JOINT RECOVERY AND SEGMENTATION OF POLARIMETRIC IMAGES USING A COMPOUND MRF AND MIXTURE MODELING

G. Sfikas[1,2], *C. Heinrich*[2], *J. Zallat*[2], *C. Nikou*[1], *N. Galatsanos*[3]

[1]Department of Computer Science, University of Ioannina, Greece,
[2] LSIIT (UMR 7005 CNRS–UDS), University of Strasbourg, France,
[3]Department of ECE, University of Patras, Greece.

ABSTRACT

We propose a new approach for the restoration of polarimetric Stokes images, capable of simultaneously segmenting and restoring the images. In order to easily handle the admissibility constraints inherent to Stokes images, a proper transformation of the images is introduced. This transformation exploits the correspondence between any Stokes vector and the covariance matrix of the two components of the electric vector of the light wave. A Bayesian model based on a mixture of Gaussian kernels is used for the transformed images. Inference is achieved using the EM framework.

To quantify the performances of this approach, the algorithm is tested with both synthetic and real data. We note that the pixels of the restored Stokes images issued from our approach are always physically admissible which is not the case for the naïve pseudo-inverse approach.

Index Terms— Polarimetric images, image segmentation, spatially varying Gaussian mixture models, Expectation-Maximization (EM) algorithm, Markov Random field (MRF).

1. INTRODUCTION

Exploiting the polarization of light has been shown to be a useful and powerful technique, overcoming many limitations that arise in radiance measurement-based classical imagery. Indeed, there is increasing evidence that recording the polarization properties of inhomogeneous objects provides a rich set of information about their local nature. This imaging modality requires the development of efficient imaging systems that can record spatially distributed polarization patterns across a scene and appropriate techniques of handling and processing the issued multicomponent images while preserving the physical integrity of the data.

Stokes imaging consists in estimating the four Stokes parameters of each pixel of an image. This is traditionally achieved by placing a Polarization State Analyzer (PSA) in front of a camera. This configuration allows acquiring polarized radiance images g that can be used to calculate the multi-component Stokes image s. Images g and s are linked by the Polarization Measurement Matrix (matrix H) that depends on the PSA configuration ($g = Hs$, pixelwise). Classically, the Stokes parameters are obtained using a pseudo-inverse approach, which is sensitive to noises that degrade the acquired intensity images. Moreover, at each location, the Stokes parameters

---

must satisfy physical constraints that can be infringed when using pseudo-inversion.

In this paper we present a Bayesian model based on a mixture of Gaussians, capable of performing recovery simultaneously with segmentation of the polarimetric image. Intuitively we conjecture that the problem of recovering the degraded image should be intertwined with the problem of segmenting the image. In other words, a good restoration should lead to a good segmentation and vice versa, which is the motivation of performing the two operations jointly.

Polarimetric imaging gives rise to intricate estimation problems because of the associated underlying physical admissibility conditions [1, 2]. Stokes images exhibit the particularity that, while they are comprised of 4 separate channels, only a subset of $\mathbb{R}^4$ constitutes admissible Stokes 4-variate vectors. In order to work around this problem, instead of assuming directly that the input Stokes vectors follow a certain distribution, we impose a mixture of Gaussian distributions on a suitable transformation of the Stokes image. Furthermore, with a properly chosen prior set on the probabilities of the underlying segmentation class labels, we achieve to produce smooth edge-preserving segmentations that in turn produce correspondingly reasonably smooth image restorations.

Polarimetric image restoration methodologies have already been presented in [1, 2], but under the hypothesis that the variables of interest were spatially piecewise constant. Our present model uses no such constraining assumption. We have tested the proposed model succesfully on sets of noise-degraded artificial Stokes images and on real data, and present both numerical and visual results.

2. MODEL DESCRIPTION

Each image to be restored consists of 4 channels. For pixel n, we group the scalars of each observed channel into vector $g^n = [g_1^n \ g_2^n \ g_3^n \ g_4^n]^T$ and let $g = \{g^n\}_{n=1}^N$ denote the set of vectors g^n. These vectors are considered as stemming from N corresponding Stokes vectors $s = \{s^n\}_{n=1}^N$, $s^n = [s_1^n \ s_2^n \ s_3^n \ s_4^n]^T$. Thus g is an indirect and noise corrupted version of s. The 4×4 observation matrix H is supposed to be known. Every Stokes channel corresponds to a specific function of the covariance between orthogonal electric vector components [3]. Consequently the Stokes vectors are subject to the following constraints:

$$s_1^n \geqslant 0, \quad (s_1^n)^2 \geqslant (s_2^n)^2 + (s_3^n)^2 + (s_4^n)^2. \tag{1}$$

We model the noise on each Stokes channel as zero mean, additive, white Gaussian. Formally this translates to

$$g^n | s^n, V \sim \mathcal{N}(H s^n, V), \tag{2}$$

where the covariance matrix is $\boldsymbol{V} = \sigma^2 I$.[1]

We take advantage of our probabilistic generative model formulation and assume a prior distribution on the real Stokes vector set $\boldsymbol{s}$. Such prior knowledge is the intuitive fact that vectors with spatially neighboring coordinates are more likely to have values close to one another. In order to implement this hypothesis, we first assume the existence of an underlying segmentation of the polarimetric image in K segments. The segmentation is defined by the set of $K \times 1$ vectors $\boldsymbol{z} = \{\boldsymbol{z}^n\}_{n=1}^N$. Each member $\boldsymbol{z}^n$ is defined as a vector with its kth variate set to 1 if the corresponding n^{th} Stokes vector $\boldsymbol{s}^n$ belongs to the kth segment; otherwise, it is set to zero. Also, every Stokes vector is assumed to belong to exactly one segment.

For each image segment k we define a probability distribution (*kernel*) generating the Stokes vectors belonging to the corresponding segment. While in image segmentation problems the chosen distribution is typically Gaussian [4], in the present problem we need a kernel choice that will assign zero probability mass to vectors not complying with the constraints given in (1). To this end, the Gaussian distribution is not convenient. To work around this particularity, we can instead consider a probability distribution on a special parametrization of the Stokes vectors. This parametrization is one of the contributions of the present work.

Let $\boldsymbol{\lambda} = [\lambda_1\ \lambda_2\ \lambda_3\ \lambda_4]^T$ be a parametrization $\boldsymbol{\varphi}$ of the Stokes vector $\boldsymbol{s} = [s_1\ s_2\ s_3\ s_4]^T$, with $\boldsymbol{s} = \boldsymbol{\varphi}(\boldsymbol{\lambda})$. Let matrices $\boldsymbol{\Phi}$ and $\boldsymbol{\Lambda}$ be

$$\boldsymbol{\Phi} = \begin{bmatrix} s_1 + s_4 & s_2 - i\,s_3 \\ s_2 + i\,s_3 & s_1 - s_4 \end{bmatrix}, \quad \boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ \lambda_3 + i\,\lambda_4 & \lambda_2 \end{bmatrix}.$$

The transformation $\boldsymbol{\varphi}$ is defined to comply with

$$\boldsymbol{\Phi} = \boldsymbol{\Lambda}\boldsymbol{\Lambda}^H. \tag{3}$$

It can be easily seen that any real vector $\boldsymbol{\lambda}$ will yield, according to transformation (3), a Stokes vector $\boldsymbol{s}$ that will necessarily satisfy the constraints (1). We can conveniently assume a Gaussian *i.i.d.* distribution hypothesis on the $\boldsymbol{\lambda}$ vectors:

$$\boldsymbol{\lambda}^n | z_k^n = 1, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k \ \sim \ \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k). \tag{4}$$

In turn, the label vectors $\boldsymbol{z}$ are multinomially *i.i.d* distributed. This distribution is parametrized by the *contextual mixing proportions* set $\boldsymbol{\Pi}$:

$$\boldsymbol{z}^n | \boldsymbol{\pi}^n \sim Mult(\boldsymbol{\pi}^n).$$

The prior probability vectors $\boldsymbol{\Pi} = \{\boldsymbol{\pi}^n\}_{n=1}^N$ are subject to the positiveness $\pi_k^n \geqslant 0, \forall k \in [1, \ldots K], \forall n \in [1, \ldots N]$ and sum-to-unity constraints $\sum_{k=1}^K \pi_k^n = 1, \forall n \in [1, \ldots N]$.

Considering the set of *contextual mixing proportions* $\boldsymbol{\Pi}$ as random variables and assuming a proper prior, we can incorporate the spatial smoothness trait, of which we have referred to earlier in the section, in an indirect way by forcing neighboring Stokes vectors to be more likely to share the same class label. We assume a Markov random field on $\boldsymbol{\Pi}$, which equivalently means that $\boldsymbol{\Pi}$ is governed by a Gibbs distribution [5], generally expressed by:

$$p(\boldsymbol{\Pi}) \propto \prod_C e^{-\ _c(\boldsymbol{\Pi})}, \tag{5}$$

where $_c$ is a function on clique c, called *clique potential* function in the literature, and the product is over all minimal cliques of the Markov random field.

An appropriate clique distribution choice would be to assume that the local differences of *contextual mixing proportions* follow a Student-t distribution, with its peak set at zero. This choice, proposed in a natural image segmentation context in [6], also provides our model the properties of an edge-preserving line-process [5]. The probability law for local differences is thus set to the Student's-t distribution:

$$\pi_k^n - \pi_k^j \sim \mathcal{S}t(0, \beta_{kd}^2, \nu_{kd}),$$
$$\forall n \in [1, \ldots N], k \in [1, \ldots K], d \in [1, \ldots D], \forall j \in \gamma_d(n). \tag{6}$$

The parameters β_{kd} control how tightly smoothed we need the Stokes vectors of segment k to be. In (6), D stands for the number of a pixel's neighborhood adjacency types and $\gamma_d(n)$ is the set of neighbors of pixel indexed by n, with respect to the dth adjacency type. In our model, we assume 4 neighbors for each pixel (*first-order* neighborhood), and partition the corresponding adjacency types into horizontal and vertical, thus setting $D = 2$. This variability of parameter aims to capture the intuitive property that smoothness statistics may vary along clusters and spatial directions [7].

One can see that the assumption in (6) is equivalent to

$$\pi_k^n - \pi_k^j \sim \mathcal{N}(0, \beta_{kd}^2 / u_k^{nj}),$$
$$u_k^{nj} \sim \mathcal{G}(\nu_{kd}/2, \nu_{kd}/2), \quad \forall n, k, d, \quad \forall j \in \gamma_d(n),$$

where $\mathcal{N}$ and $\mathcal{G}$ represent a Gaussian and a Gamma distribution respectively. This breaking-down of the Student's-t distribution allows clearer insight on how our implicit edge-preserving line-process works. Since u_k^{nj} depends on datum indexed by n, each weight difference in the MRF can be described by a different instance of a Gaussian distribution. Therefore, as $u_k^{nj} \to +\infty$ the distribution tightens around zero, and forces neighboring *contextual mixing proportions* to be smooth. On the other hand, $u_k^{nj} \to 0$ signifies the existence of an edge and consequently no smoothing. This generative model can be examined in detail in figure 1.
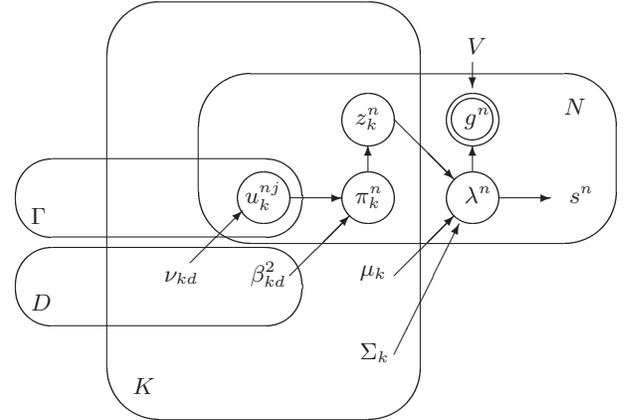
Fig. 1. Graphical model for the proposed Stokes image restoration model. Stokes vectors $\boldsymbol{s}^n$ constitute the estimated restoration, produced by observations $\boldsymbol{g}^n$. The rest of the model, namely random variable sets u_k^{nj}, π_k^n, z_k^n, constitute the prior for the proposed Stokes vector transformation $\boldsymbol{\lambda}^n$. Superscript $n \in [1, \ldots N]$ denotes pixel index, subscript $k \in [1, \ldots K]$ denotes kernel (segment) index, subscript $d \in [1, \ldots D]$ describes the neighborhood direction type and superscript $j \in [1, \ldots \Gamma]$ denotes neighbor index.

---

[1]The noise model is straightforwardly extensible to consider coloured noise.

3. MODEL INFERENCE

The problem of inference formulates in our case as, given observations $\boldsymbol{g}$, estimate the true Stokes image $\boldsymbol{s}$. To achieve this, we need to find estimates jointly for all unknown parameters $\boldsymbol{\Psi} = \{\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\beta}, \boldsymbol{\nu}\}$, $\boldsymbol{\Pi}$ and $\boldsymbol{\lambda}$; thence the restoration $\boldsymbol{s}$ can be computed given the $\boldsymbol{\lambda}$ estimate. The rest of the variables are considered as *hidden*, and are namely the labels $\boldsymbol{z}$ and the edge map $\boldsymbol{u}$. While we do not need to determine estimates for the hidden variables, they play an important role in the model inference in an indirect way.

Hence, we need to optimize the model evidence, given by

$$\ln p(\boldsymbol{g}, \boldsymbol{\lambda}, \boldsymbol{\Pi}; \boldsymbol{\Psi}), \qquad (7)$$

with respect to parameters $\boldsymbol{\Psi}, \boldsymbol{\lambda}, \boldsymbol{\Pi}$. In this sense, our method is a maximum *a posteriori* (MAP) estimation, with $\boldsymbol{\lambda}$ and $\boldsymbol{\Pi}$ being the conditioned variables. As we cannot find MAP estimates in closed form for (7), we employ the EM algorithm [8]. In EM terminology, eq.(7) is referred to as the *incomplete* likelihood, while the complete log-likelihood is expressed by

$$\ln p(\boldsymbol{g}, \boldsymbol{\lambda}, \boldsymbol{\Pi}, \boldsymbol{z}, \boldsymbol{u}; \boldsymbol{\Psi}). \qquad (8)$$

The conditional expectation of the complete likelihood is an important quantity in EM. It is defined as

$$\mathcal{E}_{\boldsymbol{z}^{(t)}, \boldsymbol{u}^{(t)} | \boldsymbol{g}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{\Pi}^{(t)}} \left\{ \ln p(\boldsymbol{g}, \boldsymbol{\lambda}, \boldsymbol{\Pi}, \boldsymbol{z}, \boldsymbol{u}; \boldsymbol{\Psi}) \right\}. \qquad (9)$$

By optimizing this expectation with respect to $\boldsymbol{\Psi}, \boldsymbol{\Pi}$ and $\boldsymbol{\lambda}$ given the observed variables and some initial estimates $\boldsymbol{\Psi}^{(0)}, \boldsymbol{\Pi}^{(0)}, \boldsymbol{\lambda}^{(0)}$ we can produce a new estimate $\boldsymbol{\Psi}^{(1)}, \boldsymbol{\Pi}^{(1)}, \boldsymbol{\lambda}^{(1)}$. In the same way, estimates are computed iteratively. It can be proved that these estimates converge to a local optimum for the incomplete likelihood of eq. (7). This is the main idea in the EM algorithm. The iteration scheme is split in two steps, the Expectation and the Maximization step.

The E-step consists in computing the joint expectation of the hidden variables $\boldsymbol{z}$ and $\boldsymbol{u}$, with respect to current iteration parameters $\boldsymbol{\Psi}^{(t)}, \boldsymbol{\lambda}^{(t)}, \boldsymbol{\Pi}^{(t)}$ where t denotes the number of current iteration. Observing the graphical model in fig.1, we can see that given $\boldsymbol{g}, \boldsymbol{\Pi}$ and $\boldsymbol{\lambda}$, $\boldsymbol{z}$ and $\boldsymbol{u}$ are conditionally independent. Therefore $\mathcal{E}_{\boldsymbol{z}, \boldsymbol{u} | \boldsymbol{g}, \boldsymbol{\lambda}, \boldsymbol{\Pi}}(\cdot) = \mathcal{E}_{\boldsymbol{z} | \boldsymbol{g}, \boldsymbol{\lambda}, \boldsymbol{\Pi}}\{(\mathcal{E}_{\boldsymbol{u} | \boldsymbol{g}, \boldsymbol{\lambda}, \boldsymbol{\Pi}}(\cdot)\}$ and we can compute these expectations separately.

Due to lack of space, we point the reader to [9] for the analytical expressions of the updates for parameters $\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\beta}, \boldsymbol{\nu}, \boldsymbol{\Pi}$, and for the expected values of $\boldsymbol{z}$ and $\boldsymbol{u}$, in a similar model proposed for natural image segmentation.

The update for the noise covariance matrix estimate $\boldsymbol{V}$ is given by

$$\boldsymbol{V}^{(t+1)} = (4N)^{-1} \sum_{n=1}^{N} (\boldsymbol{H} \boldsymbol{g}^n - \boldsymbol{s}^{n(t)})^T (\boldsymbol{H} \boldsymbol{g}^n - \boldsymbol{s}^{n(t)}) I. \quad (10)$$

Optimization with respect to the constraint-free parameters $\boldsymbol{\lambda}$ involves the following expression, after dropping constant terms from (9):

$$(\boldsymbol{g}^n - \boldsymbol{H} \boldsymbol{\varphi}(\boldsymbol{\lambda}^{n(t)}))^T \boldsymbol{V}^{-1} (\boldsymbol{g}^n - \boldsymbol{H} \boldsymbol{\varphi}(\boldsymbol{\lambda}^{n(t)})) +$$

$$\sum_{k=1}^{K} (\boldsymbol{\lambda}^{n(t)} - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\boldsymbol{\lambda}^{n(t)} - \boldsymbol{\mu}_k) <z_k^n>,$$

which after some manipulation boils down to:

$$\boldsymbol{h}^T \boldsymbol{\Omega}^1 \boldsymbol{h} + \boldsymbol{\lambda}^T \boldsymbol{\Omega}^2 \boldsymbol{\lambda} + \boldsymbol{\omega}^3 \boldsymbol{h} + \boldsymbol{\omega}^4 \boldsymbol{\lambda}, \qquad (11)$$

where we have omitted the data and iteration indices n and t for brevity. Parameters $\boldsymbol{h}, \boldsymbol{\Omega}^1, \boldsymbol{\Omega}^2, \boldsymbol{\omega}^3, \boldsymbol{\omega}^4$ are given by

$$\boldsymbol{h} \equiv \boldsymbol{H} \boldsymbol{\varphi}(\boldsymbol{\lambda}), \boldsymbol{\Omega}^1 \equiv \boldsymbol{V}^{-1}, \boldsymbol{\Omega}^2 \equiv \sum_{k=1}^{K} <z_k> \boldsymbol{\Sigma}_k^{-1},$$

$$\boldsymbol{\omega}^3 \equiv -2 \boldsymbol{g}^T \boldsymbol{V}^{-1}, \boldsymbol{\omega}^4 \equiv -2 \sum_{k=1}^{K} <z_k> \boldsymbol{\mu}_k^T \boldsymbol{\Sigma}_k^{-1}.$$

In view of (3), eq. (11) is a fourth-order polynomial over each of the variates of $\boldsymbol{\lambda}$. Setting the derivative of (11) with respect to each of the $\boldsymbol{\lambda}$ variates to zero, we can obtain optimizers for $\boldsymbol{\lambda}$ by solving the resulting third-order polynomial equations. Thus for each $n \in [1, \dots N]$, we solve iteratively four third-order polynomial equations, and repeat the operation until convergence of $\boldsymbol{\lambda}$.

Finally, in order compute the Stokes estimates $\boldsymbol{s}^n$ we simply make use of the $\boldsymbol{\varphi}$ transformation definition (11) to obtain the update

$$\begin{cases} s_1^n = \frac{1}{2} \left[(\lambda_1^n)^2 + (\lambda_2^n)^2 + (\lambda_3^n)^2 + (\lambda_4^n)^2 \right], \\ s_2^n = \lambda_1^n \lambda_3^n, \\ s_3^n = \lambda_1^n \lambda_4^n, \\ s_4^n = \frac{1}{2} \left[(\lambda_1^n)^2 - (\lambda_2^n)^2 - (\lambda_3^n)^2 - (\lambda_4^n)^2 \right]. \end{cases} \quad (12)$$

4. NUMERICAL EXPERIMENTS

We have applied the proposed recovery algorithm to two test Stokes images, one artificial image of size 64×64 and one real image of size 256×256. On the artificial image, the experiment was conducted by reproducing the blurring / noise model of (2) and applying varying levels of noise variance σ^2. We also used different assumed numbers of underlying segments K. The obtained results are shown in table 1. These are computed as the improvement over SNR for the degraded image, given by

$$ISNR = 20 \log_{10} \frac{\|\boldsymbol{s}^\star - \boldsymbol{g}\|}{\|\boldsymbol{s}^\star - \hat{\boldsymbol{s}}\|} \qquad (13)$$

where $\boldsymbol{s}^\star, \hat{\boldsymbol{s}}, \boldsymbol{g}$ represent the ground truth, the estimate, and the degraded (observed) image respectively. The pseudo-inverse estimates are as well computed for the restoration using $\hat{\boldsymbol{s}}^n = (\boldsymbol{H}^T \boldsymbol{H})^{-1} \boldsymbol{H}^T \boldsymbol{g}^n, \forall n \in [1, \dots N]$. The results clearly demonstrate that the proposed method gives a consistently better restoration compared to the general purpose pseudo-inverse estimator. Note also that the pseudo-inverse estimate will not necessarily yield values that satisfy the constraints (1). The main advantage of our method is that it takes explicitly into account the Stokes admissibility constraints (1), which is not the case for standard restoration methods. Representative visual results for the artificial image under $5dB$ noise are shown in figure 2. Corresponding results for a real Stokes image are shown in figure 3. The pseudo-inverse estimate of the real Stokes image contains 285 inadmissible vectors, out of a total of 65535 vectors; under the heavy-noise scenario of fig.2, this figure goes up to 1316 inadmissible vectors out of a total of 4096. Our method on the other hand, ensures always admissibility for *all* recovered vectors.

Runtimes for our algorithm were approximately 7 and 400 seconds respectively for the artificial and the much larger and more complex real Stokes image, for each EM iteration. We found that our algorithm converged in each case at around 10 EM iterations. Computations were done on a dual core 1.8 GHz PC workstation.

Table 1. *Restoration error results on the simulated Stokes data of figure 2. The image was degraded by varying noise levels. The presented values are the restoration ISNR (13); higher values correspond to better restorations. Results are shown for various numbers of classes K of the underlying segmentation, as well as the result of the pseudo-inverse estimate (PI).*

	PI	Proposed method		
SNR		$K=3$	$K=5$	$K=7$
$20dB$	21.4	22.4	22.3	22.7
$10dB$	11.4	14.7	14.5	14.5
$5dB$	6.3	10.9	9.9	9.9
$1dB$	2.3	7.3	7.6	7.6

    (a)   (b)   (c)   (d)
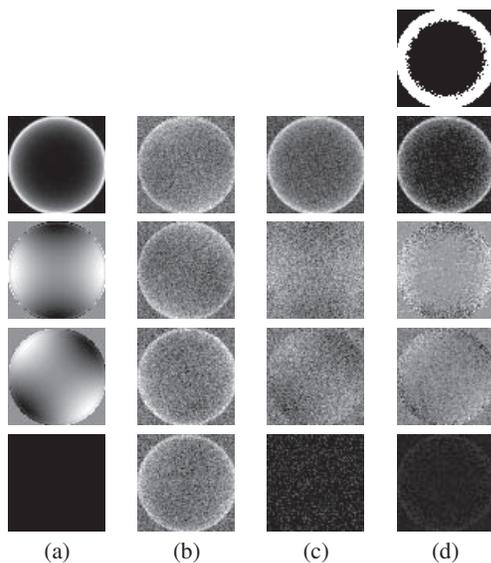
Fig. 2. Recovery result for simulated Stokes data under significant degradation. From left to right, each column shows the four channels of (a) the original Stokes image $s^\star$, (b) the degraded image g (SNR of 5 dB) (c) the non-complying to Stokes constraints pseudo-inverse recovery estimate and (d) the recovered image $\hat{s}$ obtained with our method. The corresponding segmentation of the degraded image into $K=2$ classes is shown at the top of the (d) column.

5. CONCLUSION

We have presented an image recovery methodology suitable for Stokes images. Making use of a smoothing prior which assumes an underlying image segmentation and a suitable Stokes vector parametrization, we are able to produce a good estimate of the real image that at the same time satisfies the Stokes vector constraints (1). Also, simultaneously we produce a segmentation of the input image due to the model structure. Future work could be directed towards more sophisticated prior models, adapted properly in order to handle the distinctive difficulties of the Stokes image recovery problem.

6. REFERENCES

[1] J. Zallat and C. Heinrich, "Polarimetric data reduction: a Bayesian approach," *Optics Express*, vol. 15, no. 1, pp. 83–96, 2007.
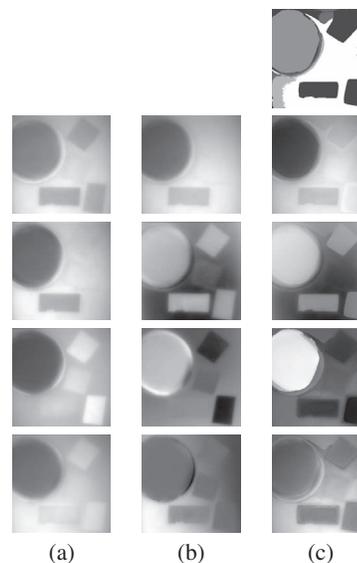
[2] J. Zallat, C. Heinrich, and M. Petremand, "A Bayesian approach for polarimetric data reduction : the Mueller imaging case," *Optics Express*, vol. 16, no. 10, pp. 7119–7133, 2008.

[3] J. Zallat, C. Collet, and Y. Takakura, "Clustering of polarization-encoded images," *Applied Optics*, vol. 43, no. 2, pp. 283–292, 2004.

[4] G. McLachlan, *Finite mixture models*, Wiley-Interscience, 2000.

[5] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 6, pp. 721–741, 1984.

[6] G. Sfikas, C. Nikou, N. Galatsanos, and C. Heinrich, "MR brain segmentation using an edge-preserving spatially variant Bayesian mixture model," in *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, New York, USA, 2008.

[7] C. Nikou, N. Galatsanos, and A. Likas, "A class-adaptive spatially variant mixture model for image segmentation," *IEEE Transactions on Image Processing*, vol. 16, no. 4, pp. 1121–1130, 2007.

[8] D. Tzikas, A. Likas, and N. Galatsanos, "The variational approximation for Bayesian inference," *IEEE Signal Processing Magazine*, vol. 25, no. 6, pp. 131–146, November 2008.

[9] G. Sfikas, C. Nikou, and N. Galatsanos, "Edge preserving spatially varying mixtures for image segmentation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Alaska, USA, 2008.

    (a)   (b)   (c)

Fig. 3. Recovery result for real Stokes data. From left to right, each column shows the four channels of (a) the observed image g (b) the non-complying to Stokes constraints pseudo-inverse recovery estimate and (c) the recovered image $\hat{s}$ obtained with our method. The corresponding segmentation of the observed image into $K=7$ classes is shown at the top of the (c) column.