Bayesian Estimation of Smooth Parameter Maps for Dynamic Contrast-Enhanced MR Images with Block-ICM

B. Michael Kelm, Natalie Mueller, Bjoern H. Menze, Fred A. Hamprecht

Interdisciplinary Center for Scientific Computing

University of Heidelberg

INF 368, 69120 Heidelberg, Germany

http://klimt.iwr.uni-heidelberg.de/mip/

Abstract

In cancer, pathological tissue often exhibits abnormal perfusion and vascular permeability. These can be estimated by monitoring the abundance of an injected contrast medium over time, using Dynamic Contrast-Enhanced (DCE) MR Imaging. The resulting spatially resolved time curves are usually interpreted in terms of a pharmacokinetic model which is fitted by maximum likelihood. However, the resulting nonlinear least squares (NLLS) problem may exhibit spurious local optima leading to false parameter estimates at individual voxels in the generated parameter map. We propose the application of a spatial prior model in form of a generalized Gaussian Markov random field. By using information from parameter estimates at neighboring voxels and computing a maximum a posteriori solution for the whole parameter map at once, false local optima at individual voxels can be avoided. Since the number of variables gets very big for common image resolutions, standard NLLS solvers cannot be employed anymore. We therefore propose a generalized iterated conditional modes (ICM) approach operating on blocks instead of sites. Results on DCE-MR images of the prostate show less speckle noise in the resulting parameter maps. Furthermore, the mean square error (MSE) in the affected voxels is significantly smaller, thus reflecting a better fit.

1. Introduction

Dynamic Contrast-Enhanced MR Imaging (DCE-MRI) emerges as a novel and powerful imaging modality which has already proved useful in various clinical applications such as the examination of breast cancer [4], bone marrow [7] brain [5] and prostate tumors [15, 9, 11].

In contrast to traditional T1- and T2-weighted MR imaging modalities which mainly carry morphological information, DCE-MRI allows to derive physiological information.

However, DCE-MRI also requires more involved postprocessing strategies. In general these can be grouped into model-based and model-free approaches.

In model-based approaches, an appropriate pharmacokinetic model which describes the expected signal evolution is derived [5]. Its parameters are often associated with meaningful physiological properties of the examined tissue and are determined with a nonlinear least squares (NLLS) approach. Because of signal noise and local minima in the NLLS criterion these model fits can fail for individual voxels, resulting in speckled parameter maps. In [5] such voxels are identified by a signal-to-noise ratio criterion and removed from the parameter map, thus discarding possibly valuable information.

A different approach to cope with noise and unanticipated signal shapes is to reduce the use of prior knowledge with model-free approaches. For example, [13] requires a labeled data set to train an artificial neural network without an explicit physiological model. As opposed to that, [10] proposes the application of an unsupervised method: the high-dimensional time curve feature vectors are projected onto a two-dimensional manifold by means of self-organizing maps. The resulting 2D coordinates can be color-coded and mapped in a diagnostic image. However, model-free approaches often lack physiological interpretability and come with the flavor of "black-box" methods.

The approach taken in this paper is model-based. In addition to a pharmacokinetic model we assume that the characteristics of the tissue vary gradually from voxel to voxel and, hence, that the parameter map that best describes the physiological properties of the tissue should exhibit some spatial smoothness. The parameter map is modeled as a Markov random field (MRF) with pairwise potentials and the recorded MR volume is regarded as a noisy observation of the nonlinearly transformed hidden parameter map. Hence, the parameter estimate at each voxel is supported by

the estimates at its neighboring voxels which helps avoiding spurious local minima. Since incorporating a spatial prior also poses a greater computational challenge, we propose an efficient block-ICM algorithm for approximate inference.

2. Dynamic Contrast-Enhanced MR Imaging

DCE-MRI is used to track the diffusion of a paramagnetic contrast medium (CM) such as Gd-DTPA and study tissue perfusion and vascular permeability *in vivo* [5]. Therefore, DCE-MRI allows to detect pathologic tissue changes and can be used in tumor diagnosis [15, 9, 11, 4, 7].

During the intravenous injection of the CM, a sequence of several T1-weighted MR image volumes is recorded at intervals of a few seconds. Hence, a 3D volume of T1 intensity time curves is obtained. Their evaluation is based on a pharmacokinetic model whose parameters characterize the tracer accumulation properties of the underlying tissue (see [12] and references therein).

The basis for the quantitative evaluation in this paper is the widely used two-compartment model by Brix *et al.* [5]. The resulting T1 intensity dynamics are described by the model function

$$\frac{S_{CM}(t)}{S_0} = \begin{cases} 1 & t \leq t_0 \\ 1 + A\, C_{CM}(t - t_0) & t_0 < t \end{cases} \quad (1)$$

where S_0 describes the T1 intensity obtained without CM and t_0 the lag time. The amplitude A depends on several tissue properties, the employed MR sequence and the infusion rate of the CM [5]. It usually increases in tumors. The concentration of the CM evolves as

$$C_{CM}(t) = v\frac{\exp(k_{el}t') - 1}{\exp(k_{el}t)} - u\frac{\exp(k_{21}t') - 1}{\exp(k_{21}t)} \quad (2)$$

with $t' \equiv t$ for $t \leq \tau$, $t' \equiv \tau$ for $t > \tau$ and

$$u^{-1} = k_{21}(k_{21} - k_{el}) \quad (3)$$
$$v^{-1} = k_{el}(k_{21} - k_{el}). \quad (4)$$

Here, k_{21} describes the exchange rate between the two compartments (blood plasma and interstitium) which is increased in tumors and k_{el} the first order elimination rate constant of the CM from the first compartment (plasma). The duration of the CM injection is described by τ. For convenience, all six model parameters are summarized as $\theta = (S_0, A, k_{el}, k_{21}, t_0, \tau)$ in the following.

Given the observed intensities y_i for a certain voxel at discrete time points $\{t_i\}_{i=1}^N$ the parameters are usually estimated by minimizing the sum of the squared residuals, i.e. by solving the NLLS problem

$$\hat{\theta} = \operatorname*{argmin}_{\theta} \sum_{i=1}^N (f_\theta(t_i) - y_i)^2 \quad (5)$$
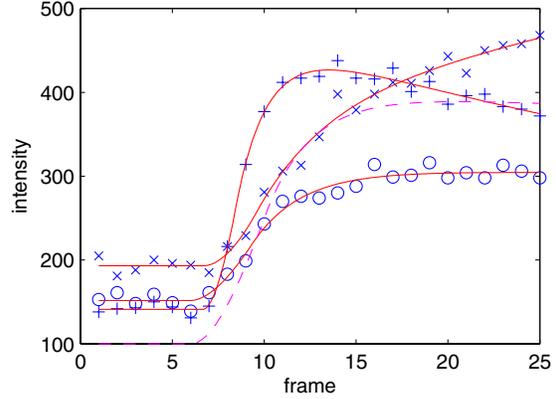

Figure 1. Three examples for the employed pharmacokinetic model from [5] fitted to recorded T1 intensity time curves. One frame was acquired every $11.25s$. The dashed time curve is obtained for $\theta_0 = (100, .3, .4, .003, 10, 6)$ used for initialization.

with $f_\theta(t) = S_{CM}(t)$. When normalized with the number of sampling points N the sum of squared residuals yields the mean square error (MSE) which is also minimum for $\hat{\theta}$. Some examples for time curves fitted with this model are given in Fig. 1.

For simplicity, regions of interest (ROIs) of the recorded MR volumes are usually analyzed voxel by voxel [5, 12]. Since the NLLS problem cannot be expected to be convex, multiple local optima may exist. Indeed, the NLLS fit with the previously described pharmacokinetic model can fail to find the "right" solution as shown later. The resulting parameter maps then contain single voxels or even regions of voxels for which the estimates are completely off. Instead of identifying and masking these voxels in the parameter map, we propose the application of a Bayesian approach which introduces a spatial smoothness prior in form of a generalized Gaussian Markov random field (GGMRF) [3]. In this way, the NLLS fit of a voxel is influenced by the data and the fits in its local neighborhood and pushed towards a better local optimum.

3. GGMRF: A Generalized Gaussian Markov Random Field Prior

The generalized Gaussian Markov random field [3] is a Markov random field [2, 14] with particular compatibility functions (the logarithm of which are known as potentials). Every voxel in the ROI is represented by a site $s \in S$ which is associated with the vector-valued random variable θ_s. Like in the single-voxel case, the observation likelihood is Gaussian, *i.e.* $y_i^s | \theta_s \sim \mathcal{N}(f_{\theta_s}(t_i), \sigma^2)$. Adding the spatial GGMRF prior on the parameter map θ yields a joint

distribution over y and θ in form of the Gibbs distribution:

$$\Pr(\theta, y) = \frac{1}{Z} \prod_{s \sim t} \Psi(\theta_s, \theta_t) \prod_s \Phi(\theta_s, y_s) \qquad (6)$$

where y and θ are vector variables obtained by stacking the site vector variables y_s and θ_s. Z is the global normalizer (partition function) and $s \sim t$ denotes pairs of neighboring sites according to the employed neighborhood system. The compatibility functions in $\Phi(\theta_s, y_s)$ and $\Psi(\theta_s, \theta_t)$ are defined by the potentials

$$\log \Phi(\theta_s, y_s) = -\frac{1}{2\sigma^2} \sum_{i=1}^N (f_{\theta_s}(t_i) - y_i^s)^2 \qquad (7)$$

$$\log \Psi(\theta_s, \theta_t) = -\frac{\alpha_{st}}{2} \|W(\theta_s - \theta_t)\|_p^p \qquad (8)$$

where $1 \le p \le 2$ and $\alpha_{st} \ge 0$ are hyper-parameters determining smoothness properties of the sought parameter map and W is a diagonal weighting matrix which accounts for the different scales of the parameters in θ_s.

The application of a GGMRF allows to vary continuously between a smoothing Gaussian MRF prior ($p = 2$) and an edge-preserving MRF ($p = 1$) with properties comparable to a weighted median filter [3]. Furthermore, the GGMRF potential defined by (8) is convex and, as opposed to robust alternatives such as the Huber potential [8], it does not have a threshold parameter at which its behavior suddenly changes.

In this paper, we will mostly be concerned with a special graph structure, namely regular lattices. Neighborhoods of different order can be defined on lattices [1]. Here, we will only consider neighborhoods up to an order of two (i.e. four and eight neighbors). Furthermore, we only model pairwise interactions, *i.e.* all compatibility functions for cliques with sizes greater than two are defined to be constant.

4. MAP Estimation with Block-ICM

Given an observed MR volume sequence $\{y_i^s\}$, the maximum a posteriori (MAP) estimate $\hat{\theta}$ is found by minimizing

$$\hat{\theta} = \arg\min_\theta \left[\sum_{s \in S} \sum_{i=1}^N (f_{\theta_s}(t_i) - y_i^s)^2 \right.$$
$$\left. + \sigma^2 \sum_{s \sim t} \alpha_{st} \|W(\theta_s - \theta_t)\|_p^p \right] \qquad (9)$$

For a realistically sized DCE-MR image of 100×100 voxels, the parameter vector θ already has $6 \cdot 10^4$ entries which makes the optimization problem (9) very hard to solve with standard NLLS algorithms. However, the problem is sparse in the sense that most of the θ_s are not directly coupled. The MRF framework provides special algorithms which can

exploit this sparsity such as the ICM (iterated conditional modes) algorithm [2].

Here, we use a generalized ICM algorithm which often converges faster than the standard ICM approach. As this algorithms considers collections of sites instead of single sites at each step, we call this approach *block-ICM*.

Given an arbitrary subset of sites $\tilde{S} \subseteq S$, it follows from the Hammersley-Clifford theorem [14] that the posterior distribution $\Pr(\theta \mid y) = \Pr(\theta_S \mid y)$ can be factored as

$$\Pr(\theta_S \mid y) = \Pr(\theta_{\tilde{S}} \mid \theta_{\partial \tilde{S}}, y) \Pr(\theta_{S \setminus \tilde{S}} \mid y) \qquad (10)$$

where $\partial \tilde{S} = \{s \mid t \sim s \wedge t \in \tilde{S} \wedge s \in S \setminus \tilde{S}\}$ is the border of $\tilde{S}$. Increasing $\Pr(\theta_{\tilde{S}} \mid \theta_{\partial \tilde{S}}, y)$ with respect to $\theta_{\tilde{S}}$ certainly cannot decrease $\Pr(\theta \mid y)$ since the second factor does not depend on any of the variables in $\theta_{\tilde{S}}$. Hence, the MAP problem (9) can be solved iteratively by solving a series of smaller MAP problems over subsets of sites

$$\theta_{\tilde{S}}^{(k+1)} = \arg\min_{\theta_{\tilde{S}}} \left[\sum_{s \in \tilde{S}} \sum_{i=1}^N (f_{\theta_s}(t_i) - y_i^s)^2 \right.$$
$$+ \sigma^2 \sum_{\substack{s \sim t \\ s,t \in \tilde{S}}} \alpha_{st} \|W(\theta_s - \theta_t)\|_p^p$$
$$\left. + \sigma^2 \sum_{\substack{s \sim t \\ s \in \tilde{S} \\ t \in \partial \tilde{S}}} \alpha_{st} \left\|W\left(\theta_s - \theta_t^{(k)}\right)\right\|_p^p \right] \qquad (11)$$

Clearly the block-ICM algorithm can be viewed as a coordinate descent approach where the potentially intersecting subsets $\tilde{S}^{(k)}$ redefine generalized coordinates $\theta_{\tilde{S}}^{(k)}$ in every descent step. Also, it is easily established that it suffices to find a configuration $\tilde{S}^{(k+1)}$ which decreases the objective (11) instead of finding the exact minimum in every descent step. The convergence to a local minimum of (9) is preserved.

The shape, size and update sequence of the subsets are design parameters of the block-ICM algorithm and should be chosen so as to trade off the problem size in each step against the number of sweeps required for convergence. If, *e.g.*, each of the subsets $\tilde{S}^{(k)}$ only contains one site s the standard ICM algorithm is recovered [2] which is known to converge often rather slowly. If, on the other hand, only one (sub)set $\tilde{S} \equiv S$ is chosen the complete MAP problem (9) which contains all variables is obtained. Hence, small subsets of sites should be chosen depending on the size of the local neighborhood and the strength of the mutual influence. Because of the locality of this influence, the size of the subsets does not have to be increased with growing lattices yielding and algorithm which scales linear with the number of sites.

5. Experimental Setup

Data from 36 patients with prostate cancer was provided by our clinical partner (German Cancer Research Center (dkfz), Heidelberg) [15]. DCE-MRI was performed on a 1.5-T clinical MR scanner (Magnetom Symphony; Siemens Medical Solutions, Erlangen, Germany) using endorectal coils (Medrad, Indianola, PA) with T1-weighted FLASH (TR/TE = 125ms/3.11ms) and a temporal resolution of 11.25s. Sequences of 25 image volumes have been acquired with a spatial resolution of $256 \times 160 \times 16$ voxels and $.78 \times .78 \times 3.99$ mm^3 voxel spacing. Regions of interest of about 100×100 voxels have been chosen in slices that contained tumorous tissue.

For comparison, the parameters of the pharmacokinetic model (1) have been determined using the voxel-wise approach first. The NLLS fit has been performed with an interior trust region method [6] as implemented in Matlab R14 (optimization toolbox). The same initialization $\theta_0 = (100, .3, .4, .003, 10, 6)$ has been used in all voxels (c.f. Fig. 1). A maximum number of 10000 iterations per voxel was allowed for in the to ensure convergence.

The proposed GGMRF approach has been tested for a first and second order neighborhood system (4N/8N) and homogeneous parameterization with $p = 2$. The influence of diagonal coupling terms in the second order neighborhood system has been reduced by choosing α_{st} for diagonal neighbors 1.4 times smaller than for horizontal/vertical neighbors.

The whole lattice was subdivided into two staggered sets of blocks with 6×6 voxels (Fig. 2). In every odd sweep, the blocks in the first set have been visited in a quincunx pattern as indicated in Fig. 2. In every even sweep, the same procedure was performed on the second set of blocks. A total of 12 sweeps have been performed. To prevent premature convergence to a local minimum, the number of optimization steps in each block was restricted to 20 iterations in the first 10 sweeps. For the remaining sweeps, up to 2000 iterations were permitted. The same optimization algorithm and the same initialization have been used for the GGMRF and the voxel-wise approach.

6. Results and Discussion

In Fig. 3, maps of the two parameters S_0 and k_{21} for patient P-19 are shown, contrasting the conventional voxel-wise approach and the spatial GGMRF approach with first and second order neighborhoods. Especially the k_{21}-map, which is an important physiological parameter [11], shows many white dots in the voxel-wise approach. At these voxels, the k_{21} estimate is very different from its surroundings which does not occur for the GGMRF approach. Furthermore, no visible difference is observed between the employed neighborhood systems. For the remainder, we will
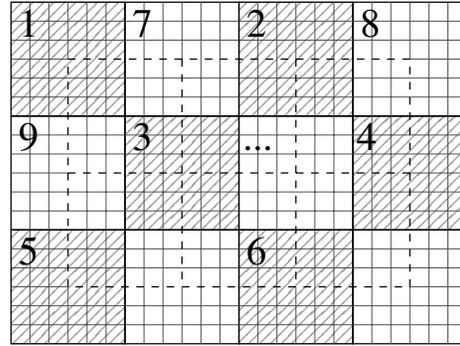

Figure 2. Blocks and update schedule used for the block-ICM algorithm. In every odd sweep square blocks of 6×6 sites are visited following the quincunx pattern as indicated by the numbering. The even sweeps are performed in the same way but shifted by 3 sites (dashed squares).

therefore confine ourselves to the GGMRF-8N on a second order lattice.

Figure 6 presents the difference in MSE between the GGMRF-8N and the voxel-wise approach. For voxels which are dark to black in Fig. 6(a), the GGMRF-8N yields smaller values. Hence, as judged by the MSE, the GGMRF-8N apparently offers a better fit in nearly all relevant voxels within the indicated boundary of the prostate. Fig. 4 shows model fits for some of those voxels for which the GGMRF yields significantly smaller MSE. Although especially the first two examples look like reasonable fits in the case of the voxel-wise approach, the solutions found with the GGMRF approach look much more convincing in all four cases.

At first sight the decreased MSE might be surprising since the use of a smoothness prior should result in estimation bias and always yield greater MSE than an unbiased estimate. However, this must only hold true if the global optimum of the NLLS objective 5 is actually found. And indeed, an increased MSE due to estimation bias can be observed in voxels in which the single-voxel approach converged to the global optimum and the spatial prior gains influence due to sharp edges in the parameter maps. For patient P-19 a few such voxels can be identified in the contrast-enhanced version of the difference image in Fig. 6(b).

That small degradations in MSE occur more often than small improvements by using the GGMRF prior can also be seen from the zoomed histogram in Fig. 6(c) which is slightly skewed to the right. It should be noted that to the far left of this histogram a few outliers are found, i.e. the black voxels in Fig. 6(a), resulting in a mean MSE difference which is clearly negative (c.f. Tab. 1).

For two of the voxels which are marked with circles in Fig. 6(b) the corresponding model fits are shown in Fig. 5. Although the GGMRF fits in the right column of Fig. 5 pro-

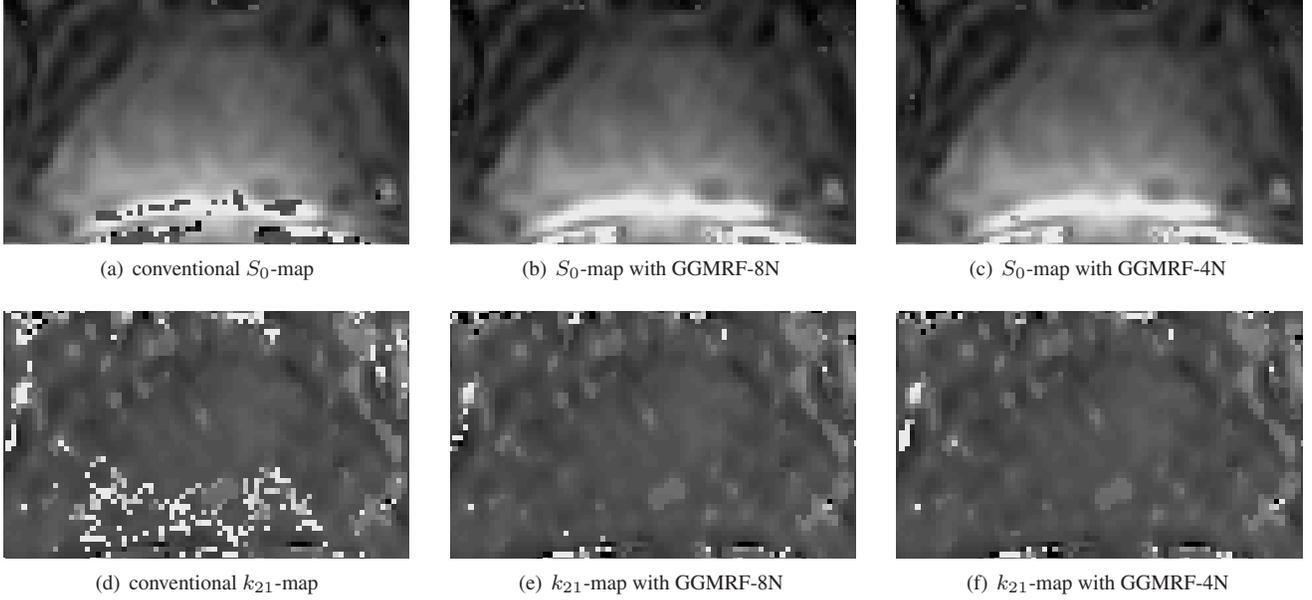|		
(a) conventional S_0-map	(b) S_0-map with GGMRF-8N	(c) S_0-map with GGMRF-4N
(d) conventional k_{21}-map	(e) k_{21}-map with GGMRF-8N	(f) k_{21}-map with GGMRF-4N

Figure 3. Comparison of parameter maps obtained for patient P-19 with conventional voxel-wise estimation and with GGMRF prior on a first and second order lattice (4N/8N). The k_{21}-map is reported to be most relevant for a diagnostic evaluation [11]. In both maps, the conventional approach produces speckles which do not occur if the spatial GGMRF prior is employed.

duce slightly greater MSE than the voxel-wise fits in the left column, both solutions look reasonable. One might even prefer the solutions found with the GGMRF approach since they are more in line with the expected pharmacokinetic behavior. It should be emphasized that this property is not explicitly influenced by the GGMRF approach but only implicitly encouraged by the surrounding voxels.

Table 1 lists numerical results for all 36 patients. For comparison we have calculated the mean MSE difference ΔMSE splitted into a negative and a positive part:

$$\Delta MSE^- = \frac{1}{|S|} \sum_{s \in S} \max(-\Delta MSE(s), 0) \quad (12)$$

$$\Delta MSE^+ = \frac{1}{|S|} \sum_{s \in S} \max(\Delta MSE(s), 0) \quad (13)$$

Then, the mean MSE difference can be calculated as $\Delta MSE = \Delta MSE^+ - \Delta MSE^-$. Hence, the GGMRF-8N approach performs better on the average if $\Delta MSE^- > \Delta MSE^+$ which is true for all but two of the 36 patients.

In Fig. 7 the $k21$-maps of several patients with various $\Delta MSE^-/\Delta MSE^+$ ratios are shown among which are the two most extreme patients P-4 and P-3 ($c.f.$ Tab. 1). Altogether, the examples shown in Figs. 3 and 7 seem to reflect quite well the numerical results in Tab. 1.

Table 1 also provides a robust estimate of the signal-to-noise ratio (SNR), $i.e.$ an estimate that is insensitive to outliers. The lowest SNR is obtained for patient P-7 for which two exemplary fits are shown in Fig. 8. Both cases confirm

that the measured image data is very noisy and show that the expected signal shape ($c.f.$ Fig. 1) can hardly be recognized. For example, in the left part of Fig. 1 it is difficult to judge which of the two very different fits should be favored.

For patient P-3 the estimated SNR is very high which indicates that the signal model fits very well. Also, both the negative and positive MSE differences (ΔMSE^- and ΔMSE^+) are very small which means that the GGMRF-8N and the voxel-wise approach perform very similar for that patient. Figures 7(i) and 7(j) show a comparison of the $k21$-maps. Apart from a few voxels, the two parameter maps look identical and the better mean MSE of the voxel-wise approach can again be ascribed to the bias introduced with the GGMRF. However, since no ground truth is available the bias-variance tradeoff cannot be analyzed on our experimental data. Nevertheless, oversmoothing seems to be no problem, neither for patient P-3 nor for the other examples. All structures present in the conventional parameter maps are preserved with the GGMRF approach.

The large gain in individual voxels cannot be ascribed to a reduction in estimation variance. Better parameter estimates for the dark voxels in Fig. 4 obviously do exist but are just not found with the voxel-wise approach. For example, the GGMRF parameter estimates in the right column of Fig. 4 all achieve lower scores on the NLLS objective (5) than the corresponding estimates obtained from the voxel-wise approach. Hence, the main benefit from using the GGMRF prior here is that it helps to avoid spurious local minima of (5) at individual sites.
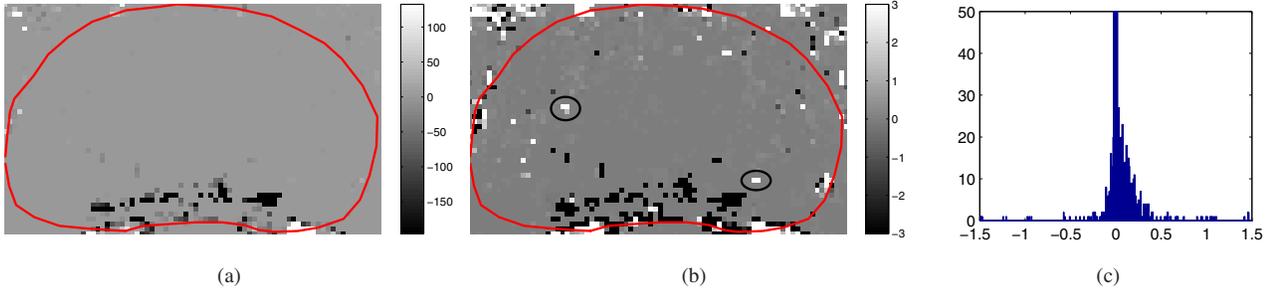
(a)　　　　　　　　　　　(b)　　　　　　　　　　　(c)

Figure 6. Difference in mean square error (MSE) for patient P-19. Darker pixels indicate sites for which the GGMRF model could find a better fit. (a) As for the parameter maps in Fig. 3, the greatest benefit is obtained in the lower half of the image. (b) Contrast-enhanced version of (a). (c) Zoomed histogram from the voxels within the prostate boundary. Positive differences prevail in the displayed range, reflecting the bias introduced by the spatial prior.
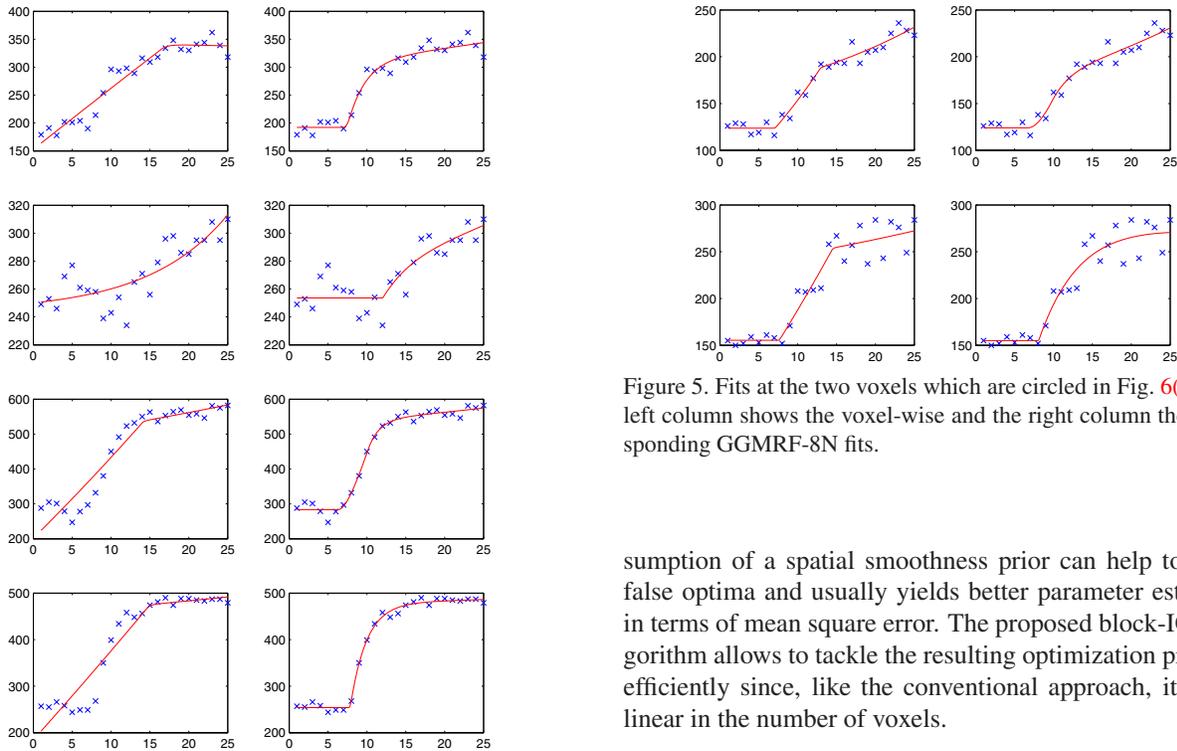

Figure 4. Comparison of selected model fits (patient P-19). In these examples the GGMRF-8N approach produced a lower MSE. The left column shows results from the voxel-wise approach whereas the right column shows the corresponding fit with the GGMRF-8N prior.

Figure 5. Fits at the two voxels which are circled in Fig. 6(b). The left column shows the voxel-wise and the right column the corresponding GGMRF-8N fits.

sumption of a spatial smoothness prior can help to avoid false optima and usually yields better parameter estimates in terms of mean square error. The proposed block-ICM algorithm allows to tackle the resulting optimization problem efficiently since, like the conventional approach, it scales linear in the number of voxels.

In the conducted experiments we could not observe considerable oversmoothing introduced by the spatial prior. On the other hand, the noticeable speckles resulting from failed parameter fits in the conventional approach could largely be removed.

7. Conclusion

We have proposed the application of a generalized Gaussian Markov random field prior for the quantification of DCE-MR image data. Since the nonlinear least squares problem used to fit a pharmacokinetic model in each voxel is not convex, the conventional approach is susceptible to spurious local optima. We have demonstrated that the as-

Acknowledgements

Patient	SNR [dB]	ΔMSE^-	ΔMSE^+	ratio
P-4	39.0	1410.241	2.186	644.981
P-30	33.8	1064.353	4.018	264.895
P-29	27.3	223.972	0.857	261.351
P-14	51.7	35.729	0.162	219.999
P-16	49.6	194.004	1.033	187.885
P-23	50.9	101.946	0.967	105.386
P-27	31.2	13.349	0.157	85.142
P-15	36.8	21.289	0.266	80.113
P-5	46.9	47.562	0.868	54.785
P-18	45.4	42.064	0.829	50.750
P-24	33.1	687.113	15.007	45.786
P-26	22.4	81.634	1.887	43.258
P-7	16.4	191.418	4.714	40.603
P-35	38.0	194.517	5.400	36.020
P-13	23.6	41.499	1.427	29.085
P-31	26.8	34.853	1.426	24.434
P-20	41.5	59.006	2.538	23.251
P-28	41.9	69.005	3.597	19.184
P-11	50.7	23.560	1.333	17.668
P-12	41.8	61.956	5.112	12.120
P-6	52.0	0.433	0.043	10.017
P-36	45.1	4.046	0.472	8.574
P-19	33.3	15.860	2.040	7.776
P-33	33.9	6.387	0.913	6.994
P-25	42.4	11.269	1.619	6.959
P-9	29.9	0.632	0.095	6.662
P-34	31.0	9.341	1.422	6.569
P-2	30.3	4.117	0.802	5.134
P-1	34.7	0.124	0.032	3.889
P-10	28.9	6.589	1.869	3.525
P-22	33.6	6.715	2.061	3.258
P-21	40.8	1.763	0.655	2.692
P-8	17.6	10.740	4.367	2.459
P-32	32.7	21.184	12.499	1.695
P-17	24.0	3.940	5.318	0.741
P-3	50.0	0.035	0.148	0.236

Table 1. Results for all 36 patients ordered by the ratio $\Delta MSE^-/\Delta MSE^+$. ΔMSE^- (ΔMSE^+) is the negative (positive) part of the mean MSE difference, *i.e.* only voxels in which the GGMRF-8N (voxel-wise) approach is better are considered. Hence, the mean MSE difference over all voxels is $\Delta MSE^+ - \Delta MSE^-$. The GGMRF-8N approach can improve the results in all but the last two cases (ratio > 1).

(a) conventional, patient P-4   (b) GGMRF-8N , patient P-4

(c) conventional, patient P-18   (d) GGMRF-8N , patient P-18

(e) conventional, patient P-28   (f) GGMRF-8N , patient P-28

(g) conventional, patient P-2   (h) GGMRF-8N , patient P-2

(i) conventional, patient P-3   (j) GGMRF-8N , patient P-3

Figure 7. Comparison of k_{21}-maps for several patients. The GGMRF-8N approach improves the MSE most for patients P-4, P-18, P-28 and P-2 and increases the MSE for patient P-3 (bold patients in Tab. 1). Nevertheless, the conventional k_{21}-map of patient P-3 in (c) shows some speckles which are avoided with the GGMRF-8N approach (d).

References

[1] N. Balram and J. M. F. Moura. Noncausal Gauss-Markov random fields: Parameter structure and estimation. *IEEE Trans Inf Theory*, 39(4):1333–1355, July 1993. 3

[2] J. E. Besag. Spatial interaction and the statistical analysis of lattice systems (with discussion). *J Roy Statist Soc Ser B*, 36:192–236, 1974. 2, 3
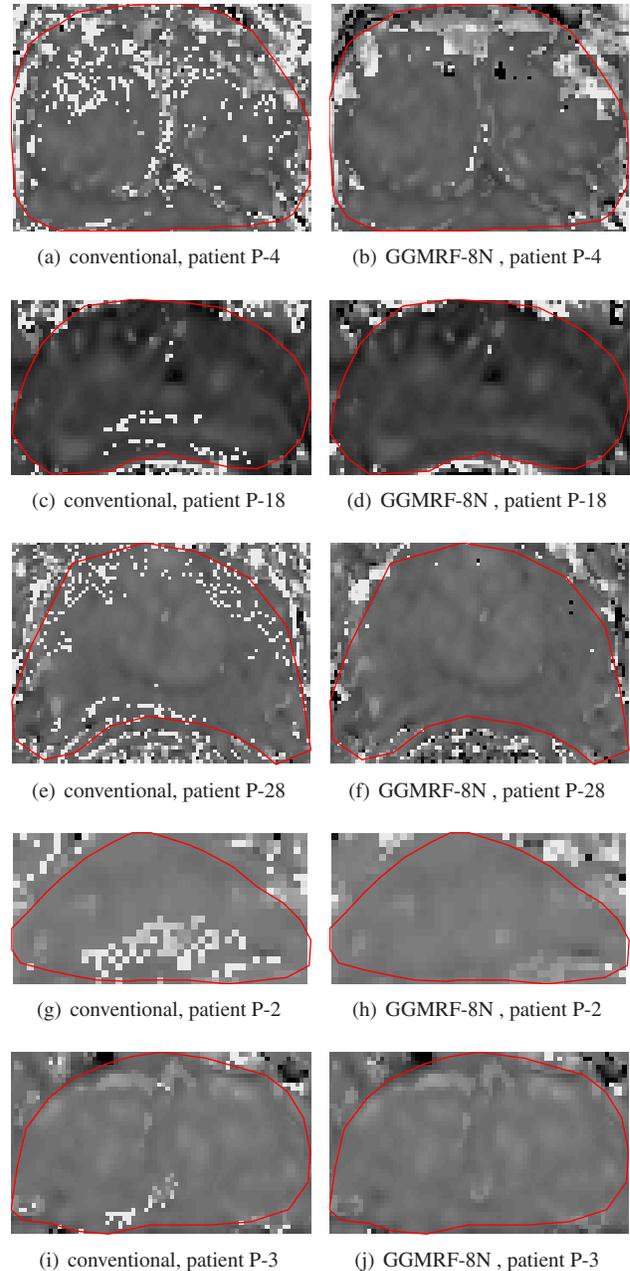
[3] C. Bouman and K. Sauer. A generalized Gaussian image model for edge-preserving MAP estimation. *IEEE Trans Image Proc*, 2(3):296–310, July 1993. 2, 3

[4] G. Brix, M. Henze, M. Knopp, R. Lucht, J. Doll, H. Junkermann, H. Hawighorst, and U. Haberkorn. Comparison of pharmacokinetic MRI and [18F] fluorodeoxyglucose PET in
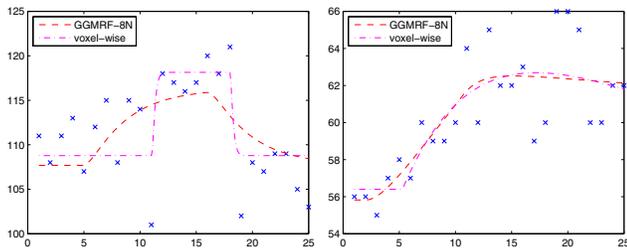
Figure 8. Two examples for fits from patient P-7 which has the lowest signal-to-noise ratio in Tab. 1. For both cases the expected signal shape is hardly recognizable in the measured data (blue crosses).

the diagnosis of breast cancer: initial experience. *Eur Radiol*, 11(10):2058–70, 2001. 1, 2

[5] G. Brix, W. Semmler, R. Port, L. Schad, G. Layer, and W. Lorenz. Pharmacokinetic parameters in CNS Gd-DTPA enhanced MR imaging. *J Comput Assist Tomogr*, 15(4):621–8, 1991. 1, 2

[6] T. Coleman and Y. Li. An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM J on Optim*, 6:418–445, 1996. 4

[7] H. Hawighorst, M. Libicher, M. Knopp, T. Moehler, G. Kauffmann, and G. Kaick. Evaluation of angiogenesis and perfusion of bone marrow lesions: role of semiquantitative and quantitative dynamic MRI. *J Magn Reson Imaging*, 10(3):286–94, Sep 1999. 1, 2

[8] P. J. Huber. *Robust Statistics*. New York: John Wiley and Sons Inc., 1981. 3

[9] F. Kiessling, M. Lichy, R. Grobholz, M. Heilmann, N. Farhan, M. S. Michel, L. Trojan, J. Ederle, U. Abel, H.-U. Kauczor, W. Semmler, and S. Delorme. Simple models improve the discrimination of prostate cancers from the peripheral gland by T1-weighted dynamic MRI. *Eur Radiol*, 14(10):1793–801, Oct 2004. 1, 2

[10] T. W. Nattkemper and A. Wismüller. Tumor feature visualization with unsupervised learning. *Med Image Anal*, 9(4):344–51, Aug 2005. 1

[11] H.-P. Schlemmer, J. Merkle, R. Grobholz, T. Jaeger, M. S. Michel, A. Werner, J. Rabe, and G. van Kaick. Can pre-operative contrast-enhanced dynamic MR imaging for prostate cancer predict microvessel density in prostatectomy specimens? *Eur Radiol*, 14(2):309–17, Feb 2004. 1, 2, 4, 5

[12] P. Tofts, G. Brix, D. Buckley, J. Evelhoch, E. Henderson, M. Knopp, H. Larsson, T. Lee, N. Mayr, G. Parker, R. Port, J. Taylor, and R. Weisskoff. Estimating kinetic parameters from dynamic contrast-enhanced T(1)-weighted MRI of a diffusable tracer: standardized quantities and symbols. *J Magn Reson Imag*, 10(3):223–32, Sep 1999. 2

[13] T. Twellmann, O. Lichte, and T. W. Nattkemper. An adaptive tissue characterization network for model-free visualization of dynamic contrast-enhanced magnetic resonance image data. *IEEE Trans Med Imaging*, 24(10):1256–66, Oct 2005. 1

[14] G. Winkler. *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*, volume 27 of *Applications of Mathematics*. Springer, 2nd edition, 2003. 2, 3

[15] C. Zechmann, K. Baudendistel, K. Aftab, L. Trojan, M.-S. Michel, H.-U. Kauczor, and S. Delorme. Dynamic contrast-enhanced T1-weighted MRI combined with MR spectroscopic imaging in patients with prostate cancer - initial experience. In *Proc Intl Soc Mag Reson Med*, volume 13, 2005. 1, 2, 4