On Global MDL-based Multichannel Image Restoration and Partitioning

Tetyana Ivanovska
University of Greifswald, Germany
tetyana.ivanovska@uni-greifswald.de

Horst K. Hahn
Fraunhofer MEVIS, Germany
horst.hahn@mevis.fraunhofer.de

Lars Linsen
Jacobs University, Germany
l.linsen@jacobs-university.de

ABSTRACT

In this paper, we address the problem of multichannel image partitioning and restoration, which includes simultaneous denoising and segmentation processes. We consider a global approach for multichannel image partitioning using minimum description length (MDL). The studied model includes a piecewise constant image representation with uncorrelated Gaussian noise. We review existing single- and multichannel approaches and make an extension of the MDL-based grayscale image partitioning method for the multichannel case. We discuss the algorithm's behavior with several minimization procedures and compare the presented method to state-of-the-art approaches such as Graph cuts, greedy region merging, anisotropic diffusion, and active contours in terms of convergence, speed, and accuracy, parallelizability and applicability of the proposed method.

Keywords: Segmentation, Denoising, Minimum Description Length, Energy Minimization, Multichannel images.

1 INTRODUCTION

The goal of image partitioning is to detect and extract all regions of an image which can be distinguished with respect to certain image characteristics. A special form of image partitioning is image segmentation, where one or a few regions of given characteristics are separated from the rest of the image. If the underlying image is supposed to be piecewise constant, image partitioning is equivalent to the restoration of that image, which is often obtained using denoising algorithms. Although there exist plenty of methods for solving image partitioning, segmentation, and restoration problems, many of them are task-specific or require intensive user interaction and triggering of a large number of parameters.

In this paper, we restrict ourselves to generic solutions using automatic methods based on energy minimization. Hence, two tasks need to be tackled: First, one needs to formulate an energy functional based on some assumptions and, second, one needs to find an appropriate and computationally feasible minimization algorithm. The main emphasis of this work is put on analysis and discussion of an energy functional, based on the assumption that the processed image is multichannel, piecewise constant, and affected by white Gaussian noise. We analyse and compare several minimization procedures and draw analogies to other approaches. Color images are used as examples, as they document the algorithmic behavior in an intuitive manner. However, all steps work are applicable for any multichannel image data. The paper is organized as follows. In Section 2 the related work in this area is described. In Section 3 we formulate the energy functional and describe two minimization procedures. Our findings are presented and discussed in Section 4.

2 RELATED WORK

Global energy minimization approaches originate from such seminal works as the ones by Mumford and Shah [24] and Blake and Zisserman [2]. The Markov Random Fields (MRF) framework (see the seminal work of Geman and Geman [8]) is a stochastic branch of the energy minimization approaches. In the last years a great breakthrough has been done in the direction of Markov random fields methods [19]. Such methods as Graph Cuts [5] are mathematically well described and allow to find a solution that lies close to the global optimum.

The minimum description length (MDL) based approach [15] uses basic considerations from information theory [19] in the formulation of the energy term, in the sense that the best image partitioning is equivalent to obtaining the minimum description length of the image with respect to some specific description language.

There exist two main tendencies in further developement of this approach. The first one consists in extending the functional to be minimized. These extentions include either different noise models (not only white Gaussian noise), or elimination of the model parameters that must be manually tuned by a user.

Kanungo [12] et al. formulated the functional for multiband and polynomial images. Lee [17] considered correlated noise. Galland [7] et al. considered speckle, Poisson, and Bernoulli noise. Zhu and Yuille [31] proposed an algorithm combining region growing, merging, and region competition which permits one to segment complex images. In several approaches [20, 16], an extended version of the functional is used and all user-defined parameters are eliminated. For example, Luo and Khoshgoftaar [20] propose to use the well known Mean-Shift method [6] to obtain the initial segmentation and start the MDL-based region merging procedure from it.

These approaches utilize region growing, as the minimization of the extended functional is infeasible. However, such an iterative technique, generally, does not lead to a stable local minimum and can give much coarser results, when compared to the relaxation technique, if the procedure of region merging is non-reversible. Moreover, the authors state, for instance, in [20], that the initial region selection has a strong impact on the efficiency and effectiveness of the region merging.

The second direction is to use a simplified or limited model with some user-defined parameters, but apply a global minimization procedure, which allows one to find at least a stable local minimum. Kerfoot and Bresler [13] formulated a full MDL-based criterion for piecewise constant image partitioning, but the class of images is limited to the class of simply-connected ones. For minimization such methods as Graph Cuts [5] have gained in popularity in the last years.

3 METHODS

3.1 Multichannel Model Description

The fundamental idea behind the Minimum Description Length (MDL) [19, 23] principle is that any regularity in the given data can be used to compress the data.

The image partitioning problem with respect to the MDL principle can be formulated as follows: Using a specified descriptive language, construct the description of an image (code) that is simplest in the sense of being shortest (when coded, needs the least number of bits) [15]. Let $L(M)$ denote the language for describing a model M and $L(D|M)$ the language for describing data D given model M. Moreover, let $|.|$ denote the number of bits in the description. The goal is to find the model M that minimizes the code length $C_l = |L(M)| + |L(D|M)|$. This corresponds to the two-part MDL code [9]. If the a priori probabilities $P(M)$ of the described models are known, then the number of bits in the description equals the negative base-two logarithm of the probability of the described models [19]: $|L(M)| = -\log_2 P(M)$.

In terms of image partitioning and restoration the code length can be written as $C_l = |L(u)| + |L(z-u)|$,where the model we are looking for is the underlying image representation (or partitioning) u that minimizes the code length. The term z describes the initial (or given) image, and the difference $r = (z-u)$ between the given image z and the partitioning u corresponds to the noise in the image. The noise describes the data with respect to model u.

A simple implementation of the MDL principle for image partitioning was presented by Leclerc [15, 23]: he assumed a piecewise constant model and derived the functional (or energy term)

$$C_l = \frac{b}{2}\sum_{i\in I}\sum_{j\in N_i}\left(1-\delta\left(u_i-u_j\right)\right) + a\sum_{i\in I}\left(\frac{z_i-u_i}{\sigma}\right)^2, \quad (1)$$

where u denotes the underlying image, z the given image, and σ^2 the noise variance. Moreover, $\delta(u_i-u_j)$ denotes the Kronecker delta (1 if $(u_i = u_j)$, else 0), I denotes the range of the image, and N_i is the neighbourhood of the ith pixel, a and b are constants. The first term in Equation (1) encodes the boundaries of the regions, whereas the second term encodes the noise in form of uncorrelated white Gaussian noise.

One can observe the similarities between the functional in Equation (1) with constants a, b, σ and the energy considered in different MRF approaches, namely in the Graph Cut methods (see [5] for more details): $E(f) = \lambda \sum_{(p,q)\in N} V_{p,q}(f_p, f_q) + \sum_{p\in I} D_p(f_p)$, where the interaction potential V between pixels p, q having labels (colors) f_p, f_q is taken from the Potts model (which corresponds to the Kronecker deltas), D is the distance between the initial and current colors of the pixel p (which corresponds to the noise values), and λ is a constant.

We expand on this approach to derive a multichannel image description length. For encoding the model, i.e., deriving $L(u)$, we have to encode the boundaries of the regions. To do so, we calculate the number of pixels that contribute to the boundary. Hence, the codelength for the boundary encoding is given by $|L(u)| = \frac{b}{2}\sum_{i\in I}\sum_{j\in N_i}\left(1-\prod_{k\in Ch}\delta\left(u_i^k-u_j^k\right)\right)$, where k denotes the channel, Ch is the range of channels (e.g., RGB in the three-channel color case), and other notations are as above. To encode the data that do not fit the model, i.e., the noise, we derive $L(z-u)$ assuming that the values in each channel are subject to white Gaussian noise with parameters $(0, (\sigma^k)^2)$. This assumption implies that the noise between channels is not correlated. Such an assumption allows for better understanding of the underlying processes and is often sufficient for many applications, since the uncorrelated noise can appear during the transmission, storing, or manipulation of images [1]. Moreover, the assumption holds when a multichannel image is combined from different independent modalities for processing.

The codelength of the noise is derived as

$$|L(z-u)| = -\sum_{i\in I}\sum_{k\in Ch}\log_2 P(r_i^k) \qquad (2)$$
$$= -\sum_{i\in I}\sum_{k\in Ch}\log_2\left(\frac{q}{\sqrt{2\pi(\sigma^k)^2}}\exp\left(-\frac{(r_i^k)^2}{2(\sigma^k)^2}\right)\right)$$
$$= \frac{1}{2\ln 2}\sum_{i\in I}\sum_{k\in Ch}\left(\frac{r_i^k}{\sigma^k}\right)^2 + const$$

where $q=1$ is the pixel precision, *const* is an additive constant, which is discarded, if $\left(\sigma^k\right)^2$ is considered constant and equal for all channels.

The resulting codelength functional becomes $C_l = \frac{b}{2}\sum_{i\in I}\sum_{j\in N_i}\left(1-\prod_{k\in Ch}\delta\left(u_i^k-u_j^k\right)\right) + \frac{1}{2\ln 2}\sum_{i\in I}\sum_{k\in Ch}\left(\frac{r_i^k}{\sigma^k}\right)^2$. This functional corresponds to the one considered in the Graph cuts method up to the constant weights.

3.2 Minimization

Having formulated the energy functionals, one needs to minimize them for a given image in order to compute the image partitioning. As it has been shown that computing the global optimum even of the simplest functional is an NP-hard problem [5], in practice one has to look for efficient approximations for it.

In the current paper, we will deal with two optimization approaches for energy minimization: a discrete one and a continuous one, namely, the α-expansion Graph Cut algorithm, introduced by Boykov [5], and the GNC-like approach, introduced by Leclerc [15], with several modifications, which have been done to check the convergence.

Graph cuts is an efficient minimizing technique that allows for finding a local minimum within a known factor of the global one. This algorithm belongs to the class of discrete optimization. Here, we give the outline of the algorithm and refer the reader to the original paper by Boykov et al [5] for further details. Let S and $\mathscr{L}$ denote image pixels (lattice) and the palette (set of all possible colors), correspondingly. The labeling u is described as $\{\mathscr{S}_l | l\in\mathscr{L}\}$, where $\mathscr{S}_l = \{p\in S | u_p = l\}$ is a subset of pixels with assigned color l. Given a label α a move from a labeling u to a new labeling u' is called an α- expansion if $\mathscr{S}_\alpha \subset \mathscr{S}'_\alpha$ and $\mathscr{S}'_l \subset \mathscr{S}_l$ for any label $l\neq\alpha$. In other words, an α-expansion move allows any set of image pixels to change their labels to α [5]. The minimum of the energy E for each label α is found by constructing a graph and finding the minimum cut for it. It is efficiently done by the algorithm developed by Boykov and Kolmogorov [3].

Start with an arbitrary partitioning u
repeat
  Set *success* $\leftarrow 0$
  for all $\alpha\in\mathscr{L}$ **do**
    Find $\hat{u} = \arg\min E(u')$ among u' within one α-expansion of u
    if $E(\hat{u}) < E(u)$ **then**
      $u \leftarrow \hat{u}$
      success $\leftarrow 1$
    end if
  end for
until *success* $\neq 0$
Return u

A Relaxation method using ideas of Graduated Non Convexity (GNC) by Blake and Zisserman [2] was proposed by Leclerc [15]. This is a continuous optimization method, where the labeling u is not selected from the given palette, as in the Graph Cuts case, but is iteratively computed. Here, $u\in\mathbb{R}^n$. The basic concept of the minimization procedure is to replace the non-convex codelength functional $C_l(u)$ by an embedding in a family of continuous functions $C_l(u,s)$, where $s\in\mathbb{R}$ is a user-defined parameter, that converge towards the target functional $C_l(u)$ when s goes to zero. $\lim_{s\to 0} C_l(u,s) = C_l(u)$. For the starting value of s, the functional $C_l(u,s)$ is convex such that standard convex minimization procedures can compute the single minimum. When s approaches zero, number and positions of the local minima of $C_l(u,s)$ become those of C_l. The minimization procedure iterates over s, which steadily decreases, and minimizes $C_l(u,s)$ for the respective value of s in each iteration step.

To obtain a continuous embedding, the discontinuous parts in functional C_l need to be replaced by a continuous approximation. The discontinuity of C_l is due to the use of the function δ. Hence, function δ is replaced by a continuous approximation that converges to δ when s goes to zero [15]. We use the approximation $\delta\left(u_i^k-u_j^k\right)\approx\exp\left(-\frac{\left(u_i^k-u_j^k\right)^2}{\left(s\sigma^k\right)^2}\right) = e_{ij}^k$.

The minimization iteration starts with a sufficiently large value $s=s^0$ and computes the (global) minimum u^0 of the convex functional $C_l(u,s^0)$. In each iteration step $T+1$, we set $s^{T+1} = rs^T$, where $0<r<1$, and compute the local minimum of $C_l(u,s^{T+1})$ starting from minimum u^T of the previous iteration step. The iteration is repeated until s is sufficiently small, i.e., until $s<\varepsilon$ with ε being a small positive threshold.

To compute the local minimum u on each iteration we apply Jacobi iterations [26]. The functional $C_l(u,s)$ is convex, if we choose a value for s that satisfies $(x_i^k - x_j^k)^2 \le 0.5\left(s\sigma^k\right)^2$ for all pixels i and j with $i\neq j$) and all channels k. Hence, when this condition is met, the local minimum must be a global one. The condition needs to be fulfilled for the starting value $s=s^0$. Then, the condition $\frac{\partial C_l(u,s^T)}{\partial u_i^k} = 0$for the local minimum at iteration step T becomes

$$\frac{2a\left(u_i^k-z_i^k\right)}{\left(\sigma^k\right)^2} + \frac{b}{2}\sum_{j\in N_i}\left[\frac{2\left(u_i^k-u_j^k\right)}{\left(s^T\sigma^k\right)^2}\prod_{l\in Ch}e_{ij}^l\right] = 0, \qquad (3)$$

where constant $a = (2\ln 2)^{-1}$. As Equation (3) cannot be solved explicitly, we use an iterative approach, where at each iteration step $t+1$ we compute

$$u_i^{k,t+1} = \frac{z_i^k + \frac{b}{a\,(s^T)^2} \sum_{j \in N_i} u_j^{k,t} \prod_{l \in Ch} e_{ij}^{lt}}{1 + \frac{b}{a\,(s^T)^2} \sum_{j \in N_i} \prod_{l \in Ch} e_{ij}^{lt}}. \qquad (4)$$

In total, we have two nested iterations. The outer iteration denoted by T iterates over s^T, while the inner iteration denoted by t iterates over $u_i^{k,t+1}$. Considering the behavior of the exponential function, the termination criterion for the inner loop is given by $\left| u_i^{k,t+1} - u_i^{k,t} \right| < s^T \sigma^k, \quad \forall i \in I$. Starting with $u = z$, the minimization procedure can be summarized by the following pseudocode:

while $s \geq \varepsilon$ **do**
  start with local minimum for u found in previous iteration
  while termination criterion for u is not met **do**
    recalculate u using Equation (4)
  end while
  update s
end while

Comparison to Anisotropic Diffusion. The derived iterative scheme for computing u_i^{t+1} in the single-channel case, cf. [15], is similar to the iteration in the well-known anisotropic diffusion approach. The continuous form of the Perona-Malik equation [25] for anisotropic diffusion is given by

$$\frac{\partial I}{\partial t} = div(g(\|\nabla I\|) \cdot \nabla I), \qquad (5)$$

where I denotes the image and function g is defined by $g(\|\nabla I\|) = \exp\left(-\left(\frac{\|\nabla I\|}{K}\right)^2\right)$ with flow constant K. The discrete version of Equation (5) is given by $I_i^{t+1} - I_i^t + \lambda \sum_{j \in N_i} \left(I_i^t - I_j^t\right) \exp\left(-\left(\frac{I_i^t - I_j^t}{K}\right)^2\right) = 0$, where λ is a normalization factor.

The local minimum of the MDL-based energy C_l in single-channel notation (for a fixed s) is defined by solving $\frac{\partial C_l(u,s)}{\partial u_i} = 0$, which leads to

$$(u_i - z_i) + \alpha \sum_{j \in N_i} \left[(u_i - u_j) \exp\left(-\frac{(u_i - u_j)^2}{(s\sigma)^2}\right)\right] = 0.$$

The continuous version of this equation can be written as

$$\int_0^{t_{end}} \frac{\partial u}{\partial t} dt = div(g(\|\nabla u\|) \cdot \nabla u), \qquad (6)$$

where $u_0 = z$ and $u_{t_{end}}$ describes the image values at the current step t_{end}. Comparing the continuous form of the Perona-Malik equation (5) with the continuous form of the MDL-based equation (6), one can immediately observe the similarity. The main difference is the integral on the left-hand side of Equation (6). The integral represents the changes between the current image and the initial image. Thus, this version of the MDL-based minimization algorithm can be considered as an "anisotropic diffusion algorithm with memory".

In general, the assumptions about the convexity of $C_l(u,s)$ do not hold anymore, when s is small. Hence, the relaxation method with Jacobi iterations can give an inadequate result. To check this we implemented the steepest descent minimization scheme [26] and compared the results.

4 RESULTS AND DISCUSSION

The presented algorithm belongs to the class of algorithms that are not purely for denoising or segmentation, but can be treated as an elegant combination of both approaches. Usually, some part of meaningful image data might be lost on the denoising step, which affects the subsequent segmentation results. Here, on the contrary, the complete information is used both for boundary preservation and noise exclusion, which is referred in the literature as image restoration [28] or reconstruction [25].

The algorithm (minimization procedures with Jacobi iterations, Gradient descent, and Graph cuts) was implemented using C/C++ programming language, optimized appropriately, and was compiled under gcc4.4. The Graph cuts code for grayscale images was kindly provided on the Vision.Middlebury web-site [22].

Figure 1 documents the general behavior of the Relaxation scheme. Starting with a synthetic 100×100 image with manually added noise, we apply the Jacobi iteration minimization procedure. The three rows in Figure 1 show (intermediate) results at iteration $T = 4$, $T = 600$, and $T = 1484$, respectively. The first column shows the currently detected underlying image and the second column the currently removed noise, i.e., the difference between the initial image and the currently detected underlying image (first column of Figure 1). For the generation of the images in column three, we picked one row of the image, i.e., a horizontal cut through the image, and show the current values of u for that single row. The third column in Figure 1 documents nicely, how individual values (first row) start assimilating and grouping together (second row) to eventually form a piecewise constant representation (third row).

To qualitatively describe the restoration results, we evaluated the difference between the resulting and initial (not noisy) images using two measures. The first one is the mean squared error (MSE) estimator [18], which is suitable to show whether the colors were reconstructed as close to

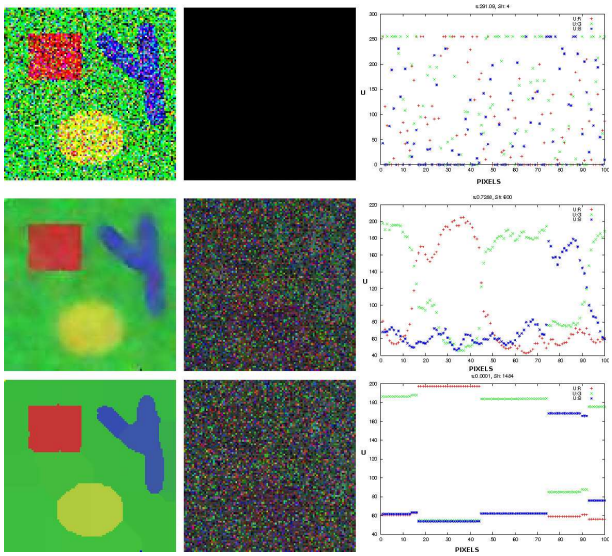

Figure 1: Different phases of MDL minimization on timesteps 4, 600, and 1484. The images from left to right are: the detected underlying image, the removed noise, image values u of one row (The pixel coordinates are given in x axis, and u values are in y axis).

the initial ones as possible, i.e., for the denoising part. For the multichannel case, MSE is defined as $MSE = \frac{1}{cwh}\sum_{c_1\in c}\sum_{y\in h}\sum_{x\in w}(In(x,y,c_1)-Res(x,y,c_1))^2$,where c is the number of color channels, w is the image width, h is the image height, and In and Res define the initial and resulting image, respectively.

The second metric is utilized to evaluate the segmentation results, i. e., oversegmentation and inaccurate boundary localization. It is based on the one proposed by Mezaris et al [21]. This metric is defined as follows. Let $S = s_1, s_2, ..., s_K$ be the segmentation mask to be evaluated, comprising K regions s_k, $k = [1,K]$, and let $R = r_1, r_2, ..., r_Q$ be the reference mask, comprising Q reference regions r_q, $q = [1,Q]$. Each region r_q is associated with a different region s_k, i. e. s_k is chosen such that the overlap $r_q \cap s_k$ is maximized. Let $A(r_q, s_k)$ denote the set of region pairs and let N_S denote the set of non-associated regions of mask S. Energy E_b is used for the evaluation, values closer to zero indicate better segmentation: $E_b = \sum_{q=1}^{Q} E_q + \sum_{s_k\in N_S} F_k$, where E_q and F_k are defined are follows.

$$E_q = \sum_{p\in(r_q - r_q\cap s_k)} f_1(p,r_q) + \sum_{p\in(s_k - s_k\cap r_q)} f_2(p,r_q) \quad (7)$$

$$F_k = \alpha \sum_{p\in s_k} f_1(p,r_q) \quad (8)$$

E_q is a weighted sum of misclassified pixels for region pair (r_q, s_k) $\forall (r_q, s_k) \in A$. F_k is a weighted sum of misclassified pixels $\forall s_k \in N_S$. f_1 and f_2 are weight functions, proposed by Villegas et al [30] to deal with the fact that the distance of a misclassified pixel from the boundary of the reference region to which it belongs affects the visual relevance of the error [21]. f_1 is used for false negatives and f_2 is used for false positives.

$$f_1(p,r_q) = d(p,r_q) * 10^{-4},$$
$$f_2 = \begin{cases} d(p,r_q) * 10^{-4} & d(p,r_q) < 10 \\ 10^{-3} & otherwise \end{cases}.$$

Moreover, d is the Euclidean distance between the pixel p and the boundary of the region r_q. α is a weight parameter which was heuristically set to 100 in our experiments, since we would like to penalize oversegmentations.

First, we run a series of tests on synthetic images with artificially added Gaussian noise applying three algorithms: Graph Cuts, Relaxation with Jacobi iterations, and Relaxation with steepest descent. Figures 2 and 3 demonstrate the high-quality results produced by both Graph Cuts and Relaxation with gradient descent methods. The exact regions are reconstructed with few misclassifications. Graph Cuts restore the closest colors to the initial image ($MSE = 830.65$).

However, the computational costs of these methods are rather high. For the Graph cuts method the execution time is dependent on the palette size, i. e., the number of colors (labels) which would be considered for expansion. Ideally, the palette should include the whole color space. In this case, the optimum will be found accurately. Boykov et al. proposed to reduce the space of labels for this case [4], taking, for instance, only unique colors of the image and their closest neighbors as labels. Such a choice allows for a massive label space reduction, but leaves enough variability for the label expansion. In our experiments the computation for a 100×100 color image with $50,000$ labels (unique colors and their closest neighbours of a noisy synthetic image) took around 10 minutes per iteration. As usually several iterations are needed, this approach appears to be infeasible for bigger datasets.

The relaxation scheme with the Gradient Descent method for a 100×100 color image takes in several hours, which makes this method inapplicable for real datasets.

The relaxation scheme with Jacobi iterations restores more regions than required, however, they have very close colors and the "weak" region boundaries can not be distinguished by a human eye. Due to such color closeness, these regions can be merged into one using a simple region-merging procedure. Starting at any pixel (seed point), the algorithm subsequently adds neighbouring pixels to the region if the distance in color space between them and the region color is lower than a certain threshold. The region color is the average color of the pixels that belong to it. The algorithm stops when all pixels belong to some regions. Computational costs of the region merging procedure are negligible

when compared to the ones of the Relaxation method. When the simple region merging is applied, we obtain $E_b = 0.0021$, which is slightly better than the other results.

The main advantage of the Jacobi iteration approach is its independence of the palette size and straightforward parallelizability, such that computations only take seconds [11], which makes this method attractive for real applications.

Next, we would like to compare the presented method to other (similar) approaches, namely, anisotropic diffusion, region merging, and a combination of these two procedures. Although anisotropic diffusion is primarily used to remove noise from digital images without blurring edges, it can be used in edge detection algorithms. By running the diffusion with an edge seeking diffusion coefficient for a certain number of iterations, the image can be evolved towards a piecewise constant image with the boundaries between the constant components being detected as edges [25]. We compared our approach to the anisotropic diffusion algorithm for multichannel images, presented by Sapiro and Ringach [27]. This method similarly to ours considers the influence of all image channels at once, which allows for better boundary preservation when compared to the methods treating each image channel separately. However, our approach has a "memory", as it always refers to the initial image instead of iterating from the result obtained on the previous time step. Such a behavior should guarantee a better boundary preservation and initial color restoration if the parameters are properly chosen. Figure 4 demonstrates that Sapiro's method successfully eliminates the noise (except from some mistakes on the image boundaries) and preserves most of the object boundaries. Our approach appears to be stable to noise, preserves object boundaries, restores the colors close to the initial ones, and produces the most uniformly colored background.

Due to the energy functional complexity, which arises for the models with no manually adjustable parameters or for more complicated noise models, it is often difficult to find a feasible minimization procedure. Thus, many authors follow a "greedy" region merging strategy [12, 20, 14, 16]. The idea of the algorithm is as follows. It starts from some initial oversegmentation. Then at each timestep it chooses two neighbouring regions and merges them to form a new region. These two regions are chosen in such a way that, when they are merged, it provides the largest energy reduction amongst all other possible merges. Although this procedure is very fast, it has several drawbacks. First, it does not guarantee convergence to a stable local minimum, especially, in the presence of noise. Second, the way the initial oversegmentation is chosen also affects the results.

Although Kanungo et al. [12], Luo and Khoshgoftaar [20], Lee [16], and Koepfler et al. [14] propose to use different functionals for color image segmentation, the region merging procedures are similar to each other. We assume that a more complicated functional could lead to a more adequate result. However, there is no guarantee about the convergence to a stable local minimum, and this issue stays unresolved.

We compared our approach to the algorithm proposed by Koepfler et al. [14]. They use the Mumford-Shah functional for the piecewise constant case, which coincides with the MDL-based functional C_l. Here, we computed both metrics to evaluate the results. As it can be observed in Figure 5, the iterative region merging produces results with some misclassified boundaries, which appears due the fact that the merging algorithm did not land close to the minimum, and not completely erased noise when compared to Figure 4. The relaxation approach can be approximated with a pipeline that utilizes several algorithms, namely, denoising, boundaries sharpening, and region merging. We constructed a pipeline from Sapiro's anisotropic diffusion and Koepfler's region merging. As it can be observed in Figure 5, the pipeline produces the best result in terms of boundary preservation and noise elimination when compared to the independent application of these methods. However, in general, each method in the pipeline requires additional adjustment of the parameters for each image.

Our comparison shows that the proposed approach produces the best results in terms of color (the lowest $MSE = 377.628$) and region restoration (the lowest $E_b = 0.0008$) for the test image.

We also compared our results to the results obtained with the Active contours without edges [29] method introduced by Chan and Vese. This is a variational approach based on energy minimization, and the energy is formulated using the Mumford-Shah functional for the piecewise constant case. For our tests, we applied the 4-phase version of the algorithm for color piecewise constant images with the parameters for each phase:$\lambda_{1,2} = 1$, $\nu = 0$, as it is recommended by the authors. We experimented with different values of parameter μ and initial contour locations. We executed the algorithm with max. 2000 iterations with the $timestep = 0.5$. Figure 6 documents that it is problematic to obtain decent results for images with low signal-to-noise ratio for the reasonable amount of time and the prior denoising is needed.

When applied to real images, our approach allows for obtaining the results with different levels of detail, which reminds one of the multi-scale theory of piecewise image modeling, introduced by Guigues et al. [10]. In Figure 7 the results illustrate this effect. Increasing the value of b allows us to reduce the

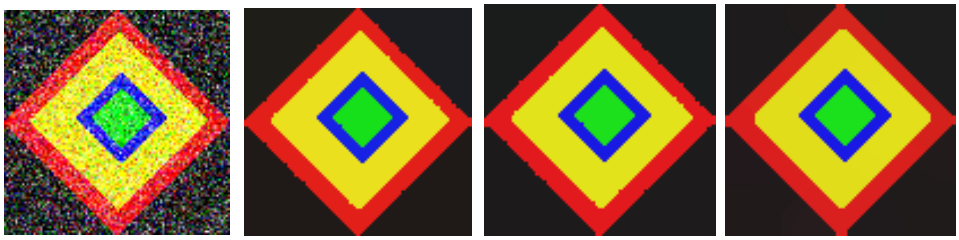

Figure 2: *First*: Input diamand image with manually added noise $N(0, 70^2)$. *Second*: Graph cuts result. $MSE = 830.65$. $E_b = 0.0027$ *Third*: Relaxation (Steepest descent on inner iterations) result. $MSE = 2273.8096$. $E_b = 0.0031$ *Fourth*: Relaxation (Jacobi iterations) result. $MSE = 2418.262$. No region merging $E_b = 651.287$. With region merging $E_b = 0.0021$

Figure 3: *First*: True image with 8 regions (5 unique colors). *Second*: Piecewise-constant regions are marked with random different colors on the Graph cuts result. The image consists of 8 regions (7 unique colors). *Third*: Regions are marked with random different colors on the Relaxation (Descent) result. The image consists of 8 regions. *Fourth and Fifth*: Piecewise-constant regions are marked with random different colors on the Relaxation (Jacobi) result. The image consists of 49 regions (before region merging) and 8 regions (after region merging).

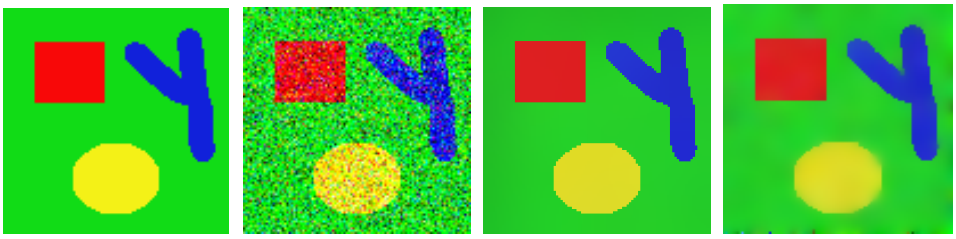

Figure 4: *First*: Initial synthetic image. *Second*: Synthetic image with added noise $N(0, 70^2)$. *Third*: Result for our approach with subsequent region merging. $MSE = 377.628$. $E_b = 0.0008$. *Fourth*: Result for Sapiro's anisotropic diffusion with parameters: $edge = 49$, $numStep = 45$. $MSE = 447.83$.

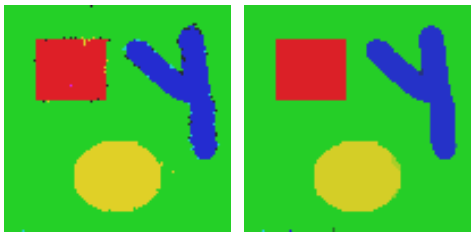

Figure 5: *First*: Result for iterative region merging with $\lambda = 10^6$. Noise is not erased completely. Further increasing the parameter v_0 does not improve the result. $MSE = 436.538$. $E_b = 6.5369$ *Second*: Result for the pipeline consisting of the anisotropic diffusion and region merging based on the Mumford-Shah functional. $MSE = 419.7864$. $E_b = 0.26$

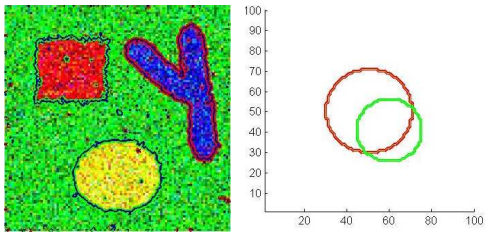

Figure 6: *First*: Result for Active contours. The resulting contours are still noisy. *Second*: Initial contours.

number of details, i.e., the number of regions, and leave only the "strongest" edges.

5 CONCLUSION AND FUTURE WORK

We generalized the existing single-channel MDL-based method to a multichannel one and applied it to color image partitioning and restoration. For synthetic color images, the results were compared to the ones obtained with Graph cuts, anisotropic diffusion, greedy region merging with energy, and active contours approaches. Our method produces best results in terms of color

Figure 7: Results with different *b* for the "woman" image. The initial image is the leftmost one. The parameter *b* is selected: 2, 5, 20 for the images from left to right, correspondingly.

restoration and boundary accuracy for our test cases. Moreover, the relaxation scheme with Jacobi iterations combined with a simple region merging procedure allows for fast computations due to straightforward parallelizability, which is important when applied to real-world problems.

We believe that this approach has potential in solving problems of image restoration. The future work directions include studying different noise models (e. g. correlated noise) for piecewise constant and piecewise smooth images as well as the models based on other MDL codes.

6 REFERENCES

[1] D. Baez-Lopez, F. H. Mendoza, and J. M. Ramirez. Noise in color digital images. *Circuits and Systems, Midwest Symposium on*, 0:403, 1998.

[2] A. Blake and A. Zisserman. *Visual Reconstruction*. MIT Press, 1987.

[3] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137, 2004.

[4] Y. Boykov, O. Veksler, and R. Zabih. Efficient restoration of multicolor image with independent noise. Technical report, 1998.

[5] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239, 2001.

[6] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619, 2002.

[7] F. Galland, N. Bertaux, and P. Refregier. Multi-component image segmentation in homogeneous regions based on description length minimization: Application to speckle, poisson and bernoulli noise. *Pattern Recognition*, 38(11):1926 – 1936, 2005.

[8] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6):721–741, 1984.

[9] P. Grunwald. *A tutorial introduction to the minimum description length principle*. 2004.

[10] L. Guigues, J. Cocquerez, and H. Le Men. Scale-sets image analysis. *International Journal of Computer Vision*, 68:289–317, 2006.

[11] T. Ivanovska, L. Linsen, H. K. Hahn, and H. Voelzke. Gpu implementations of a relaxation scheme for image partitioning: Glsl vs. cuda. *Computing and Visualization in Science*, 14(5):217–226, 2012.

[12] T. Kanungo and B. Dom et al. A fast algorithm for mdl-based multi-band image segmentation. *Computer Vision and Pattern Recognition*, pages 609–616, 1994.

[13] I.B. Kerfoot and Y. Bresler. Theoretical analysis of multispectral image segmentation criteria. *Image Processing, IEEE Transactions*, 8(6):798–820, 1999.

[14] G. Koepfler, C. Lopez, and M. Morel. A multiscale algorithm for image segmentation by variational method. *SIAM Journal of Numerical Analysis*, 31:282–299, 1994.

[15] Y.G. Leclerc. Constructing simple stable descriptions for image partitioning. *International Journal of Computer Vision*, 3(1):73–102, 1989.

[16] T. C. M. Lee. A minimum description length-based image segmentation procedure, and its comparison with a cross-calidation-based segmentation procedure. *Journal of the American Statistical Association*, 95(449):259–270, 2000.

[17] T. C. M. Lee and T. Lee. Segmenting images corrupted by correlated noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:481–492, 1998.

[18] E. L. Lehmann and George Casella. *Theory of Point Estimation (Springer Texts in Statistics)*. Springer, 2003.

[19] S. Z. Li. *Markov random field modeling in image analysis*. Springer-Verlag New York, Inc., 2001.

[20] Q. Luo and T.M. Khoshgoftaar. Unsupervised multiscale color image segmentation based on mdl principle. *IEEE Transactions on Image Processing*, 15(9):2755–2761, 2006.

[21] V. Mezaris, I. Kompatsiaris, and M. G. Strintzis. Still image objective segmentation evaluation using ground truth. In *In Proceedings of the Fifth COST 276 Workshop on Information and Knowl-*

edge Management for Integrated Media Communication, 2003.

[22] http://vision.middlebury.edu/MRF/ Middlebury Vision.

[23] A. Mitiche and I. B. Ayed. *Variational and Level Set Methods in Image Segmentation*. Springer Berlin Heidelberg, 2011.

[24] D. Mumford and J. Shah. Boundary detection by minimizing functionals. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1985.

[25] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(7):629–639, 1990.

[26] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C++: The Art of Scientific Computing*. Cambridge University Press, 2002.

[27] G. Sapiro and D. L. Ringach. Anisotropic diffusion of multivalued images with applications to color filtering. *Image Processing, IEEE Transactions on*, 5(11):1582–1586, 1996.

[28] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A comparative study of energy minimization methods for markov random fields with smoothness-based priors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(6):1068–1080, 2008.

[29] L. A. Vese and T. F. Chan. A multiphase level set framework for image segmentation using the Mumford and Shah model. *International Journal of Computer Vision*, 50(3):271–293, December 2002.

[30] P. Villegas, X. Marichal, and A. Salcedo. Objective evaluation of segmentation masks in video sequences. In *In Proceedings of Workshop om Image Analysis for Multimedia Interactive Services*, 1999.

[31] S.C. Zhu and A.L. Yuille. Region competition: unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *IEEE Trans. on PAMI*, 1996.