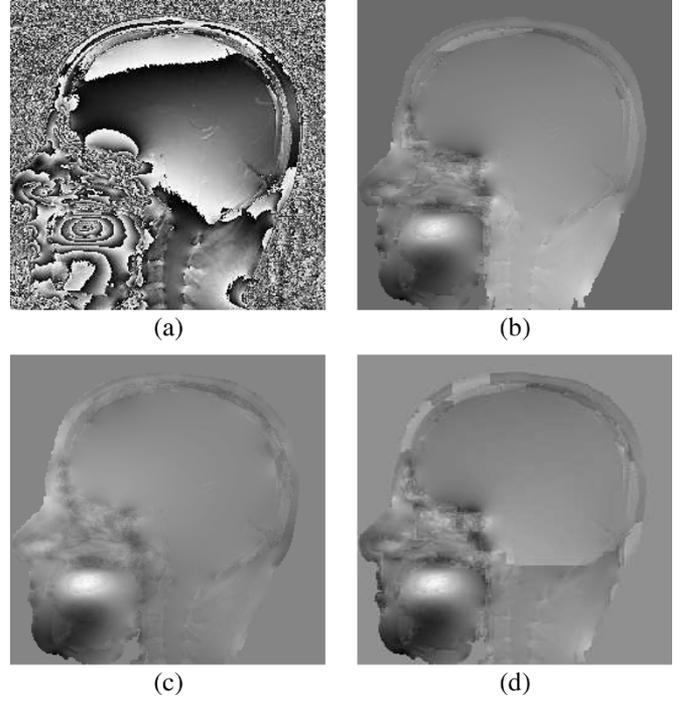


Fig. 7. MR phase images of a human head: (a) wrapped phase image, and unwrapped phase images by (b) the proposed method, (c) the LS method, and (d) the algorithm of Chen and Zebker.

image reconstructed using the algorithm of Chen and Zebker, above the neck and about the skull.

IV. DISCUSSION

The proposed method has a number of appealing properties. First, as in the LS method and the algorithm of Chen and Zebker, the proposed method yields an exact solution if the smoothness condition in (6) is satisfied. This can be shown using the following inequality:

$$\begin{aligned} & \min_{\Delta \mathbf{k}} \sum_{m,n} [V(\psi_{m+1,n} + 2\pi k_{m+1,n}, \psi_{m,n} + 2\pi k_{m,n}) \\ & \quad + V(\psi_{m,n+1} + 2\pi k_{m,n+1}, \psi_{m,n} + 2\pi k_{m,n})] \\ & \leq \sum_{m,n} \left[\min_{\Delta_x k_{m,n}} V(\psi_{m+1,n} + 2\pi k_{m+1,n}, \psi_{m,n} + 2\pi k_{m,n}) \right. \\ & \quad \left. + \min_{\Delta_y k_{m,n}} V(\psi_{m,n+1} + 2\pi k_{m,n+1}, \psi_{m,n} + 2\pi k_{m,n}) \right] \end{aligned}$$

where the left-hand side of the inequality corresponds to the proposed algorithm. When the potential is a nonnegative function of the difference between its two arguments, the right-hand side gives the solution $\Delta_x k_{m,n} = \langle -(\Delta_x \psi_{m,n} / 2\pi) \rangle$, and $\Delta_y k_{m,n} = \langle -(\Delta_y \psi_{m,n} / 2\pi) \rangle$, which is the same as the solution in Lemma 1.2. The equality holds if $\Delta \mathbf{k}$ is a gradient field, which is satisfied when (6) holds. Therefore, when the true phase is sufficiently smooth, the solution of the proposed method yields exactly the same solution as in Lemma 1.2.

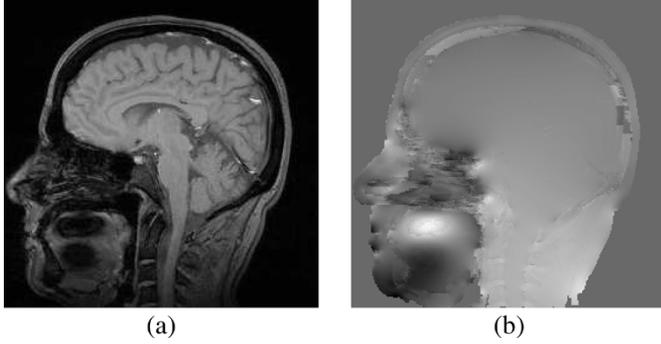


Fig. 8. (a) MR magnitude image of the human head. (b) The unwrapped phase image by the proposed method weighted by the magnitude.

Second, it is easy to incorporate a quality map into the proposed method to further improve its robustness. In MR phase unwrapping, the magnitude image, $|I_{m,n}|$, serves as a natural quality map that assigns greater weights to pixels having large magnitude values. To do so, we can simply modify the cost function as follows:

$$\begin{aligned}
 U(\mathbf{k}) = & \sum_{m,n} [V(\psi_{m,n} + 2\pi k_{m,n}, \psi_{m-1,n} + 2\pi k_{m-1,n}) \\
 & \times \frac{(|I_{m,n}| + |I_{m-1,n}|)}{2} \\
 & + V(\psi_{m,n} + 2\pi k_{m,n}, \psi_{m,n-1} + 2\pi k_{m,n-1}) \\
 & \times \frac{(|I_{m,n}| + |I_{m,n-1}|)}{2}]. \quad (26)
 \end{aligned}$$

Fig. 8 shows how the magnitude-weighting helps to improve the phase unwrapping results in the nasal region.

Third, as with all other path-following methods, the proposed algorithm satisfies the congruence constraint. The complex image with unwrapped phase is exactly the same as the measured complex image with true phase, a result desirable in most applications. This is in contrast to the LS method, where this property does not hold in general, unless (6) is true.

A drawback of the SICM algorithm, common to most other path-following methods, is that the search for the optimal path always starts from a reference pixel, whose phase value is assumed to be free from wrapping ambiguity. If this assumption fails, the resulting unwrapped phase map may have a constant offset. In most applications, however, this constant offset is not problematic—for example, in correction of artifacts due to field inhomogeneities based on field map measurements.

In addition to the above properties common to most existing methods, the proposed algorithm has several unique properties that may be advantageous in certain applications. Specifically, it has been shown experimentally that the algorithm is robust to noise and is able to better handle abrupt phase variations; this may be a result of using the MRF model and the SICM optimization method. More importantly, the SICM algorithm and related higher-order derivatives are able to efficiently handle any model based on local interactions. As an example, Fig. 9 illustrates how the SICM algorithm may be generalized to fit a third-order MRF. This flexibility offers an advantage over the existing methods (e.g., the algorithm of Chen and Zebker) which are based on models having interactions between adjacent pixels only. The efficiency and generality of the SICM algorithm may make it possible to use data-dependent and

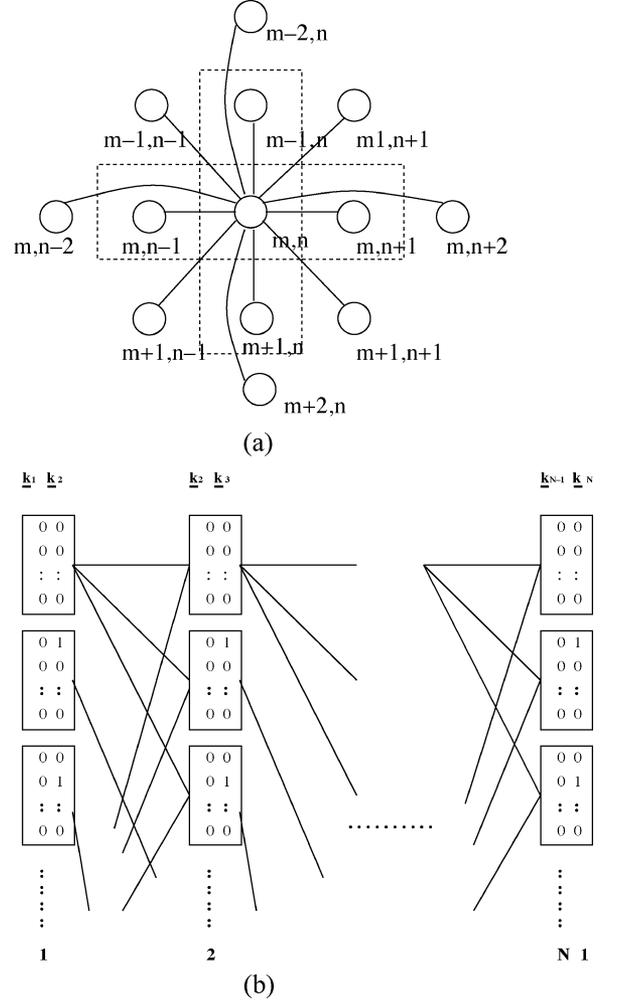


Fig. 9. (a) Structure of a third-order MRF and (b) the corresponding dynamic programming realization. In searching for the optimal sequence of k_1, k_2, \dots, k_N , each unit in the n th column denotes a realization of the k_n, k_{n+1} pair.

spatially varying MRF models to solve more challenging phase unwrapping problems, where existing algorithms fail because of inadequacy of the phase model used. It is also worth noting that although higher-order MRF's may be more accurate, the increased capability comes at the expense of computational complexity, which increases exponentially with the order of the model. Further investigations into the trade-offs associated with more complex MRF models are needed and would be a direction of our future work.

In general, the SICM algorithm is efficient in computation and memory usage. The high computational efficiency results primarily from the fact that ICM reduces a 2-D problem to a series of 1-D problems to which dynamic programming can be applied. The high efficiency in memory usage is due to its repeated use of the same memory at each step of dynamic programming [41], making the method especially attractive for hardware implementation. The proposed algorithm is not as fast as the LS method (same speed for a fixed image size). However, it is difficult to make a fair comparison with the algorithm of Chen and Zebker, because its computational complexity has not yet been quantified. The speeds of both the proposed method and the algorithm of Chen and Zebker depend on the number of wraps, as

well as the content of the image. For the MR phase images we have evaluated, the proposed method converged faster than the algorithm of Chen and Zebker. But the algorithm of Chen and Zebker may be faster for unwrapping synthetic aperture radar phase images with hundreds of wraps, since the complexity of the proposed method increases exponentially with the number of wrap counts. To overcome this problem, we may modify the proposed method by restricting the wrap count at each pixel to be -1 , 0 or $+1$ at each iteration, and then improve the results iteratively. However, this reduction in computational complexity is at the expense of performance (associated with the suboptimal search strategy used).

V. CONCLUSION

Two-dimensional phase unwrapping is a very challenging problem; the challenges lie in both modeling the true phase and solving the underlying optimization problem efficiently. In this paper, we propose a novel algorithm for unwrapping MR phase images. Based on the phase characteristics of MR images, we use a flexible MRF for modeling the phase, accommodating the smooth characteristics of the true phase of an MR image, while also permitting abrupt phase changes. Using this statistical model, we also propose an efficient optimization method based on ICM along with a dynamic programming approach for fitting the unwrapped phase to the model. Experimental results demonstrate that the proposed algorithm is robust in the presence of noise, and able to effectively handle rapid, large phase variations encountered in gradient-echo imaging. It will hopefully prove useful for a variety of practical applications, such as field mapping, flow imaging, and temperature mapping, applications which depend on an accurate unwrapping of phase images.

ACKNOWLEDGMENT

The authors would like to thank the editor, associate editor, and anonymous reviewers for their careful reading and helpful comments and suggestions that have greatly enhanced this paper.

REFERENCES

- [1] Z.-P. Liang, "A model-based method for phase unwrapping," *IEEE Trans. Med. Imag.*, vol. 15, no. 6, pp. 893–897, Dec. 1996.
- [2] S. M. Song, S. Napel, N. J. Pelc, and G. H. Glover, "Phase unwrapping of MR phase images using Poisson equation," *IEEE Trans. Image Process.*, vol. 4, no. 5, pp. 667–676, May 1995.
- [3] K. Itoh, "Analysis of the phase unwrapping problem," *Appl. Opt.*, vol. 21, no. 14, p. 2470, Jul. 1982.
- [4] D. C. Ghiglia, G. A. Mastin, and L. A. Romero, "Cellular automata method for phase unwrapping," *J. Opt. Soc. Amer. A*, vol. 4, no. 1, pp. 267–280, Jan. 1987.
- [5] C. W. Chen and H. A. Zebker, "Network approaches to two-dimensional phase unwrapping: intractability and two new algorithms," *J. Opt. Soc. Amer. A*, vol. 17, no. 3, pp. 401–414, Mar. 2000.
- [6] S. Chavez, Q. S. Xiang, and A. Li, "Understanding phase maps in MRI: a new outline phase unwrapping method," *IEEE Trans. Med. Imag.*, vol. 21, no. 8, pp. 966–977, Aug. 2002.
- [7] J. M. Huntley, "Three-dimensional noise-immune phase unwrapping algorithm," *Appl. Opt.*, vol. 40, pp. 3901–3908, 2001.
- [8] M. Jenkinson, "Fast, automated, N-dimensional phase unwrapping algorithm," *Magn. Reson. Med.*, vol. 49, pp. 193–197, 2003.
- [9] D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. New York: Wiley, 1998.
- [10] O. Loffeld, C. Arndt, and A. Hein, "Estimating the derivative of modulated phases," in *Proc. IEEE ICASSP*, 1997, pp. 2841–2844.
- [11] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: two-dimensional phase unwrapping," *Radio Sci.*, vol. 23, pp. 713–720, Jul.–Aug. 1988.
- [12] M. Costantini, "A novel phase unwrapping method based on network programming," *IEEE Trans. Geosci. Remote Sens. E.*, vol. 36, no. 3, pp. 813–821, May 1998.
- [13] T. J. Flynn, "Two-dimensional phase unwrapping with minimum weighted discontinuity," *J. Opt. Soc. Amer. A*, vol. 14, no. 10, pp. 2692–2701, Oct. 1997.
- [14] G. Nico, G. Palubinska, and M. Datcu, "Bayesian approach to phase unwrapping: theoretical study," *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2545–2556, Sep. 2000.
- [15] C. W. Chen and H. A. Zebker, "Two-dimensional phase unwrapping with use of statistical models for cost functions in nonlinear optimization," *J. Opt. Soc. Amer. A*, vol. 18, no. 2, pp. 338–351, Feb. 2001.
- [16] M. D. Pritt and J. S. Shipman, "Least-squares two-dimensional phase unwrapping using FFTs," *IEEE Trans. Geosci. Remote Sens. E.*, vol. 32, no. 3, pp. 706–708, May 1994.
- [17] D. Kerr, G. H. Kaufmann, and G. E. Galizzi, "Unwrapping of interferometric phase-fringe maps by the discrete cosine transform," *Appl. Opt.*, vol. 35, no. 5, pp. 810–816, Feb. 1996.
- [18] M. D. Pritt, "Phase unwrapping by means of multigrid techniques for interferometric SAR," *IEEE Trans. Geosci. Remote Sens. E.*, vol. 34, no. 3, pp. 728–738, May 1996.
- [19] D. J. Bone, "Fourier fringe analysis: the two-dimensional phase unwrapping problem," *Appl. Opt.*, vol. 30, no. 25, pp. 3627–3632, 1991.
- [20] D. P. Towers, T. R. Judge, and P. J. Bryanston-Cross, "Automatic interferogram analysis techniques applied to quasi-heterodyne holography and ESPI," *Opt. Laser Eng.*, vol. 14, pp. 239–281, 1991.
- [21] N. H. Ching, D. Rosenfeld, and M. Braun, "Two dimensional phase unwrapping using a minimum spanning tree algorithm," *IEEE Trans. Image Process.*, vol. 1, no. 3, pp. 355–365, Jul. 1992.
- [22] L. An, Q. S. Xiang, and S. Chavez, "A fast implementation of the minimum spanning tree method for phase unwrapping," *IEEE Trans. Med. Imag.*, vol. 19, no. 8, pp. 805–808, Aug. 2000.
- [23] M. Takeda and T. Abe, "Phase unwrapping by a maximum cross-amplitude spanning tree algorithm: a comparison study," *Opt. Eng.*, vol. 35, no. 8, pp. 2345–2351, 1996.
- [24] J. M. B. Dias and J. M. N. Leitao, "The $Z\pi M$ algorithm: a method for interferometric image reconstruction in SAR/SAS," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 408–422, Apr. 2002.
- [25] J. M. N. Leitao and M. A. T. Figueiredo, "Absolute phase image reconstruction: a stochastic nonlinear filtering approach," *IEEE Trans. Image Process.*, vol. 7, no. 6, pp. 868–882, Jun. 1998.
- [26] R. Koetter, B. J. Frey, N. Petrovic, and D. C. Munson Jr, "Unwrapping phase images by propagating probabilities across graphs," in *Proc. Int. Conf. Acoustics Speech and Signal Processing (ICASSP)*, 2001, pp. 1845–1848.
- [27] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Roy. Statist. Soc., B*, vol. 36, pp. 192–236, 1974.
- [28] S. Z. Li, *Markov Random Modeling in Computer Vision*. Tokyo, Japan: Springer-Verlag, 1995.
- [29] C. Bouman and K. Sauer, "A unified approach to statistical tomography using coordinate descent optimization," *IEEE Trans. Image Process.*, vol. 5, no. 3, pp. 480–492, Mar. 1996.
- [30] K. Held, E. R. Kops, B. J. Krause, W. M. Wells, R. Kikinis, and H.-W. Müller-Gärtner, "Markov random field segmentation of brain MR images," *IEEE Trans. Med. Imag.*, vol. 16, no. 6, pp. 878–886, Dec. 1997.
- [31] R. M. Leahy and J. Qi, "Statistical approaches in quantitative positron emission tomography," *Statist. Computing*, vol. 10, pp. 147–165, 2000.
- [32] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.
- [33] J. Qi, R. M. Leahy, S. R. Cherry, A. Chatzioannou, and T. H. Farquhar, "High resolution 3-D Bayesian image reconstruction using the microPET small animal scanner," *Phys. Med. Biol.*, vol. 43, pp. 1001–1013, 1998.
- [34] P. Green, "Bayesian reconstruction from emission tomography data using a modified EM algorithm," *IEEE Trans. Med. Imag.*, vol. 9, no. 1, pp. 84–93, Mar. 1990.
- [35] S. Geman and D. McClure, "Statistical methods for tomographic image reconstruction," *Bull. Int. Statist. Inst.*, vol. 52, pp. 5–21, 1987.
- [36] M. Chan, G. T. Herman, and E. Levitan, "Bayesian image reconstruction using a high-order interacting MRF model," in *Proc. Int. Conf. Image Anal. Process. (ICIAP)*, 1995, pp. 609–614.
- [37] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680, 1983.
- [38] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-6, no. 6, pp. 721–741, Jun. 1984.
- [39] R. E. Bellman and S. E. Dreyfus, *Applied Dynamic Programming*. Princeton, NJ: Princeton Univ. Press, 1962.
- [40] J. Besag, "On the statistical analysis of dirty pictures," *J. R. Statist. Soc., B*, vol. 48, no. 3, pp. 259–302, 1986.
- [41] A. J. Viterbi, "Convolutional codes and their performance in communication systems," *IEEE Trans. Commun. Technol.*, vol. Com-19, no. 10, pp. 751–772, Oct. 1971.