Robust Markov Random Field Outlier Detection and Removal in Subsampled Images

ROBUST MARKOV RANDOM FIELD OUTLIER DETECTION AND REMOVAL IN SUBSAMPLED IMAGES

Paul McCool[(1)], *Yoann Altmann*[(1)], *Antonios Perperidis*[(1)(2)], *Steve McLaughlin*[(1)*]

[(1)]Heriot-Watt University,School of Engineering and Physical Sciences, Edinburgh, UK.
{*P.McCool, Y.Altmann, A.Perperidis, S.McLaughlin* }*@hw.ac.uk*
[(2)]IRC Proteus Hub, Queens Medical Research Institute, University of Edinburgh, UK.

ABSTRACT

Certain imaging technologies, such as fibred optical microscopy, operate with irregularly-spaced sparse sub-samples from their field of view. In this work, we address the problem of data restoration for applications where the observed irregularly distributed samples are corrupted by additive observation noise and sparse outliers (such as broken and damaged fibre cores). This problem is formulated as joint outlier detection and de-noising of irregularly sampled data and is addressed within a classical Bayesian framework. A Markov Random Field is considered to capture the intrinsic spatial correlation of the underlying intensity field and binary labels are used to locate the spatial position of the outliers. A Markov Chain Monte Carlo method is then used to perform Bayesian inference using the posterior distribution associated with the resulting Bayesian model. Simulations conducted on simulated data show the potential benefits of the proposed method in terms of data restoration and outlier identification.

Index Terms— Irregular sampling, Bayesian estimation, Markov chain Monte Carlo methods, anomaly detection.

1. INTRODUCTION

Fibred confocal imaging systems, such as Cellvizio, generally generate a rectangular image of a bundle of irregularly-spaced optical fibre cores [1]. While each individual optical fibre in the bundle spreads across several pixels, it provides only a single piece of intensity information about the field of view (FOV). Moreover, the cladding between the individual fibres further reduces amount of information acquired by the imaging system. This arrangement can be thought of as the FOV being sparsely sampled such that only a few percent of the pixels have intensity information. In this paper and without loss of generality, a single pixel (the closest to the centre of each fibre) is chosen to represent the value of the intensity field at the fibre location. Due to the observation conditions the measured intensity are corrupted by noise (assumed to be additive) that needs to be mitigated. Moreover, outliers or significant local intensity variations can arise, caused either by broken or faulty fibres, or by the presence of bright and localised materials (e.g., stained bacteria). A specific methodology is thus required to detect and quantify such anomalies, which is precisely the aim of this work.

Although the reconstruction of full images (i.e., full intensity fields) is on prime interest for visual assessment of internal tissues (e.g., lung alveoli), in this work we mainly focus on restoring (denoising and outlier removal) intensity values only at the fibre locations for subsequent post-processing (not addressed in this paper). First, from a computational point of view, the observable intensity information can be summarized using only the reduced number of observed samples and the full image (usually reconstructed using an arbitrary interpolation scheme) simply embeds additional/prior information that can be directly used during the post-processing steps. Thus, restoring only a reduced number of parameters allows a more efficient data storage and post-processing. Second, for very sparse sampling patterns, state-of-the art interpolation methods,e.g., based on local smoothness assumptions (e.g., total-variation regularization), via Fourier/wavelet transforms or Gaussian Process regression [2] provide interpolated images that are similar to those obtained using standard linear interpolation. Thus, we propose to mainly focus on the restoration of the observed data, given that a satisfactory full image can be obtained via interpolation after the denoising/outlier removal step [1].

This work significantly differs from compressive sensing approaches which rely on sparse sampling, assuming that the signal of interest (e.g., the image) has a sparse representation in an arbitrary basis[3]. Although compressive sensing approaches have been refined to reduce noise [4] and the effects of outliers [5][6], such methods do not apply here as we are not primarily interested in recovering the full image, but on detecting and quantifying the potential anomalies.

In [7] a robust blind source separation technique was proposed to analyse hyperspectral data. The positions of the outliers within the data cube were modelled *via* a structured binary mask. This approach has been extended in [8], where an Ising model was used as prior model for the outlier positions. In a similar manner, here the observations are modeled as the actual intensity values, corrupted by Gaussian noise and sparse outliers. We propose a Bayesian method for joint outlier detection and de-noising, while accounting for the spatial correlation affecting the underlying intensity field of interest. Adopting a classical Bayesian framework, we assign prior models to the unknown model parameters. More precisely, a Gaussian Markov random field (GMRF) is used to model the underlying light intensity field (within the FOV) and binary labels are introduced to identify the outlier positions. Classical conjugate prior models are used for the other unknown parameters, including the variance of the additive Gaussian noise and the parameter controlling the amount of spatial correlation of the light field. The originality of work of this work in terms of modeling is the combination of the ℓ_0 based outlier model with a GMRF adapted for irregularly sampled data. The joint posterior distribution of these parameters is then derived and a Markov chain Monte Carlo (MCMC) method is used to generate samples according to the posterior of interest. These samples are then used to assess the outlier presence (outlier values and confidence measures) and denoise the intensity values.

*Part of this work was supported by the EPSRC via grant EP/K03197X/1.

The remainder of the paper is organized as follows. Section 2 introduces the observation model associated with the joint denoising and outlier detection problem considered. Section 3 presents the proposed hierarchical Bayesian model modelling the dependence between the different model parameters, and the resulting posterior distribution. Section 4 describes the Gibbs sampler that is used to sample from the posterior of interest and subsequently approximate appropriate Bayesian estimators. Results of experiments conducted on controlled data are shown and discussed in Section 5 and conclusions are reported in Section 6.

2. PROBLEM FORMULATION

Consider a set of N observed samples $\mathbf{y} = [y_1, \ldots, y_N]^T$, irregularly distributed in a two-dimensional (image) plane. These samples are assumed to be corrupted by additive Gaussian noise and potential additive outliers. In a similar manner to [8, 9, 7], the vector $\mathbf{y}$ can be modelled as

$$\mathbf{y} = \mathbf{x} + \mathbf{r} + \mathbf{e} \tag{1}$$

with $\mathbf{x} = [x_1, \ldots, x_N]^T$, $\mathbf{r} = [r_1, \ldots, r_N]^T$ and $\mathbf{e} = [e_1, \ldots, e_N]^T$, where x_n is the actual intensity value for the nth sample, r_n are the outliers and e_n is noise. The problem addressed in this paper consists of recovering $\mathbf{x}$ and the outliers in $\mathbf{r}$ from the observed samples in $\mathbf{y}$. As discussed above, once $\mathbf{x}$ has been estimated through the MCMC process described in the rest of this paper, a full (rectangular) enhanced image can be obtained via interpolation from $\mathbf{x}$. The next Section details the proposed Bayesian model used to solve the denoising/outlier detection problem considered.

3. BAYESIAN MODEL

3.1. Likelihood

Eq. (1) shows that $y_n|x_n, r_n, \sigma^2 \sim \mathcal{N}\left(y_n; x_n + r_n, \sigma^2\right)$. Assuming independence between noise sequences of the N observed samples, the likelihood of the observation vector $\mathbf{y}$ can be expressed as

$$f(\mathbf{y}|\mathbf{x}, \mathbf{r}, \sigma^2) \propto \sigma^{-N} \exp\left[-\frac{\|\mathbf{y} - \mathbf{x} - \mathbf{r}\|^2}{2\sigma^2}\right] \tag{2}$$

where $\propto$ means "proportional to" and $\|\cdot\|$ is the standard ℓ_2 norm.

3.2. Parameter priors

3.2.1. Prior for the underlying intensity field $\mathbf{x}$

For many applications, the intensity values of the scene to be recovered are likely to be spatially correlated. An interesting way to take possibly correlated intensities is to consider MRFs to build a prior for $\mathbf{x}$. MRFs assume that the distribution of a given intensity x_n conditionally to the other intensity values of the image equals the distribution of this parameter conditionally to its spatial neighbours, i.e., $f(x_n|\mathbf{x}_{\setminus x_n}) = f(x_n|\mathbf{x}_{\mathcal{V}_n})$, where $\mathcal{V}_n$ is the index set of the neighbours of x_n, $\mathbf{x}_{\setminus x_n}$ denotes the vector $\mathbf{x}$ whose element x_n has been removed and $\mathbf{x}_{\mathcal{V}_n}$ is the subset of $\mathbf{x}$ composed of the elements whose indexes belong to $\mathcal{V}_n$. In this paper, a Delaunay triangulation scheme is used on the N samples to define the neighouhood structure. In this study, we specify $f(x_n|\mathbf{x}_{\mathcal{V}_n})$ as

$$f(x_n|\mathbf{x}_{\mathcal{V}_n}, \gamma^2) \propto \exp\left[-\frac{1}{\gamma^2} \sum_{n' \in \mathcal{V}_n} \frac{(x_n - x_{n'})^2}{d_{n,n'}}\right], \tag{3}$$

where $d_{n,n'}$ denotes the distance between the spatial locations n and n' and γ^2 controls the global correlation between intensities. Eq. (3) promotes smooth intensity variations between neighbours while ensuring that the prior dependence between neighbours decreases as $d_{n,n'}$ increases. In this work $d_{n,n'}$ is the standard euclidean distance, but other distances could be used (e.g., distances on manifolds). It can be shown that the resulting joint prior $f(\mathbf{x}|\gamma^2)$ can be expressed as

$$f(\mathbf{x}|\gamma^2) \quad \propto \quad \gamma^{-d/2} \exp\left[-\frac{\mathbf{x}^T \mathbf{D} \mathbf{x}}{2\gamma^2}\right] \tag{4}$$

where

$$[\mathbf{D}]_{i,j} = \begin{cases} \sum_{j \in \mathcal{V}_i} 1/d_{i,j} & \text{if } i = j \\ 0 & \text{if} j \notin \mathcal{V}_i \\ 1/d_{i,j} & \text{else} \end{cases} \tag{5}$$

and $d = \text{rank}(\mathbf{D})$. It is important to mention that when $d < N$, the prior (4) is not proper. However, as will be seen is Section 4, the conditional distribution of $\mathbf{x}$ is proper (when combining (4) and 2). Note that if a regular sampling grid is used (4) can be replaced by a (degenerated) Gaussian prior based on a Laplacian filter as in [10], leading to a posterior which is easier to sample.

3.2.2. Prior for the noise variance

A Jeffreys' prior is chosen for the noise variance σ^2, i.e., $f(\sigma^2) \propto \sigma^{-2} \mathbf{1}_{\mathbb{R}^+}\left(\sigma^2\right)$ where $\mathbf{1}_{\mathbb{R}^+}(\cdot)$ denotes the indicator function defined on $\mathbb{R}^+$, which reflects the absence of knowledge about this parameter.

3.2.3. Priors of the outliers

As in [9, 7, 11], the outliers are assumed to be sparse, i.e., for most of the spatial locations, the outliers are expected to be exactly equal to zero. To model the outlier sparsity, we factorize the outlier vector as $\mathbf{r} = \mathbf{z} \odot \mathbf{t}$, where $\mathbf{z} = [z_1, \ldots, z_N]^T \in \{0, 1\}^N$ is a label vector, $\mathbf{t} \in \mathbb{R}^N$ and $\odot$ denotes the Hadamard (termwise) product. This decomposition allows one to decouple the location of the sparse components from their values. More precisely, $z_n = 1$ if an outlier is present in the nth observed location with value equal to $r_n = t_n$. A conjugate Gaussian prior is used for $\mathbf{t}$, i.e.,

$$f(\mathbf{t}|s^2) = \prod_n \mathcal{N}\left(t_n; \mu, s^2\right), \tag{6}$$

where μ and s^2 control the prior mean and variance of the outliers. Note that (6) allows the outliers to be negative. Other conjugate priors, such as truncated Gaussian priors, could be used instead of (6), e.g., to enforce outlier positivity. The next section presents the prior considered for the label vector $\mathbf{z}$.

3.2.4. Label matrix

In this work, we do not assume any particular spatial structure for the outlier positions. However, the anomalies have the same probability to be present in each region of the scene. To model this prior belief, we assign each label the following Bernoulli prior $f(z_n = 1|\pi) = 1 - f(z_n = 0|\pi) = \pi$. Moreover, we assume that the probability of outlier presence π is also unknown and we include this parameter within the inference process by assigning π a conjugate Beta prior $\mathcal{B}e(\pi; \alpha, \beta)$. In this paper, we fix (α, β) to $(\alpha, \beta) = (0.1, 1)$ as we expect the proportion of outliers to be relatively small (the prior

mean of π is $\alpha/(\alpha+\beta)$). Note that to encode the potentially structured spatial distribution of the outliers, a Markov model could also be used (e.g., an Ising model as in [11]).

3.3. Hyperparameters

To reflect the lack of prior knowledge about the outlier variance in (6) and regularization parameter γ^2 in (4), the following weakly informative inverse-Gamma priors are assigned to s^2 and γ^2

$$s^2 \sim \mathcal{IG}(\eta,\nu), \quad \gamma^2 \sim \mathcal{IG}(\eta,\nu),$$

where (η,ν) are fixed to $(\eta,\nu)=(10^{-3},10^{-3})$ to obtain a weakly informative prior. Similarly, we assign μ the following conjugate Gaussian prior

$$\mu|\bar{\mu},\xi^2 \sim \mathcal{N}\left(\mu;\bar{\mu},\xi^2\right) \tag{7}$$

where $(\bar{\mu},\xi^2)$ are fixed and user-defined parameters (which might depend on the dynamics of the image to be recovered). In this work, we arbitrarily fixed $(\bar{\mu},\xi^2)=(200,10000)$ but these parameters can be changed depending on the prior knowledge available about the outliers. The next section derives the joint posterior distribution of the unknown parameters associated with the proposed Bayesian model and studies an MCMC methods to sample from this posterior. These samples can then be used to approximate appropriate Bayesian estimators.

4. SAMPLING STRATEGY

4.1. Joint posterior distribution

Assuming the parameters $\mathbf{x},\mathbf{z},\mathbf{t}$ and σ^2 are a priori independent, the joint posterior of the parameter vector $\boldsymbol{\theta}=\left\{\mathbf{x},\mathbf{z},\mathbf{t},\sigma^2\right\}$ and hyperparameters $\boldsymbol{\Phi}=\{\mu,s^2,\pi,\gamma^2,\}$ can be expressed as

$$f(\boldsymbol{\theta},\boldsymbol{\Phi}|\mathbf{y}) \propto f(\mathbf{y}|\boldsymbol{\theta})f(\boldsymbol{\theta}|\boldsymbol{\Phi})f(\boldsymbol{\Phi}) \tag{8}$$

where

$$\begin{aligned} f(\boldsymbol{\theta}|\boldsymbol{\Phi}) &= f(\mathbf{x}|\gamma^2)f(\mathbf{z}|\pi)f(\mathbf{t}|\mu,s^2)f(\boldsymbol{\sigma}^2), \\ f(\boldsymbol{\Phi}) &= f(\gamma^2)f(\pi)f(\mu)f(s^2), \end{aligned} \tag{9}$$

and $f(\mathbf{z}|\pi)=\prod_n f(z_n|\pi)$.

4.2. Gibbs sampler

To overcome the challenging derivation of Bayesian estimators associated with $f(\boldsymbol{\theta},\boldsymbol{\Phi}|\mathbf{y})$, we propose to use an efficient Markov Chain Monte Carlo (MCMC) method to generate samples asymptotically distributed according to (8). More precisely, we consider a Gibbs sampler described in the next part of this section. The principle of the Gibbs sampler is to sample according to the conditional distributions of the posterior of interest [12, Chap. 10]. In this paper, we propose to sample sequentially the elements of $\boldsymbol{\theta}$ and $\boldsymbol{\Phi}$ using moves that are summarized below.

Labels: Sampling each label z_n from its conditional distribution can be achieved by drawing in $\{0,1\}$ with known probabilities. Moreover, the elements of $\mathbf{z}$ can be updated in a parallel manner using $f(\mathbf{z}|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\Phi})=\prod_n f(z_n|\mathbf{y},\boldsymbol{\theta},\boldsymbol{\Phi})$.

Intensity field: Sampling each intensity value x_n from its conditional distribution reduces to sampling from a Gaussian distribution (using the conjugacy of (2) and (3)). Note that it is possible to exploit the structure of the MRF to cluster sub-samples into groups of subsamples that are not neighbours, and thus update several intensities in a parallel manner by updating the groups sequentially.

Outlier values $\mathbf{t}$: In a similar manner to the intensity field, it can be shown that the conditional distribution of $\mathbf{t}$ is a multivariate Gaussian distribution with diagonal covariance matrix, from which is easy to sample (the elements of $\mathbf{t}$ can be sampled independently and in a parallel manner).

Noise and outlier variances and MRF parameter: Sampling σ^2,s^2 and γ^2 can be achieved by sampling from independent inverse-Gamma distributions.

Outlier mean: Using the conjugacy of (6) and (7), updating μ reduces to sampling from a standard Gaussian distribution.

Probability of outlier presence: Due to the conjugacy of the hierarchical prior model $f(\mathbf{z}|\pi)f(\pi)$, the conditional distribution of π reduces to a standard Beta distribution.

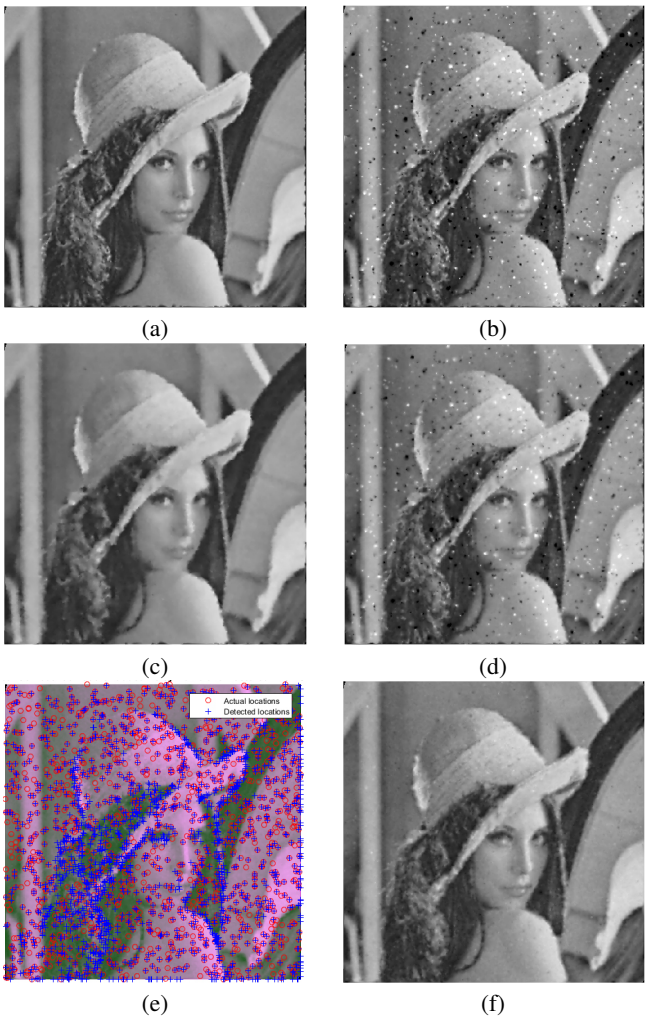
After N_{MC} iterations (including N_{bi} burn-in samples that are discarded), the label vector is estimated using marginal maximum a posteriori (MAP) estimation. This estimator is then used to compute the minimum mean square error (MMSE) of $\mathbf{r}$ conditioned upon $\mathbf{z}=\widehat{\mathbf{z}}_{\text{MAP}}$, i.e., $\widehat{\mathbf{r}}=(\widehat{\mathbf{r}}_{\text{MMSE}}|\widehat{\mathbf{z}}_{\text{MAP}})\odot\widehat{\mathbf{z}}_{\text{MAP}}$. Finally, the remaining parameters are estimated using the empirical averages of the generated samples (MMSE estimates). It is interesting to note that thanks to the conjugate Gaussian prior (6), the vector $\mathbf{t}$ could have been marginalized. However, this marginalization would lead to non-standard conditional distributions for σ^2 and s^2 and accept/reject procedures would have to be used to update these variables. Instead, we propose not to marginalize $\mathbf{t}$ and estimate the outlier values, as it would be achieved when considering more complex prior than (6) (e.g., to handle outlier positivity).

In this work, we assume that the rank of the matrix $\mathbf{D}$ is known (it only needs to be computed once). However its computation can be challenging for large data sets, i.e., for large values of N. In such cases, it is still possible to estimate γ^2 without significant additional computational cost, e.g., using maximum marginal likelihood estimation [13], in a similar manner to [8, 14].

5. EXPERIMENTS

In order to assess the effectiveness of this approach for detecting outliers, Gaussian noise was added to the standard 'Lena' reference image. Then approximately 8% of the pixels were sampled to form the sub-set, the coordinates were noted and neighbours were calculated using Delaunay Triangulation. Artificial outliers, drawn from a normal distribution (with varying mean and variance), were then added to a varying fraction of the sub-set. Figure 1 shows an example of interpolation from sub-samples, corruption by noise ($\sigma^2=10$) and 5% of outliers $(\mu,s^2)=(255,10000)$, outlier detection and processing.

Figure 1(a) shows an image formed by natural neighbour interpolation from around 30,000 randomly sub-sampled pixels of the 262,144 pixels in the Lena test image. Figure 1(b) shows an interpolated image where 5% of the samples are corrupted with positive and negative outliers drawn from a Gaussian distribution. The visual impact of the outliers is exacerbated by the interpolation process. Figure 1(c) shows an image that has been interpolated from the noisy subsamples that have been median filtered with their subsample neighbours. Figure 1(d) shows the result of MCMC processing without outlier detection. Figure 1(e) shows a graphical representation of the true outlier locations and the detected outliers. This figure illustrates the outlier detection ability of the proposed method which detects most of the actual outliers (those not detected

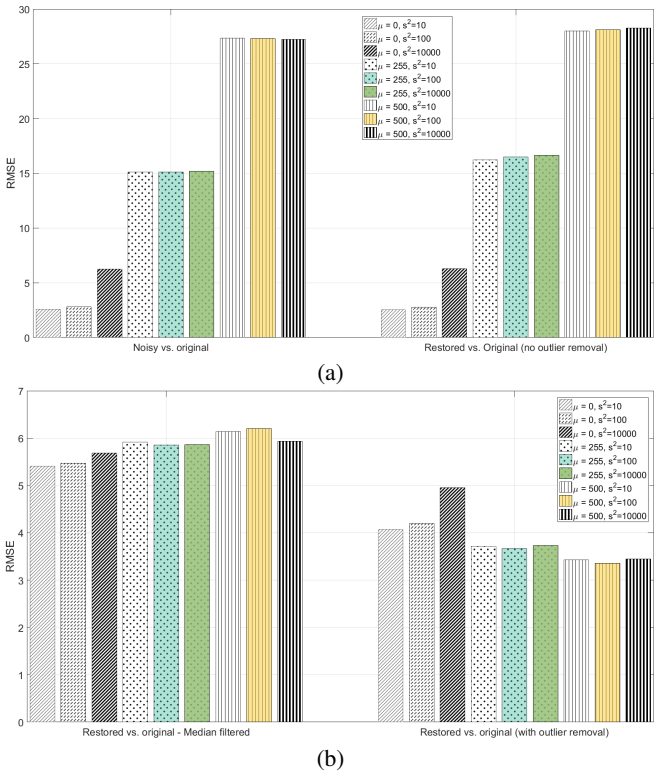


(a) (b) (c) (d) (e) (f)

Fig. 1. (a) Image interpolated from roughly 11% of sub-samples of Lena image, (b) Image interpolated where sub-samples have been corrupted by Gaussian noise and 5% have been corrupted by outliers, (c) Noisy image filtered by median filter, (d) Processed using MCMC without the outlier detection mechanism, (e) Graphical representation of outlier detection, (f) Interpolated from processed sub-sample values

are generally of low amplitude). Note that the the proposed methods also detects sharp intensity changes. Figure 1(f) shows an interpolated image after the sub-samples have been processed using MCMC. We evaluated our algorithm by selecting $\sigma^2 \in \{0, 1, 10\}$, $\mu \in \{0, 255, 500\}$and $s^2 \in \{10, 100, 10000\}$.

The criterion used to compare the original, noisy and recovered samples is the Root Mean Square Error (RMSE), defined as $RMSE = \frac{1}{N}\sqrt{\sum_{n=1}^{N}(x_n - \hat{x}_n)^2}$, where x_n is the nth of the original sample, and $\hat{x}_n$ is the nth recovered sample. The results from a representative sub-set of parameter combinations are shown in figure 2 for 5% outliers. Outlier percentages of 1% and 10% were also tested and are not shown here due to space constraints, but the trends in behaviour are the same. Median filtering can be used to mitigate, but not explicitly detect, outliers. Note that each sample is median-filtered using its direct neighbours. The MCMC algorithm is compared against median filtering as shown in figure 2.



(a)

(b)

Fig. 2. RMSE for $\sigma^2 = 10$ 5% of outliers

Figure 2 shows that MCMC has lower RMSE when outlier detection is implemented. If the outlier mean and covariance are similar to those of the sub-samples with added Gaussian noise, then the RMSE is higher because more edge points than outliers are removed. Additionally, spatial regularisation introduces estimation bias, which might locally increase the number of detected outliers (around sharp edges).

6. CONCLUSION

In this paper, an outlier detection method based on Markov Chain Monte Carlo has been shown to 1) satisfactorily detect anomalies in the data and 2) improve the reconstruction of RMSE of non-uniform image sub-samples of test images that have been corrupted by outliers. The improvement in terms of reconstruction has been shown to be better than what is achievable using a median filter. One of the main advantages of the proposed approach is that it does not rely on strong assumptions for the sampling pattern. Future work will seek to extend the method to structured outliers and color/multispectral images and apply it to real optical endomicroscopy images. Although very interesting, algorithmic improvement in terms of computational complexity (e.g., using optimization methods) is out of scope of this paper and will also be considered for future work.

7. ACKNOWLEDGEMENT

We would like to thank Engineering and Physical Sciences Research Council (EPSRC, United Kingdom) Interdisciplinary Research Collaboration grant EP/K03197X/1 for funding this work.

8. REFERENCES

[1] N. Ayache, T. Vercauteren, G. Malandain, F. Oberrietter, N. Savoire, and A. Perchant, "Processing and mosaicing of fibered confocal images," *MICCAI Workshop on Microscopic Image Analysis with Applications in Biology 2006*, 2006.

[2] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.

[3] E. Candes, N. Braun, and M. Wakin, "Sparse signal and image recovery from compressive samples," in *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*, April 2007, pp. 976–979.

[4] D.-H. Lee, C.-P. Hong, M.-W. Lee, H.-J. Kim, J.-H. Jung, W.-H. Shin, J.-G. Kang, S.-J. Kang, and B.-S. Han, "Sparse sampling mr image reconstruction using bregman iteration: A feasibility study at low tesla mri system," in *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE*, Oct 2011, pp. 3446–3449.

[5] B. Moore and B. Natarajan, "A general framework for robust compressive sensing based nonlinear regression," in *Sensor Array and Multichannel Signal Processing Workshop (SAM), 2012 IEEE 7th*, June 2012, pp. 225–228.

[6] A. Ramirez, R. Carrillo, G. Arce, K. Barner, and B. Sadler, "An overview of robust compressive sensing of sparse signals in impulsive noise," in *Signal Processing Conference (EUSIPCO), 2015 23rd European*, Aug 2015, pp. 2859–2863.

[7] G. E. Newstadt, A. H. III, and J. Simmons, "Robust spectral unmixing for anomaly detection," in *Proc. IEEE-SP Workshop Stat. and Signal Processing*, Gold Coast, Australia, July 2014.

[8] Y. Altmann, S. McLaughlin, and A. Hero, "Robust linear spectral unmixing using anomaly detection," *IEEECOMP*, vol. 1, no. 2, pp. 74–85, Jun 2015.

[9] X. Ding, L. He, and L. Carin, "Bayesian robust principal component analysis," *IEEE Trans. Image Processing*, vol. 20, no. 12, pp. 3419–3430, Dec 2011.

[10] C. Gilavert, S. Moussaoui, and J. Idier, "Efficient gaussian sampling for solving large-scale inverse problems using mcmc," *IEEE Trans. Signal Process.*, vol. 63, no. 1, pp. 70–80, Jan. 2015.

[11] Y. Altmann, N. Dobigeon, S. McLaughlin, and J.-Y. Tourneret, "Residual component analysis of hyperspectral images - application to joint nonlinear unmixing and nonlinearity detection," *IEEE Trans. Image Processing*, vol. 23, no. 5, pp. 2148–2158, May 2014.

[12] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. New York: Springer-Verlag, 2004.

[13] M. Pereyra, N. Whiteley, C. Andrieu, and J.-Y. Tourneret, "Maximum marginal likelihood estimation of the granularity coefficient of a Potts-Markov random field within an mcmc algorithm," in *Proc. IEEE-SP Workshop Stat. and Signal Processing*, Gold Coast, Australia, July 2014.

[14] Y. Altmann, M. , and J. Bioucas-Dias, "Collaborative sparse regression using spatially correlated supports - application to hyperspectral unmixing," *IEEE Trans. Image Processing*, vol. 24, no. 12, pp. 5800–5811, Dec. 2015.