A Multiclass Anisotropic Mumford-Shah Functional for Segmentation of D-dimensional Vectorial Images

J. F. Garamendi[1] and E. Schiavi[2]

[1]*Departament de Tecnologies de la Informació i les Comunicacions, Universitat Pompeu Fabra, Barcelona, Spain*

[2]*Departamento de Matemática Aplicada, Ciencia e Ingenieria de los Materiales y Tecnología Electrónica, Universidad Rey Juan Carlos, Móstoles, Spain*

jf.garamendi@upf.edu, emanuele.schiavi@urjc.es

Keywords: Multi-channel Image Segmentation, Variational Methods, Mumford-Shah.

Abstract: We present a general model for multi-class segmentation of multi-channel digital images. It is based on the minimization of an anisotropic version of the Mumford-Shah energy functional in the class of piecewise constant functions. In the framework of geometric measure theory we use the concept of common interphases between regions (classes) and the value of the jump discontinuities of the (weak) solution between adjacent regions in order to define a minimal partition energy functional. The resulting problem is non-smooth and non-convex. Non-smoothness is dealt with highlighting the relationship of the proposed model with the well known Rudin, Osher and Fatemi model for image denoising when piecewise constant solutions (i.e partitions) are considered. Non-convexity is tackled with an optimal threshold of the ROF solution which we which generalize to multi-channel images through a probabilistic clustering. The optimal solution is then computed with a fixed point iteration. The resulting algorithm is described and results are presented showing the successful application of the method to Light Field (LF) images.

1 INTRODUCTION

Image segmentation is, possibly, one of the most important steps of any image analysis, recognition or image quantification process. Among the most established methods, minimization of a feasible energy functional has proven to be an efficient, accurate and sound based mathematical framework for digital image processing. A key stone in image segmentation is the celebrated Mumford and Shah model (Mumford and Shah, 1989) which is a piecewise model producing a partition (segmentation) of a digital image. Several mathematical and computational difficulties arise when this model is considered. Convex formulations have been proposed for solving a more general model in terms of multi-labelling problems, (Pock et al., 2009) and (Brown et al., 2011). A simplified version of the Mumford and Shah model in a curve evolution framework (level sets) is also considered in (Chan and Vese, 2001). Their proposal amounts to consider a coupled system of parabolic equations and, as such, it presents a strong dependency on the initialization of the numerical scheme and a extremely low stabilization rate to the steady state solution. Also, this model do not properly takes account of the local inter-phases between adjacent classes (some are counted twice and some are missed).

The framework we propose is a generalization of the model for binary (2-classes) segmentation of 2D grayscale images proposed by Osher and Vese (Osher and Paragios, 2003). We use the concept of weights for the inter-phase lengths introduced for binary segmentation and we generalize it to M-Classes segmentation on D-Dimensional vectorial images. This generalization leads to an anisotropic version of the Mumford-Shah model (hereafter AMS model). Moreover, we show that the resulting AMS model for multi-class multi-channel segmentation has the same minimum as the vectorial version of the celebrated Rudin, Osher and Fatemi (Rudin et al., 1992) denoising model (from now on VROF model) when restricted to (minimized in) the set of piecewise constant functions (which are $SBV(\Omega)$ functions). In fact the energy of the two functionals is exactly the same when this set is considered for minimization. In turn this allows to solve the segmentation problem using the well-established theory for the ROF model where existence and uniqueness was proved (Chambolle and Lions, 1997) and for which efficient numerical schemes where developed (Chambolle, 2004;

Garamendi J. and Schiavi E.
A Multiclass Anisotropic Mumford-Shah Functional for Segmentation of D-dimensional Vectorial Images.
DOI: 10.5220/0006127804680475
In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017)*, pages 468-475
ISBN: 978-989-758-225-7

Garamendi et al., 2013; Chambolle and Pock, 2011).

Possible applications of the proposed multiclass model are segmentation of a given Optical Flow, which is a 2D vectorial field, multimodal magnetic resonance image, considering each modality as a single channel, and in general multiclass segmentation task. In this work, as a specific application of the above general setting, we consider Light Field images (LF) partition. There is a growing interest for light field imaging applied to computer vision due to the new hand-held cameras such as Lytro[1] or Raytrix[2]. In fact, compared to conventional imaging, light field imaging increases the directional information of the scene. Plenoptic cameras (Lippman, 1908; Ng, 2006; Ng et al., 2005; Perwass and Wietzke, 2010) capture LF images from scene keeping the light direction information. The purpose of this cameras is capturing the amount of light (radiance) traveling along each ray that intersects the sensor. So, for each spatial position in the 2-D space of the sensor, it stores the amount of light coming from a certain space directions. So one acquisition, or data set, has the information about how the scene appears from a certain number of possible viewpoints. More details can be found in (Wanner et al., 2013a), (Reddy et al., 2013).

The mathematical modeling of the light filed images is usually done considering two planes of $\mathbb{R}^2$. One of them defines the spatial coordinates in a single view and the other one defines the view itself. Let $\Omega \subset \mathbb{R}^2$ and $\Pi \subset \mathbb{R}^2$ be bounded Lipschitz domains representing the spatial image domain and the angular domain respectively, and let $\mathbf{f}$ be a multi-channel data. The LF image can be modeled as the 4-D function

$$\begin{aligned} \mathbf{f}: \quad \Omega \times \Pi \quad &\to \mathbb{R}^M, \\ (\bar{p}, \bar{q}) \quad &\mapsto \mathbf{f}(\bar{p}, \bar{q}) \end{aligned}$$

where $\bar{p} := (x,y) \in \Omega$ and $\bar{q} := (s,t) \in \Pi$ represent coordinate pairs in the sensor plane spatial domain and in the view angular domain respectively. This provides a model of a lightfield color image as a vectorial 4D function, in such a way $\mathbf{f}(x,y,s,t)$ represents the color ($M = 3$ in the case of RGB color images) at pixel (x,y) corresponding to ray (s,t).

The structure of LF images allows for a very precisely disparity map computation with a very small cost (Wanner et al., 2013b), so one can assumes that this information is available as an additional feature to the intensity (gray) color, making this modality of image very well suited to segmentation. In this case, for a color image in a RGB color space, $M = 4$, the three first components corresponding to color information and the fourth to depth information. Notice that the user can add other information to the channels as for example local variance, texture, etc. This vectorial 4D dimensional structure of the images (color components and depth) allows to test the model presented in this work.

[1]www.lytro.com

[2]www.raytrix.de

2 NOTATION AND DEFINITIONS

Let $\Omega \subset \mathbb{R}^D$ be a bounded Lipschitz domain representing the digital image domain and let $\mathbf{f} : \Omega \subset \mathbb{R}^D \to \mathbb{R}^M$ be a given D-dimensional noisy image representing the data, where $M = 1$ for scalar images and $M > 1$ for vector valued (multichannel) images. As usual in image processing we assume $\mathbf{f} \in [L^\infty(\Omega)]^M$, i.e. f essentially bounded.

Given an image and chosen the number $N \geq 2$ of classes into which we wish to partition the given image, the segmentation problem can be formulated as the determination of a partition of the domain Ω into a collection of sets $\{\Omega_i\}_{i=1..N}$ of finite perimeter (Cacciopoli sets) in Ω such that no overlap and no vacuum can occur, e.g. $\Omega_i \cap \Omega_j = \emptyset$, $i \neq j$, $\Omega = \bigcup_{i=1}^N \Omega_i \cup \Gamma_i$ where the boundary of each class is denoted by $\Gamma_i = \partial\Omega_i \cap \Omega$. Given a partition $P(\Omega)$ we define $\bar{\chi} = (\chi_i)$, $i = 1..N$ as the associated vectorial characteristic function. For almost every point $x \in \Omega_i \subset \Omega$ we have $\bar{\chi} : \Omega \to \mathbb{R}^N$, $\bar{\chi}(x) = \bar{e}_i$ where $\bar{e}_i$ is a vector of the canonical base of $\mathbb{R}^N$. Notice that the dimension of $\bar{\chi}$ depends on the number of classes and it is independent from number of channels M. Also $\sum_{i=1}^{N} \chi_i(x) = 1, \text{a.e.} x \in \Omega$.

For a given vector valued function $\mathbf{u} : \Omega \to \mathbb{R}^M$ the vectorial TV norm, denoted as $\mathbf{TV}$, is defined by the finite positive measure (Ambrosio et al., 2000; Bresson and Chan, 2008)

$$|D\mathbf{u}|(\Omega) = \int_\Omega |D\mathbf{u}| \doteq \sup_{\mathbf{P} \in K} \left\{ \int_\Omega \langle \mathbf{u}, \nabla \cdot \mathbf{P} \rangle d\mathbf{x} \right\} \quad (1)$$

where $\mathbf{P} : \Omega \to \mathbb{R}^{M \times D}$ is a matrix dual function, $\nabla \cdot$ is the divergence operator and the product $\langle .,. \rangle$ is the Euclidean scalar product defined as $\langle \mathbf{v}, \mathbf{w} \rangle \doteq \sum_{i=1}^M \langle v_i, w_i \rangle$ from where $\langle \mathbf{u}, \nabla \cdot \mathbf{P} \rangle = \sum_{i=1}^M \langle u_i, \nabla \cdot \mathbf{p}_i \rangle$. The set K of functions of the dual variable $\mathbf{P}$ is

$$K \doteq \left\{ \mathbf{P} \in C_c^1 \left(\Omega; \mathbb{R}^{M \times D} \right) : |\mathbf{P}| \leq 1 \right\} \quad (2)$$

where $|\cdot|$ is the L^2 norm such that $|\mathbf{P}| = \sqrt{\sum_{i=1}^M \langle \mathbf{p}_i, \mathbf{p}_i \rangle}$.

Setting $M = 1$ and denoting the dual variable $\mathbf{P}$ in vectorial form $\mathbf{p}$ (scalar case) we can use (1) to define the perimeter of each subset Ω_i of the partition in form (Klann and Ramlau, 2013):

$$\text{per}(\Omega_i) \doteq |D\chi_i|(\Omega) = \int_\Omega |D\chi_i| = \text{TV}(\chi_i) = \quad (3)$$

$$\sup_{\mathbf{p}\in K}\left\{\int_\Omega \chi_i \nabla\cdot\mathbf{p}d\mathbf{x}\right\} = \sup_{\mathbf{p}\in K}\left\{\int_{\Omega_i} \nabla\cdot\mathbf{p}d\mathbf{x}\right\} = |\Gamma_i| \quad (4)$$

where χ_i is the characteristic function of the set Ω_i and TV denotes the (scalar) Total Variation operator.

As a key feature of our framework we introduce the concept and notation for the *common interface* between subsets Ω_i of the partition $P(\Omega)$ of Ω. This provides a finer decomposition of the **TV** functional defined in (1) into common inter-phases which recover the discontinuity set of the solution weighting our believing of the classification through the *strength* of the boundary (the jump of the solution). Sharp transitions are then proportionally weighted and properly (locally) considered. For this some notations are introduced below.

Let $\Gamma = \bigcup_{i=1}^N \Gamma_i$ be the (discontinuity) *jump set* of the solution $\mathbf{u}$ defined through a partition and let $\Gamma_i = \bigcup_{j=1}^N \Gamma_{ij}$, $i \neq j$, be the inter-phase (perimeter) of class i, i.e the boundary of the class Ω_i. Let moreover Γ_{ij} to denote the *local* inter-phases between the classes Ω_i and Ω_j where the length of the inter-phases is defined as

$$|\Gamma| = \int_\Gamma d\mathcal{H}^{D-1} \quad (5)$$

$$|\Gamma_i| = \int_{\Gamma_i} d\mathcal{H}^{D-1} \quad (6)$$

$$|\Gamma_{ij}| = \int_{\Gamma_{ij}} d\mathcal{H}^{D-1} \quad (7)$$

The perimeter $|\Gamma_i|$ can be computed summing up the contribution of all the common interfaces between the different classes of the partition

$$\text{per}(\Omega_i) = |\Gamma_i| = \sum_{i\neq j} |\Gamma_{ij}| \quad (8)$$

that is the $D-1$-dimensional Hausdorff measure ($|\emptyset| = 0$) of the (reduced) boundary of Ω_i. Obviously $|\Gamma_{ij}| = |\Gamma_{ji}|$.

Given a partition $P(\Omega)$ we define $\bar{\chi} = (\chi_i)$, $i = 1..N$ as the associated vectorial characteristic function. In this setting a vectorial piecewise constant function $\mathbf{u} : \Omega \subset \mathbb{R}^D \to \mathbb{R}^M$ taking exactly N different positive values determined by a N-classes partition $P(\Omega)$ and coefficient matrix $\mathbf{C} \in \mathbb{R}^{M\times N}$, is determined via the vectorial mapping

$$\mathbf{F} : P(\Omega) \times \mathbb{R}^{M\times N} \to [L^\infty(\Omega)]^M \quad (9)$$

in form $\mathbf{u} = \mathbf{F}(\bar{\chi}, \mathbf{C}) \doteq \mathbf{C}\bar{\chi}$. We can evaluate at each pixel $x \in \Omega$ using this mapping in form $\mathbf{F}(x) : \mathbb{R}^N \times \mathbb{R}^{M\times N} \to \mathbb{R}^M$ to have

$$\mathbf{u}(x) = \mathbf{F}(x) = \mathbf{F}(\bar{\chi}(x), \mathbf{C}) \doteq \mathbf{C}\bar{\chi}(x) = \mathbf{C}\bar{e}_i = \bar{c}_i \in \mathbb{R}^M \quad (10)$$

where $i \in \{1..N\}$ is such that $x \in \Omega_i$. In other words this vectorial mapping selects the vector column $\bar{c}_i$ of the matrix $\mathbf{C}$ corresponding to the class Ω_i to which the pixel $x \in \Omega$ belongs. To each pixel corresponds exactly only one column which is a vector of constants which measure the sharpness of the boundary depending on the *jump* of the solution at the local interfaces. In fact there is a different *jump* at any local inter-phase. Notice that the dimension of the partition only depends on the number of classes N into which we classify (segment) the image domain Ω and do not depends on the number of channels M.

As example, for the simplest case of binary segmentation ($N = 2$) of gray level scalar images ($M = 1$) the associated coefficient matrix is simply $\mathbf{C} = \bar{c} = (c_1, c_2) \in \mathbb{R}^{1\times 2}$ and we look for a piecewise constant (binary) function in form

$$u = F(\bar{\chi}, \bar{c}) \doteq \bar{c}\cdot\bar{\chi} = c_1\chi_1 + c_2\chi_2 \quad (11)$$

where χ_1 and $\chi_2 = 1 - \chi_1$ are the characteristic functions of the binary partition $P(\Omega) = \{\Omega_1, \Omega_2\}$.

3 SEGMENTATION MODEL

In this section we propose the generalization of the anisotropic version of the Mumford Shah functional, for multi-class segmentation of multichannel images. Its minimization leads to find a region-based segmentation defined in the image domain Ω by a piecewise constant function $\mathbf{u} = \mathbf{F}(\bar{\chi}, \mathbf{C}) = \mathbf{C}\bar{\chi}$ into exactly N-classes obtained from multichannel data $\mathbf{f}$. For that we propose to minimize the following energy functional:

$$J(\mathbf{C}, \Gamma) = \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} |\mathbf{c}_j - \mathbf{c}_i| \left|\Gamma_{ij}\right| + \frac{1}{2\lambda}\int_\Omega (\mathbf{C}\bar{\chi} - \mathbf{f})^2 dx \quad (12)$$

where vectors $\mathbf{c}_i \in \mathbb{R}^M$, $i = 1..N$ are positive columns vectors of the coefficient matrix $\mathbf{C} \in \mathbb{R}^{M\times N}$, $\mathbf{f} \in [L^\infty(\Omega)]^M$ is the given (noisy) vector valued image and $\lambda \in \mathbb{R}^+$ is a weighting parameter that acts as a trade-off between the fidelity to the data and the regularity of the solution, i.e. the smaller the lambda is, the less regular the solution will be.

Finally, for piecewise constant functions $\mathbf{u} = \mathbf{C}\bar{\chi}$ the last term verifies

$$||\mathbf{u}-\mathbf{f}||^2_{[L^2(\Omega)]^M} = \int_\Omega |\mathbf{u}-\mathbf{f}|^2 dx = \int_\Omega (\mathbf{C}\bar{\chi}-\mathbf{f})^2\, dx \tag{13}$$

and models the L^2 fidelity norm measuring the likelihood of the data $\mathbf{f}$ assuming a Gaussian mixture distribution of the colors.

As example, for the simplest case of binary segmentation ($N = 2$) of gray level scalar images ($M = 1$), we recover the model proposed in Osher and Vese in (Osher and Paragios, 2003), that is a modified version of the Chan-Vese functional (Chan and Vese, 2001) in which the length term is weighted by the jump $|c_2 - c_1|$, resulting in the following functional

$$\begin{aligned} J(\mathbf{C},\Gamma) &= |c_2-c_1|\int_\Gamma d\mathcal{H}^1 + \frac{1}{2\lambda}\int_{\Omega_1}|c_1-f|^2dx \\ &+ \frac{1}{2\lambda}\int_{\Omega_2}|c_2-f|^2dx \end{aligned} \tag{14}$$

where Γ is the unique common interface of the partition $\Gamma = \Gamma_{12} = \Gamma_{21}$.

A fundamental point is that this energy functional can be written as the Rudin-Osher-Fatemi denoising model (Rudin et al., 1992)

$$J(u) = \int_\Omega |Du| + \frac{1}{2\lambda}\int_\Omega |u-f|^2dx \tag{15}$$

when its minimization is constrained to binary piecewise constant functions (Osher and Paragios, 2003) and where we denoted as Du the generalized gradient of u and with $\int_\Omega |Du|$ the total variation of u. Note in $J(\mathbf{C},\Gamma)$ that the minimization relates the discontinuity set of the solution Γ with its length through the TV operator. The term $|c_2 - c_1|$ is a weighting factor defined by the partition in which the optimal constants can be explicitly computed. This is not true anymore when multichannel ($M > 1$) data are considered and a coupling is present. A non-linear system of equations has to be solved.

The key observation, first obtained in Chambolle (Chambolle and Darbon, 2008), when studying the Chan, Essedoglu, Nikolova model (Nikolova et al., 2006) and here generalized to the multi-class multi-channel framework, is that the functionals 12 and 15 are the same when restricted to piecewise constant solutions. In fact we have the following theorem:

Theorem. Let N and λ be positive fixed parameters and let $\Omega \subset \mathbb{R}^D$ be a given bounded open domain representing the image domain. Let $\mathbf{f} \in L^\infty(\Omega)^M$ be a vectorial data function $M \geq 1$ where each scalar component $f_i(x) : \Omega \to \mathbb{R}$ represents a channel. Let $\mathbf{u}$ be a vector valued piecewise constant function such as $\mathbf{u} = \bar{F}(\bar{\chi},\mathbf{C}) = \mathbf{C}\bar{\chi}$ obtained from multichannel data $\mathbf{f}$.

Then, the functional (12) coincides with the vectorial version of the ROF energy functional

$$J(\mathbf{u}) = \int_\Omega |D\mathbf{u}| + \frac{1}{2\lambda}||\mathbf{u}-\mathbf{f}||^2_{L^2(\Omega;\mathbb{R}^M)} \tag{16}$$

Proof. Using the geometric measure theory of Ambrosio, Fusco y Pallara (Ambrosio et al., 2000) as well as the vectorial approach of Bresson, (Bresson and Chan, 2008) the Total variation term $\mathbf{TV}(\mathbf{u}) = \int_\Omega |D\mathbf{u}|$ in the functional (16) can be decomposed in form:

$$\int_\Omega |D\mathbf{u}| = \int_\Omega |\nabla\mathbf{u}|dx + |D^c\mathbf{u}|(\Omega) + \int_\Gamma |\mathbf{u}^+ - \mathbf{u}^-|d\mathcal{H}^{D-1},$$

where $\nabla\mathbf{u}$ denotes the Lebesgue part of the gradient of $\mathbf{u}$, the term $D^c\mathbf{u}$ is the Cantor part of the measure $D\mathbf{u}$ and Γ, as before, is the set of jumping points of $\mathbf{u}$ being $\mathbf{u}^+$, $\mathbf{u}^-$ the jump functions. The two first terms vanish since $\mathbf{u}$ is piecewise constant and $\mathbf{u} \in \mathbf{SBV}(\Omega)$ (the space of Special Bounded Variation functions defined in (Ambrosio et al., 2000)). As a consequence it has no Cantor part and the only contribution to the energy is provided by the jump of the solution at the local inter-phases Γ_{ij}. We then have:

$$\begin{aligned} \int_\Omega |D\mathbf{u}| &= \int_\Gamma |\mathbf{u}^+ - \mathbf{u}^-|d\mathcal{H}^{D-1} = \\ &\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\int_{\Gamma_{ij}}|\mathbf{c}_j-\mathbf{c}_i|d\mathcal{H}^{D-1} = \\ &\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}|\mathbf{c}_j-\mathbf{c}_i|\int_{\Gamma_{ij}}d\mathcal{H}^{D-1} = \\ &\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}|\mathbf{c}_j-\mathbf{c}_i||\Gamma_{ij}| \end{aligned} \tag{17}$$

□

As a straightforward consequence of the above theorem, fixing the number N of classes into which we wish to divide the image domain, the minimization of the energy functional (16) restricted to the set $\mathbf{S}$ of vectorial piecewise constant functions

$$\mathbf{S} = \left\{\mathbf{u} = \mathbf{C}\bar{\chi} \,/\, \mathbf{c}_i \in \mathbb{R}^M, \sum_{i=1}^N \chi_i(x) = 1, a.e. x \in \Omega\right\}$$

with $\mathbf{S} \subset \mathbf{BV}(\Omega;\mathbb{R}^M)$, is equivalent to the minimization of the proposed energy functional (12). In both cases the optimization problem is non-convex because the set of partitions of exactly N-classes, in which we look subsets of $\mathbb{R}^N$, is a non-convex collection (Nikolova et al., 2006). Nevertheless is well known that the ROF model has exactly one solution (a global minimum of the energy functional) when minimized in the larger space of Bounded Variation $\mathbf{BV}(\Omega)$. So, in order to solve a convex problem, we propose to minimize (16), over the larger space $\mathbf{BV}(\Omega;\mathbb{R}^M)$. The solution is now allowed to take values on the continuous interval $[0,1]^M$ and we obtain a solution in set $\mathbf{S}$ using a thresholding step based on a maximum probability criterion. The numerical solution of the ROF model can be obtained using known algorithms such as dual, staggered or primal-dual algorithms (Chambolle, 2004; Garamendi et al., 2013; Chambolle and Pock, 2011). Finally the optimal constants of the selected piecewise solution are computed with a simple fixed point algorithm.

With a view to the resolution of the original non-convex problem, after minimization of (16) the computed solution $\mathbf{u}^* \in \mathbf{BV}(\Omega;\mathbb{R}^M)$, which is the unique global minimum in $\mathbf{BV}(\Omega;\mathbb{R}^M)$ needs to be projected to the set $\mathbf{S}$ by a thresholding step to obtain a piecewise representative $\mathbf{u_{th}} \in \mathbf{S} \subset \mathbf{BV}(\Omega)$ which approximate the global piecewise argmin value $\mathbf{u}$ of $J(\mathbf{u})$ in $\mathbf{S}$: $J(\mathbf{u}) \leq J(\mathbf{u_{pw}})$, $\forall \mathbf{u_{pw}} \in \mathbf{S}$.

As initial example, for scalar images, the thresholding step is done computing a vector $\mathbf{t} \in \mathbb{R}^{N-1}$ of thresholds that generates a partition $P(\Omega)$, defining the N-classes

$$\Omega_i = \{x \in \Omega \mid t_{i-1} \leq \mathbf{u}^*(x) < t_i\}\,.$$

For multichannel problems, the thresholding step can be recasted into the form of a probability clustering problem in the M-dimensional space. Each point $x \in \Omega$ is assigned to a class according to its probability computed through the vector valued intensity level $\mathbf{u}^*(x)$. Considering the histogram of $\mathbf{u}^*(x)$ inside each class a probability density function p_i is generated and each pixel is assigned to a class following the criterion:

$$x \in \Omega_i \quad \Longleftrightarrow \quad p_i(\mathbf{u}^*(x)) \geq p_j(\mathbf{u}^*(x)), \quad \forall j \neq i$$

where p_i and p_j are the probability density functions for the class i and j respectively. Notice that this definition covers the scalar case, where the threshold values are those where $p_i(\mathbf{u}(x)^*) = p_j(\mathbf{u}^*(x))$.

After the thresholding step a partition is generated allowing the computation of the *local inter-phases*. The optimal constants $\mathbf{C}$ can then be computed. Given λ, N, M, $\mathbf{f}$ and a partition of Ω say $P = \{\Omega_i\}_{i=1}^N$ from the thresholding step, we compute the first order necessary optimality conditions for the best constants through the partial derivatives of J defined in (12) with respect to the constants $\mathbf{c}_i \in \mathbb{R}^M$, $i = 1..N$:

$$\frac{\partial J}{\partial \mathbf{c}_i} = \sum_{\substack{j=1 \\ j\neq i}}^{N} \left(\frac{\mathbf{c}_i - \mathbf{c}_j}{|\mathbf{c}_i - \mathbf{c}_j|}\right)|\Gamma_{ij}| + \frac{1}{\lambda}\left(\mathbf{c}_i|\Omega_i| - \int_{\Omega_i} \mathbf{f}dx\right) \tag{18}$$

Imposing the first order necessary condition for optimality we can deduce the following fixed-point algorithm: Let $\mathbf{c}_i^0$ the initial guess (in the experiments we initialize with the mean value), we compute $\mathbf{c}_i^{k+1}$ as

$$\mathbf{c}_i^{k+1} = \frac{\displaystyle\sum_{\substack{j=1 \\ j\neq i}}^{N} \left(\frac{\mathbf{c}_j^k}{|\mathbf{c}_i^k - \mathbf{c}_j^k|}\right)|\Gamma_{ij}| + \frac{1}{\lambda}\int_{\Omega_i} \mathbf{f}dx}{\displaystyle\sum_{\substack{j=1 \\ j\neq i}}^{N} \left(\frac{1}{|\mathbf{c}_i^k - \mathbf{c}_j^k|}\right)|\Gamma_{ij}| + \frac{1}{\lambda}|\Omega_i|} \tag{19}$$

Notice that the new values $\mathbf{c}_j$ can be used as soon as they are computed.

Once we are done with $\mathbf{C}$ the piecewise constant function $\mathbf{u}{\in}\mathbf{S}$ is obtained by $\mathbf{u}{=}\mathbf{C}\bar{\chi}$.

This leads to a relaxation scheme in which for segmenting a vector (M-components) valued D-dimensional image into N clases minimizing the proposed anisotropic Mumford-Shah energy functional (12) it suffices to solve the vectorial ROF model (16) and threshold the solution. The proposed numerical scheme for a given vectorial image $\mathbf{f}$ can be summarized as follows:

1. *Minimize the vectorial ROF energy functional (16) and let* $\mathbf{u}^*$ *be the (unique) minimum of it.*
2. *Threshold the image* $\mathbf{u}^*$ *into N classes to obtain* $\bar{\chi}$*.*
3. *Compute the optimal constants* $\mathbf{C}$
   (a) *Initialize* $\mathbf{C}$*, as the mean value inside the regions determined by* $\bar{\chi}$*.*
   (b) *Use the iterative scheme (19) to obtain* $\mathbf{C}$
4. *Compute a minimum* $\mathbf{u}$ *of (12) as* $\mathbf{u}{=}\mathbf{C}\bar{\chi}$

4 EXPERIMENTS

The model has been tested on LF images downloaded from the Heidelberg Benchmark database for synthetic Light Field images (Wanner et al., 2013a). The structure of LF images allows for a very precisely disparity map computation with a very small cost (Wanner et al., 2013b), so one can assumes that this information is available as an additional feature to the intensity (gray) color, also. The images used in the experiments were generated using the open source software Blender, providing a ground truth for the segmentation. We used the CIE_lab color space such that distance between 3D points on this space corresponds to perceptual color difference (Rousson and Deriche, 2002). Each color in the space CIE_lab is defined by a vector of three components, the first one providing information about luminosity and the second and third ones about the chromaticity. We consider these three values as channels in our model and moreover we add the depth information as a fourth channel. The problem is to segment the 4D dimensional image $\mathbf{f}(x,y,s,t)$ of 4 components (color and depth) into several classes. In figure 1 it can be seen a view for a fixed s and t and the depth math corresponding to the same s and t. The number of classes are 7 for the Horses dataset, 6 for the Buddha and Still Life datasets and 4 for Papillon dataset.

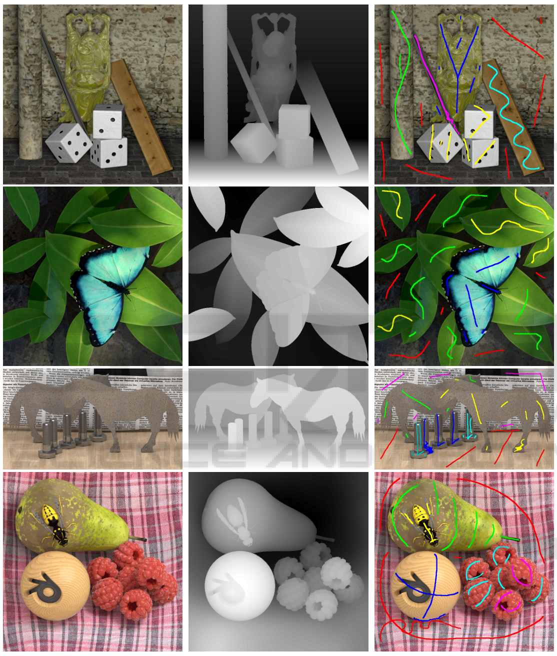

Figure 1: Heidelberg database. In the first column we show the central image ($s=5$, $t=5$) of 81 possible views of the sets Buddha, Papillon, Horses and Still Life. In the second column the depth image is used as a fourth channel. Third column, scribbles used in the projection step.

To evaluate the method we use the percentage of pixels well classified and the Jaccard similarity index, which measures the overlap (agreement) between two binary images X and Y, by taking the ratio between the size of their intersection and the size of their union: $J(X,Y) = |X \cap Y|/|X \cup Y|$. This metric yields values between 0 and 1, where 0 means complete dissimilarity and 1 stands for identical images.

Although it is not required for the model, in this experiment we used scribbles in the threshold step to estimate the probability density function at each class i as a multivariate Gaussian distribution defined by its vectorial mean $\bar{\mu}_i$ and its covariance matrix Σ_i computed from of $\mathbf{u}^*$. Then

Table 1: Percentage of well classified pixels for the different datasets (accuracy ratio).

Buddha	97.73 %
Papillon	98.37 %
Horses	97.39 %
Still Life	98.2 %

$$x \in \Omega_i \iff p_i(\mathbf{u}^*(x)\,|\,\bar{\mu}_i,\Sigma_i) \geq p_j(\mathbf{u}^*(x)\,|\,\bar{\mu}_i,\Sigma_i), \forall j \neq i$$

The scribbles we used for computing the mixture Gaussian probabilities in the threshold step are shown in figure 1. We would like to remark that although the scribbles are drawn on the original image, the vectorial mean $\bar{\mu}_i$ and the covariance matrix Σ_i are estimated from $\mathbf{u}^*$.

The results are resumed in tables 1 and 2 and shown in figure 2 where we show the computed partition (first column of figure 2) and the piecewise constant function $\mathbf{u} = \mathbf{C}\bar{\chi}$ (last column) produced by the method. The parameter λ of the model was chosen empirically with values $\lambda = 0.674$ for Buddha, $\lambda = 0.0554$ for Papillon $\lambda = 0.0765$ for Horses and $\lambda = 0.0554$ for Still Life. In the four tested datasets, the accuracy ratio (table 1) is around 98% and the Jaccard similarity index (table 2) goes from 0.7 to 0.99, depending on the size of the region (class) considered. The accuracy results we obtained are comparable with those in (Wanner et al., 2013b) being the difference less than 1% for Buddha, Papillon and Horses datasets and around 0.3% for Still Life dataset. It is worth to say that we have used less features than in (Wanner et al., 2013b). In fact we use only the color and the given depth while, in (Wanner et al., 2013b), the eigenvalues of Hessian matrix has been used.

5 CONCLUSIONS

We have extended the Anisotropic Mumford-Shah energy functional originally proposed by Osher and Vese (Osher and Paragios, 2003) for the binary segmentation of 2D gray color images to the general case of multi-class segmentation of vectorial images of any dimension. The extension has been done using the formalism for the local or common inter-phases between classes to decompose the boundary of a specific region into curves that correspond to the edge between a region and its neighbours (regions). As a further contribution, we proved the equivalence between the proposed energy functional and the well known

Table 2: Jaccard similarity index. Those datasets marked as '-' means that there is no class with this color.

	Red	Green	Blue	Yellow	Cyan	Pink	Purple	Gray
Buddha	0.96	0.95	-	0.96	0.95	0.94	-	0.81
Papillon	0.99	0.96	0.99	0.94	-	-	-	-
Horses	0.98	0.93	0.73	0.96	0.84	0.98	-	0.88
Still Life	0.99	0.98	0.98	0.88	-	0.88	0.81	-

Figure 2: Results. First column shows the classes detected for the central view ($s = 5, t = 5$). Second Column shows the ground truth. Third Column shows the final piecewise constant computed $\mathbf{u}$ in which the color corresponds to the optimal constant.

Rudin-Osher-Fatemi denoising model when it is minimized over the piecewise constant function set. This relationship allows us to rewrite the original problem of segmentation, which is non-convex, as a convex problem. Finally we show how to project the solutions of the convex problem into the original space of piecewise constant functions.

Convincing results were shown on the Heidelberg Collaboratory group dataset, where the ground truth is provided. Notice that depth information was added as a feature in a new channel. This shows and exemplifies how it is possible to use this general multichannel model for specific image modalities and tasks processing. The accuracy of the results is around 96-98% which is pretty good taking into account that in comparison with other methods for LF image segmentation, we use less features (color and depth).

ACKNOWLEDGEMENTS

The authors acknowledge partial support by TIN2015-70410-C2-1-R (MINECO/FEDER, UE) and by GRC reference 2014 SGR 1301, Generalitat de Catalunya.

REFERENCES

Ambrosio, L., Fusco, N., and Pallara, D. (2000). *Functions of Bounded Variation and free discontinuity problems*. Clarendon Press, Oxford University.

Bresson, X. and Chan, T. (2008). Fast dual minimization of the vectorial total variation norm and applications to color image processing. *Inverse Problems and Imaging*, 2(4):455–484.

Brown, E. S., Chan, T. F., and Bresson, X. (2011). Completely Convex Formulation of the Chan-Vese Image Segmentation Model. *International Journal of Computer Vision*, 98(1):103–121.

Chambolle, A. (2004). An Algorithm for Total Variation Minimization and Applications. *Journal of Mathematical Imaging and Vision*, 20(1/2):89–97.

Chambolle, A. and Darbon, J. (2008). On Total Variation Minimization and Surface Evolution using Parametric Maximum Flows Antonin Chambolle and Jérôme Darbon. *International Journal of Computer Vision*, 84(April):288–307.

Chambolle, A. and Lions, P.-L. (1997). Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(2):167–188.

Chambolle, A. and Pock, T. (2011). A first-order primal-dual algorithm for convex problems with applications to imaging. In *J. Math. Imaging Vision*, volume 40, pages 120–145.

Chan, T. F. and Vese, L. a. (2001). Active contours without edges. *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, 10(2):266–77.

Garamendi, J. F., Gaspar, F. J., Malpica, N., and Schiavi, E. (2013). Box relaxation schemes in staggered discretizations for the dual formulation of total variation minimization. *IEEE transactions on image processing*, 22(5):2030–43.

Klann, E. and Ramlau, R. (2013). Regularization Properties of Mumford–Shah-Type Functionals with Perimeter and Norm Constraints for Linear Ill-Posed Problems. *SIAM Journal on Imaging Sciences*, 6(1):413—-436.

Lippman, R. (1908). La photographie intégrale. *Acadéemie des sciences*, pages 446–551.

Mumford, D. and Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685.

Ng, R. (2006). *Digital light field photography*. PhD thesis.

Ng, R., Levoy, M., Duval, G., Horowitz, M., and Hanrahan, P. (2005). Light Field Photography with a Hand-held Plenoptic Camera. *Informational*, pages 1–11.

Nikolova, M., Esedoglu, S., Chan, T. F., Esedo, S., and Glu, . (2006). Algorithms for Finding Global Minimizers of Image Segmentation and Denoising Models. *SIAM Journal on Applied Mathematics*, 66(5):1632–1648.

Osher, S. and Paragios, N. (2003). *Geometric level set methods in imaging, vision, and graphics*.

Perwass, C. and Wietzke, L. (2010). *www.raytrix.de*.

Pock, T., Chambolle, A., Cremers, D., and Bischof, H. (2009). A convex relaxation approach for computing minimal partitions. In *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2009*, pages 810–817.

Reddy, D., Bai, J., and Ramamoorthi, R. (2013). External mask based depth and light field camera. *Workshop Consumer Depth Cameras for Vision.*

Rousson, M. and Deriche, R. (2002). A variational framework for active and adaptative segmentation of vector valued images. In *Proceedings - Workshop on Motion and Video Computing, MOTION 2002*, pages 56–61. Institute of Electrical and Electronics Engineers Inc.

Rudin, L., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268.

Wanner, S., Meister, S., and Goldluecke, B. (2013a). Datasets and Benchmarks for Densely Sampled 4D Light Fields. In *Vision, Modeling \& Visualization*, pages 225—-226.

Wanner, S., Straehle, C., and Goldluecke, B. (2013b). Globally Consistent Multi-label Assignment on the Ray Space of 4D Light Fields. *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1011–1018.