Cartoon-texture evolution for two-region image segmentation

Laura Antonelli · Valentina De Simone · Marco Viola

Received: 6 October 2021 / Accepted: 11 June 2022 / Published online: 11 July 2022
© The Author(s) 2022

Abstract
Two-region image segmentation is the process of dividing an image into two regions of interest, i.e., the foreground and the background. To this aim, Chan et al. (SIAM J Appl Math 66(5):1632–1648, 2006) designed a model well suited for smooth images. One drawback of this model is that it may produce a bad segmentation when the image contains oscillatory components. Based on a cartoon-texture decomposition of the image to be segmented, we propose a new model that is able to produce an accurate segmentation of images also containing noise or oscillatory information like texture. The novel model leads to a non-smooth constrained optimization problem which we solve by means of the ADMM method. The convergence of the numerical scheme is also proved. Several experiments on smooth, noisy, and textural images show the effectiveness of the proposed model.

Keywords Image segmentation · Cartoon-texture decomposition · Non-smooth optimization · ADMM method

Laura Antonelli, Valentina De Simone and Marco Viola have contributed equally to this work.

Laura Antonelli
laura.antonelli@cnr.it

Valentina De Simone
valentina.desimone@unicampania.it

Marco Viola
marco.viola@unicampania.it

1 Institute for High Performance Computing and Networking (ICAR), National Research Council (CNR), Via Pietro Castellino, 111, 80131 Naples, Italy

2 Department of Mathematics and Physics, University of Campania “Luigi Vanvitelli”, Viale Abramo Lincoln, 5, 81100 Caserta, Italy
1 Introduction

Image segmentation is a fundamental task in image processing and computer vision. It consists in dividing an image into non-overlapping regions of shared features, such as intensity, smoothness, and texture, which are related to the final goal of the segmentation. Thus, the division into regions is not unique, and the image segmentation can be regarded as a strongly ill-posed problem.

Let $f$ be an image defined in a domain $\Omega \subset \mathbb{R}^d$ ($d \geq 2$), segmenting $f$ consists in finding a decomposition of the domain $\Omega$ into a set of non-empty pairwise-disjoint regions $\Omega_i$, $i = 1, \ldots, m$.

A segmentation of $f$ can be expressed through a curve $C^*$ that matches the boundaries of the decomposition of $\Omega$, i.e. $C^* = \bigcup_{i=1}^{m} \partial \Omega_i$ and/or a piecewise-constant function $f^*$ defined on $\Omega$ that approximates $f$.

The research on image segmentation has made several advances in the last decades and various approaches have been developed, including thresholding, region growing, edge detection and variational methods [1–3]. Variational models, based on optimizing energy functionals, have been widely investigated, proving to be very effective on different images; curve evolution [4], anisotropic diffusion [5] and the Mumford-Shah model [6] are good representatives of these methods. Other recent approaches to image segmentation include learning-based methods, which often exploit deep-learning techniques [7–9], watershed [10], random walk methods [11], graph cuts [12, 13], epidemiological models on images [14]. However, in this case, a large amount of data must be available to train learning networks, thus making those approaches impractical in some applications.

Two-region segmentation is here considered, where the domain of the given image $\bar{f}$ is separated in two regions of interest, so $m = 2$ and $\Omega = \Omega_{in} \cup \Omega_{out}$, i.e. $\Omega_{in}$ and $\Omega_{out}$ are the foreground and the background of the image, respectively. Although the choice of $m = 2$ significantly simplifies the segmentation problem, it has a lot of application fields, such as biological and medical imaging, text extraction, compression of screen content and mixed content documents, and can be used as a computational kernel for more complex segmentation tasks [15–20].

A widely-used two-region model was introduced by Chan and Vese in [21] and, together with its variations, is regarded as state of the art in the segmentation community. These models are currently used in medical and astronomical application fields and have lately been associated with machine learning frameworks (see, e.g. [8, 22–26]). The Chan-Vese model is a special case of the most popular Mumford-Shah one [6] restricted to piecewise constant functions. The solution is the best approximation to $\bar{f}$ among all the functions that take only two values, $c_{in}$ and $c_{out}$. As is the case of many variational models for image processing, the model results in a non-convex optimization problem and may have various local minima. Chan et al. [27] propose a convexed relaxation model, here denoted as CEN, which considers the case of $f$ taking values in $[0, 1]$, and sets one of the two regions as

$$\Omega_{in} = \{ x : f(x) > \alpha \} \text{ for a.e. } \alpha \in (0, 1).$$
The CEN model first computes the values $c_{in}$ and $c_{out}$, and then, given $\lambda > 0$, it determines $f$ by solving the convex minimization problem

$$\min_{0 \leq f \leq 1} \int_{\Omega} |\nabla f| \, dx + \lambda \int_{\Omega} ((c_{in} - \bar{f}(x))^2 f(x) + (c_{out} - \bar{f}(x))^2 (1 - f(x))) \, dx.$$  

We note that the aforementioned models assume that each image region is defined as a smooth or constant function. However, images may not be piecewise smooth or flat as a whole, but they may contain some non-smooth regions. In practice, imposing smoothness on such kind of images may lead to a destructive averaging of the image content [28], which can produce an inaccurate segmentation. Exploiting information on the non-smooth structure of an image can help to improve the CEN model to be effective on a larger set than the one of smooth-images as done, e.g. in [29], thanks to the introduction of spatially-varying regularization methods.

In this paper we will design a new model for two-region image segmentation that, starting from a rough cartoon-texture decomposition $\tilde{f} = \bar{u} + \bar{v}$ of the initial image, produces a cartoon-texture-driven decomposition of $\bar{u}$ as $\bar{u} = u + v$ and simultaneously provides a segmentation of $u$. In the new model a Kullback-Leibler divergence term is used to force $v$ to be close to $\bar{v}$, thus allowing it to further extract small-scale oscillatory components from the starting cartoon part $\bar{u}$. Thanks to this additional term, the segmentation process is shown to have improved robustness with respect to noise and texture in the initial image.

The rest of the paper is organized as follows: in Sect. 2 we recall the cartoon-texture decomposition of an image, in Sect. 3 we introduce the proposed model, which results in a non-smooth convex optimization problem, and in Sect. 4 we introduce an ADMM scheme for the problem solution and analyze its convergence. Sect. 5 is devoted to numerical experiments and comparison with the original CEN model and with state-of-the-art models suited for textural image segmentation. Finally, we draw our conclusions in Sect. 6.

### 2 Cartoon-texture decomposition

An image $f$ is usually described as a superposition of two components, i.e.,

$$f = u + v,$$

where $u$ is the geometric component and $v$ is the oscillatory one. The geometric component, commonly referred to as ‘cartoon’, consists of the piecewise-constant parts of an image, including homogeneous regions, contours, and sharp edges. In contrast, the oscillatory component includes the patterns which can be observed in the image, such as texture or noise. Both texture and noise can indeed be seen as repeated patterns of small scale details, with noise being characterized by random and uncorrelated values. The cartoon-texture decomposition of an image plays an important role in computer vision [30], with a wide range of applications to, e.g., image restoration, segmentation, image editing, and remote sensing. It is an underdetermined linear inverse problem with many solutions, usually described by...
variational models able to force the cartoon and the texture into different functional spaces in order to produce the required decomposition.

Following the idea of Meyer [31], the general image decomposition problem can be formulated as

\[
\min_{(u,v) \in X \times Y} g_1(u) + g_2(v) \quad \text{s.t. } u + v = f,
\]

where \(X\) and \(Y\) are suitable function spaces and \(g_1\) and \(g_2\) are functionals that model the cartoon regions and the texture patterns, respectively. Several choices have been proposed in literature for both \(X\), \(Y\) and \(g_1\), \(g_2\) \([32, 33]\). A widely used choice to model the cartoon is \(g_1(u) = \text{TV}(u)\), due to its ability to induce piecewise smooth \(u\) with bounded variations \([34, 35]\). Some alternative approaches impose a sparse representation of the cartoon under a given system, such as wavelet frames \([36]\) or curvelet systems \([37]\). Modeling the texture component is a more complex task, due to the difficulty of conceptualizing mathematical properties able to encompass all the texture types. Many models use the space of oscillatory functions equipped with appropriate norms able to represent textured or oscillatory patterns \([34, 35, 38]\). An alternative approach assumes that, under suitable conditions, textures can be sparsified, i.e., a texture patch can be represented by few atoms in a given dictionary or by specific transforms \([39]\).

Since the existing methods for cartoon-texture decomposition are beyond the scope of this paper, here we simply assume that we are able to obtain a decomposition of the given image:

\[
\tilde{f} = \tilde{u} + \tilde{v},
\]

with the aim of using the different information on the two components to improve the effectiveness of the CEN model. In our experiments we will consider the algorithm described in \([40]\). Figure 1 shows the decomposition produced by one iteration of the algorithm, which results in

\[
\tilde{u}(x) = \omega(\rho_\sigma(x))L_\sigma \ast \tilde{f} + (1 - \omega(\rho_\sigma(x)))\tilde{f}, \quad \tilde{v}(x) = \tilde{f}(x) - \tilde{u}(x),
\]

where \(L_\sigma\) is a low-pass filter, \(\ast\) is the convolution operator, \(\omega : [0, 1] \rightarrow [0, 1]\) is an increasing function that is constant and equal to zero near zero and constant and equal to 1 near 1, and \(\rho_\sigma(x)\) is the relative reduction rate of local TV.

---

Fig. 1 Cartoon-texture decomposition of airplane image after the application of (4)-(5)
\begin{align}
\rho_\sigma(x) &= \frac{LTV_\sigma(u(x)) - LTV_\sigma(L_\sigma \ast f(x))}{LTV_\sigma(f(x))} \in [0, 1] 
\end{align}

with \( LTV_\sigma(f(x)) = (L_\sigma \ast |\nabla f|) \).

We note that the cartoon-texture decomposition produced by (4) is not unique, but it depends on the choice of \( \sigma \) [40]. Anyway, we will show that a rough decomposition is enough for our model, hence there’s no need for an accurate tuning of \( \sigma \).

3 The C-TETRIS model

We here introduce the Cartoon-Texture Evolution for Two-Region Image Segmentation (C-TETRIS) model. As mentioned in the previous sections, starting from the decomposition (3), the main idea behind C-TETRIS is to simultaneously produce the segmentation of \( \bar{u} \) and its cartoon-texture decomposition. In detail, it decomposes \( \bar{u} \) as \( \bar{u} = u + v \), where \( v \) is enforced to be close to \( \bar{v} \), and computes a segmentation of \( u \) by solving the problem

\[
\min_{u \in [0, 1], v} \mathcal{E}_{\text{CEN}}(u, c_{\text{in}}, c_{\text{out}}; \bar{u}) + \mu D_{\text{KL}}(v; \bar{v})
\]

s.t. \( 0 \leq u \leq 1, \quad u + v = \bar{u}, \)

where \( \mathcal{E}_{\text{CEN}} \) represents the objective function of problem (1), \( D_{\text{KL}}(v; \bar{v}) \) denotes the Kullback-Leibler (KL) divergence of \( v \) from \( \bar{v} \), defined as

\[
D_{\text{KL}}(v; \bar{v}) = \int_\Omega v(x) \log \left( \frac{v(x)}{\bar{v}(x)} \right) dx,
\]

where we set

\[
v(x) \log \left( \frac{v(x)}{\bar{v}(x)} \right) = \begin{cases} 0 & v(x) = 0, \\ \infty & \bar{v}(x) = 0, \end{cases}
\]

and \( \mu > 0 \). The KL divergence measures the amount of information lost if \( \bar{v} \) is used to approximate \( v \) and appears in many models of imaging science, where it is usually employed as a fidelity term. Simply speaking, the C-TETRIS model extracts from \( \bar{u} \) the “remaining texture” and produces its best approximation among all the functions that take only two values.

In the following we consider the discrete version of (6). Let

\[
\Omega_{n_x,n_y} = \{(i,j) : 0 \leq i \leq n_x - 1, 0 \leq j \leq n_y - 1\}
\]

be a discretization of \( \Omega \) consisting of an \( n_x \times n_y \) grid of pixels and

\[
|\nabla_x u|_{i,j} = |\delta^+_x u|_{i,j}, \quad |\nabla_y u|_{i,j} = |\delta^+_y u|_{i,j}
\]
where $\delta^+_x$ and $\delta^+_y$ are the forward finite-difference operators in the $x$- and $y$-directions, with unit spacing, and the values $u_{i,j}$ with indices outside $\Omega_{n,n}$ are defined by replication. The discrete version of the (6) leads to the following non-smooth constrained optimization problem:

$$
\min_{u, c_{in}, c_{out}, v} \quad E_{\text{CEN}}(u, c_{in}, c_{out}; \bar{u}) + \mu D_{\text{KL}}(v; \bar{v})
$$

s.t. $0 \leq u \leq 1,$

$$
u + v = \bar{u},
$$

where we denoted by $E_{\text{CEN}}$ the discrete version of $E_{\text{CEN}}$, defined as

$$
E_{\text{CEN}}(u, c_{in}, c_{out}; \bar{u}) = \sum_{i,j} (|\nabla_x u|_{i,j} + |\nabla_y u|_{i,j}) + 
\lambda \sum_{i,j} (u_{i,j}(c_{in} - \bar{u}_{i,j})^2 + (1 - u_{i,j})(c_{out} - \bar{u}_{i,j})^2),
$$

and we denoted with $D_{\text{KL}}$ the discrete version of the Kullback-Leibler divergence $D_{\text{KL}}$, defined as

$$
D_{\text{KL}}(v; \bar{v}) = \sum_{i,j} v_{i,j} \log \left(\frac{v_{i,j}}{\bar{v}_{i,j}}\right).
$$

It is worth noting that the first term in $E_{\text{CEN}}$ corresponds to the discrete Total Variation (TV) of the image $u$. We here opted for the use of a modified version of the TV functional, in which the $\ell_2$ norm is replaced by the $\ell_1$ one (as proposed in [41]), since in the case of image restoration it is known to produce sharper piece-wise constant images. Nevertheless, a preliminary comparison between the models equipped with the $\ell_1$ and the $\ell_2$ version, respectively, showed no difference in terms of segmentation quality.

4 Minimizing the C-TETRIS model

We here focus on the solution of the minimization problem in (8). One can observe that, although the problem is in general nonconvex, it becomes convex when either the pair $(c_{in}, c_{out})$ or the pair $(u, v)$ are fixed. Suppose, for the moment, that the values of $c_{in}, c_{out}$ have been determined and consider the minimization problem in $u$ and $v$ only, which can be written as

$$
\min_{u, v} \sum_{i,j} (|\nabla_x u|_{i,j} + |\nabla_y u|_{i,j}) + \lambda r^\top u + \mu D_{\text{KL}}(v; \bar{v})
$$

s.t. $0 \leq u \leq 1,$

$$
u + v = \bar{u},
$$

where we defined, for each $(i, j),$
Cartoon-texture evolution for two-region image segmentation

Problem (9) is a non-smooth convex optimization problem subject to linear and bound constraints which we propose to solve by the Alternate Directions Method of Multipliers (ADMM) [42]. To this aim, we reformulate problem (9) as

\[
\min_{u, d_x, d_y, v} \|d_x\|_1 + \|d_y\|_1 + \lambda r^T u + \mu D_{KL}(v; \bar{v})
\]

s.t. \( d_x = \nabla_x u, \)
\( d_y = \nabla_y u, \)
\( u + v = \bar{u}, \)
\( 0 \leq u \leq 1. \)

(10)

Starting from (10), it is straightforward to check that the objective function and the constraints of the problem can be split in two blocks. Indeed, by introducing the variable \( z = [d_x^T, d_y^T, v^T]^T, \) one can further reformulate (10) as

\[
\min_{u, z} F(u) + G(z)
\]

s.t. \( Hu - z = b, \)

(11)

where we defined

\[
F(u) = \lambda r^T u + \chi_{[0,1]}(u), \quad G(z) = \|d_x\|_1 + \|d_y\|_1 + \mu D_{KL}(v; \bar{v}),
\]

\[
H = \left[\nabla_x^T, \nabla_y^T, -I\right]^T, \text{ and } b = [0, 0, -\bar{u}^T]^T,
\]

and we used \( \chi_{[0,1]}(u) \) to indicate the characteristic function of the hypercube \([0, 1]^{n_x \times n_y}. \)

Consider the Lagrangian and the augmented Lagrangian functions associated with problem (11), defined respectively as

\[
\mathcal{L}(u, z, \xi) = F(u) + G(z) + \xi^T (Hu - z - b),
\]

\[
\mathcal{L}_A(u, z, \xi; \rho) = F(u) + G(z) + \xi^T (Hu - z - b) + 0.5\rho \|Hu - z - b\|_2^2,
\]

where \( \rho > 0, \) and \( \xi \) is a vector of Lagrange multipliers.

Starting from given estimates \( u^0, z^0, \) and \( \xi^0, \) at each iteration \( k \) ADMM updates the estimates as

\[
u^{k+1} = \arg\min_u \mathcal{L}_A(u, z^k, \xi^k; \rho),
\]

\[
z^{k+1} = \arg\min_z \mathcal{L}_A(u^{k+1}, z, \xi^k; \rho),
\]

\[
\xi^{k+1} = \xi^k + \rho (Hu^{k+1} - z^{k+1}).
\]

(12)

Since \( F(u) \) and \( G(z) \) in (11) are closed, proper and convex, and \( H \) has full rank, the convergence of ADMM can be proved by exploiting the classical result from [43], which we report in the following.
Theorem 1 Consider problem (11) where \( F(u) \) and \( G(z) \) are closed, proper and convex functions and \( H \) has full rank. Consider the summable sequences \( \{ \varepsilon_k \} \), \( \{ \nu_k \} \subset \mathbb{R}_+ \) and let

\[
\begin{align*}
\left\| u^{k+1} - \arg\min_u L_A(u, z^k, \xi^k; \rho) \right\| & \leq \varepsilon_k, \\
\left\| z^{k+1} - \arg\min_z L_A(u^{k+1}, z, \xi^k; \rho) \right\| & \leq \nu_k, \\
\xi^{k+1} & = \xi^k + \rho(H u^{k+1} - z^{k+1}) .
\end{align*}
\]

If there exists a saddle point \((u^*, z^*, \xi^*)\) of \( L(u, z, \xi) \), then \( u^k \to u^* \), \( z^k \to z^* \) and \( \xi^k \to \xi^* \). If such saddle point does not exist, then at least one of the sequences \( \{ z^k \} \) or \( \{ \xi^k \} \) is unbounded.

Theorem 1 guarantees the convergence of the ADMM scheme even if the subproblems are solved inexactly, provided that the inexactness of the solution can be controlled.

So far we have been concerned with the solution of problem (9) when the values of \( c_{in} \) and \( c_{out} \) are known in advance which, however, is not the case in practice. By following the example of [27], we adopt a two-step scheme in which we alternate updates of \( u \) and \( z \), determining the shape of the two regions, and updates of \( c_{in} \) and \( c_{out} \). Observe that, by fixing \( u = u^k \) and \( z = z^k \), the restriction of problem (8) to \( c_{in} \) and \( c_{out} \) can be written as the unconstrained convex quadratic optimization problem

\[
\min_{c_{in}, c_{out}} \sum_{i,j} \left( u_{ij}^k (c_{in} - \bar{u}_{ij})^2 + (1 - u_{ij}^k) (c_{out} - \bar{u}_{ij})^2 \right).
\]

Hence, we propose to update the values of \( c_{in} \) and \( c_{out} \) after each ADMM step by taking the exact minimizer of problem (13), i.e., by setting

\[
\begin{align*}
\xi_{in}^k & = \frac{\sum_{i,j} u_{ij}^k \bar{u}_{ij}}{\sum_{i,j} u_{ij}^k}, & \text{and} & \quad \xi_{out}^k & = \frac{\sum_{i,j} (1 - u_{ij}^k) \bar{u}_{ij}}{\sum_{i,j} (1 - u_{ij}^k)} .
\end{align*}
\]

It is worth pointing out that such a modification alters the original ADMM scheme making it an inexact alternate minimization scheme for the problem in \( u, z, c_{in}, \) and \( c_{out} \). Nevertheless, as also shown for the original CEN model, the experiments carried out in this work show that in all the cases under analysis the values of \( c_{in} \) and \( c_{out} \) stagnate after the first few iterations, thus recovering in practice the convergence properties shown for the case of fixed \( c_{in} \) and \( c_{out} \).

4.1 Solving the ADMM subproblems

We will now focus on how the subproblems in (12) can be solved in practice. First, by expliciting the form of the augmented Lagrangian functions, we can rewrite the ADMM scheme as

\[
\sum_{i,j} u_{ij}^k \bar{u}_{ij} + \rho \left( H u_{ij}^k - z_{ij}^k \right) .
\]
\[ u^{k+1} = \underset{\xi \in \partial_1}{\arg\min} \lambda \, r^T u + (\xi^k)^T (H u - \xi^k - b) + \frac{\rho}{2} \left\| H u - \xi^k - b \right\|_2^2, \]
\[ \xi^{k+1} = \arg\min_{\xi} G(z) + (\xi^k)^T (H u^{k+1} - z - b) + \frac{\rho}{2} \left\| H u^{k+1} - z - b \right\|_2^2, \]
\[ \zeta^{k+1} = \xi^k + \rho (H u^{k+1} - \zeta^{k+1} - b). \]

It is straightforward to check that the minimization problem over \( z \) can be split into three independent minimization problems, respectively on \( d_x, d_y, \) and \( v, \) leading to the following scheme

\[ u^{k+1} = \underset{\xi \in \partial_1}{\arg\min} \lambda \, r^T u + (\xi^k)^T (H u - \xi^k - b) + \frac{\rho}{2} \left\| H u - \xi^k \right\|_2^2, \]
\[ d_x^{k+1} = \underset{d_x}{\arg\min} \left\| d_x \right\|_1 + (\xi^k)^T (\nabla_x u^{k+1} - d_x) + \frac{\rho}{2} \left\| \nabla_x u^{k+1} - d_x \right\|_2^2, \]
\[ d_y^{k+1} = \underset{d_y}{\arg\min} \left\| d_y \right\|_1 + (\xi^k)^T (\nabla_y u^{k+1} - d_y) + \frac{\rho}{2} \left\| \nabla_y u^{k+1} - d_y \right\|_2^2, \]
\[ v^{k+1} = \arg\min_{v} \mu D_{KL}(v; \overline{v}) + (\xi^k)^T (-u^{k+1} - v + \overline{v}) + \frac{\rho}{2} \left\| u^{k+1} + v - \overline{v} \right\|_2^2, \]
\[ \zeta^{k+1} = \xi^k + \rho (H u^{k+1} - \zeta^{k+1} - b), \]

where we split the Lagrange multipliers vector \( \xi \) as \( \xi = [\xi_x^T, \xi_y^T, \xi_v^T]^T. \) The scheme presented in (15) can be further simplified by exploiting the linearity of the constraints \( H u - z = b, \) as suggested in [44]. In detail, by introducing the vectors \( b_x^k = \frac{\xi^k}{\rho}, b_y^k = \frac{\xi^k}{\rho}, \) and \( b_v^k = -\frac{\xi_v^k}{\rho} - \overline{u}, \) one can rewrite (15) equivalently as

\[ u^{k+1} = \underset{\xi \in \partial_1}{\arg\min} \lambda \, r^T u + \frac{\rho}{2} \left\| \nabla_x u - d_x + b_x^k \right\|_2^2 \]
\[ + \frac{\rho}{2} \left\| \nabla_y u - d_y + b_y^k \right\|_2^2 + \frac{\rho}{2} \left\| u + v + b_v^k \right\|_2^2, \]
\[ d_x^{k+1} = \underset{d_x}{\arg\min} \left\| d_x \right\|_1 + \frac{\rho}{2} \left\| \nabla_x u^{k+1} - d_x + b_x^k \right\|_2^2, \]
\[ d_y^{k+1} = \underset{d_y}{\arg\min} \left\| d_y \right\|_1 + \frac{\rho}{2} \left\| \nabla_y u^{k+1} - d_y + b_y^k \right\|_2^2, \]
\[ v^{k+1} = \arg\min_{v} \mu D_{KL}(v; \overline{v}) + \frac{\rho}{2} \left\| u^{k+1} + v + b_v^k \right\|_2^2, \]
Problem (16) is a strongly convex bound-constrained quadratic optimization problem. To obtain an approximate solution $u^{k+1}$, by following [29, 45], we consider the optimality conditions of the unconstrained version of the problem, i.e., the solution to the linear system

\[
(-\Delta + I)u = -\frac{\lambda_r}{\rho} + (\nabla^T_x (b_x^k - d_x^k)) + (\nabla^T_y (b_y^k - d_y^k)) + (b_k^v - v^k),
\]

where $\Delta$ represents the finite-difference discretization of the Laplacian. We first solve the system by Gauss-Seidel method and then project the solution in $[0, 1]^{n_x \times n_y}$.

As regards the updates in (17)-(19), one has to note that they are proximal operators [46, 47] of closed proper and convex functions. In detail, the proximal operator in (17) and (18) can be computed in closed form by means of the well-known soft-thresholding operator, defined as

\[
[S(x, \gamma)]_{i,j} = \text{sign}(x_{i,j}) \cdot \max (|x_{i,j}| - \gamma, 0).
\]

Finally, the proximal operator in (19) can be computed as

\[
[\text{prox}_{\gamma D_{KL}(x, \tilde{x})}(x)]_{i,j} = \gamma W(\gamma^{-1} x_{i,j} e^{\gamma^{-1} x_{i,j} - \tilde{x}_{i,j}}),
\]

where $W(x)$ is the Lambert $W$ function satisfying $W(y)e^{W(y)} = y$ which, although not available in closed form, can be approximated with high precision.

5 Numerical experiments

In this section, we test the effectiveness of C-TETRIS in producing two-region segmentation on various image sets. The first set contains three pairs of real-life images with corresponding ground truth coming from the database [48]: man is a smooth image whereas flowerbed and stone show an object foreground on a textured background. The second set consists of four images available from the Berkeley database [49] which are in general considered to be smooth: the real-life images airplane and squirrel, and the medical images brain and ultrasound. The third set of images consists of noisy versions of the famous cameraman image from MIT Image Library\(^1\) which we use to test the robustness of the C-TETRIS model with respect to the noise. The fourth and last set of images consists of three textural images: tiger and bear, taken from [49], and spiral, taken from [21]. We here provide some further details on the numerical experiments. The C-TETRIS algorithm was implemented in

\[
b_x^{k+1} = b_x^k + \nabla_x u_x^{k+1} - d_x^{k+1},
b_y^{k+1} = b_y^k + \nabla_y u_y^{k+1} - d_y^{k+1},
b_v^{k+1} = b_v^k + u_v^{k+1} + v^{k+1} - \bar{u}.
\]

\(^1\) https://libguides.mit.edu/findingimages.
MATLAB using the Image Processing Toolbox, where the cartoon-texture decomposition was initially performed by one iteration of the algorithm described in [40], using a Gaussian filter with $\sigma = 2$ as $L_\sigma$, and the following function $\omega$ [40]:

$$
\omega(x) = \begin{cases} 
0, & x \leq l_1, \\
(x - l_1)/(l_2 - l_1), & l_1 < x < l_2, \\
1, & x \geq l_2,
\end{cases}
$$

(21)

where the weights $l_1$ and $l_2$ have been set to 0.25 and 0.5, respectively. We would like to remark that extensive testing showed that the accuracy of the produced segmentation is only slightly influenced by the variation of the Gaussian smoothing parameter, $\sigma$, or by the number of steps performed to obtain the cartoon-texture decomposition. Among the several available implementations of CEN we chose the one proposed by the authors of [45]. Although the code is written in C programming language, a MEX interface is available for testing in MATLAB. This implementation is based on split Bregman iterations with the following stopping criterion:

$$
|\text{diff}^k - \text{diff}^{k-1}| \leq \text{tol} \quad \text{and} \quad k > \text{maxit},
$$

(22)

where

$$
\text{diff}^k = \frac{\text{sd}(f^k)}{\text{sd}(f^k) \cdot \text{sd}(f^{k-1})}, \quad \text{sd}(f^k) = \sum_{i,j} (f_{i,j}^k - f_{i,j}^{k-1})^2,
$$

tol is a given tolerance and maxit is the maximum number of SB iterations. In order to make a fair comparison, all the algorithms presented in the next section use the stopping criterion (22), where we set maxit = 50 and tol = $10^{-6}$ (tol = $10^{-8}$ for the noisy images). The parameter $\lambda$ in (1) and in (9), has a scaling role and was set according to the level of required details in the segmentation. In particular, in each test for CEN model we used the value proposed by the authors in the available code, which we indicate as $\lambda_{CEN}$, based on this empirical rule: $\lambda_{CEN} = 10^a$ with $a \in \{-1, 0, 1\}$ from larger to smaller regularization/smoothing. To balance the presence of the KL term, for C-TETRIS we perform a grid search and select a parameter $\lambda$ with a variation of at most 5% from $\lambda_{CEN}$. The parameter $\mu$ was set as $\mu = 10^c$ with $c \in \{-2, -1, 0\}$. Finally, the Bregman parameter $\rho$ was set to 1.

Before proceeding with the experiments on the four image sets described above, we show an example of the functioning of the proposed model. We consider an image for each of the four sets and report in Fig. 2 the starting cartoon-texture decomposition and the components $u$ and $v$ after the first ADMM iteration, at an intermediate iteration and at the last iteration. We note that, as the ADMM advances, the remaining texture is progressively subtracted from the cartoon, allowing a clearer distinction of background and foreground.

\footnote{\url{http://htmlpreview.github.io/?https://github.com/xbresson/old_codes/blob/master/codes.html}.}
5.1 Results on ground truth images

First of all, in order to assess the accuracy of the C-TETRIS segmentation model, a comparison with ground truth data is presented in Fig. 3. The quality of the produced segmentations confirms the greater ability of C-TETRIS with respect to CEN in separating foreground objects from the background, especially on the flowerbed and stone images, where textured background is present. Furthermore, quantitative analysis measuring the similarity between the segmented images and the corresponding ground truth is given in Table 1. The segmentation errors have been evaluated using four traditional measures\(^3\). The Rand Index (RI) \(^{50}\) counts the fraction of pairs of pixels whose labellings are consistent between the computed segmentation and the ground truth, the Global Consistency Error (GCE) \(^{51}\) measures the distance between two segmentations assuming that one segmentation must be a refinement of the other, the Variation of Information (VI) \(^{52}\) computes the distance between two segmentations as the average conditional entropy of one segmentation given the other, and the Boundary Displacement Error (BDE) \(^{53}\) computes the average boundary pixels displacement error between two segmented images. As we can note in Table 1, the segmentations produced by C-TETRIS, have smaller values of GCE, VI, and BDE, than the ones produced by CEN, as well as they present the highest values of the RI measures, showing a greater consistency with the corresponding ground truth in the partitioning of foreground objects from the background.

5.2 Results on smooth images

In Fig. 4, we show a comparison between C-TETRIS and CEN on the segmentation of the set of smooth images.

For the sake of completeness we report also the segmentation results produced by CEN on the cartoon of the images. In general, the segmentations produced by C-TETRIS are comparable with or better than the ones produced by CEN. The segmentation of airplane shows the great effectiveness of the proposed model to separate accurately a non-uniform background from the object, due to the ability of C-TETRIS to remove the remaining texture in the cartoon, as showed in Fig. 2. We note that in general there are no significant differences in the quality of the segmentation results between CEN applied to the original image and CEN applied to the cartoon. However, in the case of ultrasound the segmentation on the cartoon produces unreliable result, due to the loss of contrast introduced by decomposition.

\(^3\) The software used for the four measures of segmentation error is available at: https://people.eecs.berkeley.edu/yang/software/lossysegmentation/.

\(^4\) The error of one boundary pixel is defined as its distance from the closest pixel in the other boundary image.
### Cartoon-texture evolution for two-region image segmentation

| Cartoon-texture decomposition | ADMM iterations |
|-----------------------------|-----------------|
|                             | first | intermediate | last |
| Cartoon                     | ![Image](image1.png) | ![Image](image2.png) | ![Image](image3.png) |
| Texture                     | ![Image](image4.png) | ![Image](image5.png) | ![Image](image6.png) |
| Cartoon                     | ![Image](image7.png) | ![Image](image8.png) | ![Image](image9.png) |
| Texture                     | ![Image](image10.png) | ![Image](image11.png) | ![Image](image12.png) |
| Cartoon                     | ![Image](image13.png) | ![Image](image14.png) | ![Image](image15.png) |
| Texture                     | ![Image](image16.png) | ![Image](image17.png) | ![Image](image18.png) |

---

© Springer
In Table 2, two global metrics are listed to measure the contrast between the given image and its cartoon. In particular we used

\[ m_1 = f_{\text{max}} - f_{\text{min}} \]

and the Michelson formula [54]:

\[ m_2 = (f_{\text{max}} - f_{\text{mean}})/(f_{\text{max}} + f_{\text{mean}}) \]

where \( f_{\text{max}} \), \( f_{\text{min}} \) and \( f_{\text{mean}} \) are the maximum, the minimum and the mean value respectively of the given image intensity. We can note that the cartoon part of ultrasound shows the largest reduction of the both metrics with respect to the original image.
5.3 Results on noisy images

In Fig. 5 a comparison between C-TETRIS and CEN on the set of noisy images is shown. The *cameraman* image was corrupted by different source of noise using the MATLAB `imnoise` function. In detail: the option ‘gaussian’ was used with different values for the standard deviation to obtain images affected by Gaussian noise with signal-to-noise ratio (SNR) equal to 20 and 15, respectively; by rescaling the pixels of the original image and using the option ‘poisson’ we obtained images affected by Poisson noise with SNR equal to 35 and 30, respectively; finally, the option ‘salt & pepper’ was used to create images affected by impulsive noise on 5% and 15% of the pixels. We note that C-TETRIS is more accurate in separating background and foreground, especially when the noise level increases. In this case, indeed, the noise is recognised as texture part and classified as foreground.

![Fig. 4 Segmentations of smooth images by CEN and C-TETRIS. The results of the segmentation produced by CEN on the cartoon part of the image are also shown.](image)

| original image | cartoon | CEN | CEN on cartoon | C-TETRIS |
|----------------|---------|-----|----------------|----------|
| ![Original Image](image) | ![Cartoon](image) | ![CEN](image) | ![CEN on Cartoon](image) | ![C-TETRIS](image) |
| *airplane* | | | |
| ![Original Image](image) | ![Cartoon](image) | ![CEN](image) | ![CEN on Cartoon](image) | ![C-TETRIS](image) |
| *squirrel* | | | |
| ![Original Image](image) | ![Cartoon](image) | ![CEN](image) | ![CEN on Cartoon](image) | ![C-TETRIS](image) |
| *brain* | | | |
| ![Original Image](image) | ![Cartoon](image) | ![CEN](image) | ![CEN on Cartoon](image) | ![C-TETRIS](image) |
| *ultrasound* | | | |
5.4 Results on textural images

Here we analyze the results of the C-TETRIS model on images containing textural components which require a two-region segmentation. We compared C-TETRIS with the Spatially Adaptive Regularization (SpAReg) model [29], which modifies the CEN model as follows:

$$\min_f \sum_{i,j} \left( |\nabla_x f|_{i,j} + |\nabla_y f|_{i,j} + \lambda_{i,j} (r^T f)_{i,j} \right)$$

s.t. $0 \leq f \leq 1$

where each entry of the matrix $\Lambda = (\lambda_{i,j})$ weighs the pixel $(i, j)$ according to local texture information as follows:

$$\lambda_{i,j} = \max \left\{ \frac{\lambda_{\min}}{\lambda_{\max}}, 1 - (\rho_\sigma)_{i,j} \right\} \lambda_{\max}.$$  \hfill (24)

$(\rho_\sigma)_{i,j}$ was defined applying the Eq. (5) to the given image $\tilde{f}$, and $0 < \lambda_{\min} < \lambda_{\max} < \infty$ is a suitable range to drive the level of regularization, depending on the image to be segmented. In all the tests we set $\lambda_{\min} \leq \lambda_{\text{CEN}} < \lambda_{\max}$. We also include in the comparison a well-known segmentation model designed for textural images [55], that we denote as HTB. While C-TETRIS and SpAReg, being based on the original CEN model, classify foreground and background as regions with different intensities, the HTB model classifies them as regions with different textural components. In detail, it finds a contour that maximizes the KL distance between the probability density functions of the regions inside and outside the evolving (closed) active contour, which is aimed at separating textural objects of interest from the background. The feature used to characterize the texture is based on principal curvatures $\chi$ of the intensity image considered as a 2-D manifold embedded in $\mathbb{R}^3$. In detail, the objective function of the HTB model is

$$\text{KL}(p_{\text{in}}, p_{\text{out}}) = \sum_{i,j} ((p_{\text{in}})_{i,j} - (p_{\text{out}})_{i,j}) (\log (p_{\text{in}})_{i,j} - \log (p_{\text{out}})_{i,j}),$$

where $p_{\text{in}}, p_{\text{out}}$ are the probability distribution of the texture feature $\chi$ in $\Omega_{\text{in}}$ and $\Omega_{\text{out}}$, respectively, assuming a Gaussian distribution. We consider the implementation of HTB model provided in [45].

Figure 6 compares the segmentations produced by C-TETRIS with the ones produced by SpAReg and HTB, respectively. Firstly, we note that C-TETRIS outperforms...
| Noise Type       | SNR   | CEN       | C-TETRIS |
|------------------|-------|-----------|----------|
| Gaussian noise   | 20    | ![Image](image1.png) | ![Image](image2.png) |
| Gaussian noise   | 15    | ![Image](image1.png) | ![Image](image2.png) |
| Poissonian noise | 35    | ![Image](image1.png) | ![Image](image2.png) |
| Poissonian noise | 30    | ![Image](image1.png) | ![Image](image2.png) |
| Salt & Pepper    | 5%    | ![Image](image1.png) | ![Image](image2.png) |
| Salt & Pepper    | 15%   | ![Image](image1.png) | ![Image](image2.png) |
both SpAReg and HTB on tiger and spiral, where the textural object region was well identified and separated from the background. On the bear test image, C-TETRIS seems to identify the main object better than SpAReg; however, it mistakenly includes in the foreground region some parts of the background below the bear. Both models are outperformed by HTB, which is the only model able to include the upper part of the image in the background region. In our opinion, the inaccurate result produced by the
other two models is mainly due to the inhomogeneity of the background intensity that adverses its separation from the foreground region.

6 Conclusion

In this paper, a new model named Cartoon-Texture Evolution for Two-Region Image Segmentation (C-TETRIS) is proposed. C-TETRIS intends to improve the CEN model, which is specifically designed for smooth images, to produce good results on a wider set of images. Indeed, starting from a rough cartoon-texture decomposition of the image to be segmented, \( \tilde{f} = \tilde{u} + \tilde{v} \), where \( \tilde{u} \) and \( \tilde{v} \) describe the cartoon and the texture components respectively, C-TETRIS is able to simultaneously produce a decomposition of \( \tilde{u} \) as \( \tilde{u} = u + v \), where \( v \) is enforced to be close to \( \tilde{v} \) and the best approximation among all the functions that take only two values of \( u \). This is realized by combining the CEN model on \( u \) and a Kullback-Leibler divergence of \( v \) from \( \tilde{v} \). The proposed model leads to a non-smooth constrained optimization problem solved by means of the ADMM method, for which a convergence result is provided. Numerical experiments show that, as the ADMM advances, C-TETRIS progressively subtracts from \( \tilde{u} \) the remaining texture, leading to a clearer distinction between background and foreground of the image. The experiments show that the proposed model is able to produce accurate two-region segmentation, comparable with or better than the one produced by state-of-the-art segmentation models, for several images also corrupted by noise or containing textural components. Furthermore, C-TETRIS seems to be independent of the type and level of noise. Future work will deal with the extension of the proposed combination of cartoon-texture decomposition and KL divergence term to more advanced image segmentation models.

Acknowledgements This work was partially supported by Istituto Nazionale di Alu Matematica - Gruppo Nazionale per il Calcolo Scientifico (INdAM-GNCS), by the Italian Ministry of University and Research under Grant No. PON03PE_00060_5, and by the VALERE Program of the University of Campania “L. Vanvitelli”. We would like to thank Simona Sada (ICAR-CNR) for her technical support.
Data availability  The authors confirm that all data generated or analysed during this study are included in this article. The repositories of image tests are also reported.

Declarations

Conflict of interest  The authors have no financial or proprietary interests in any material discussed in this article.

Open Access  This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Zhang, J., Chen, K., Yu, B., Gould, D.A.: A local information based variational model for selective image segmentation. Inverse Problems Imaging 8, 293–320 (2014). https://doi.org/10.3934/ipi.2014.8.293
2. Gwet, D.L.L., Otesteanu, M., Libouga, I.O., Bitjoka, L., Popa, G.D.: A review on image segmentation techniques and performance measures. Int. J. Inf. Control Comput. Sci. 12,0(12) (2018). https://doi.org/10.5281/zenodo.2579976
3. Antonelli, L., De Simone, V., di Serafino, D.: A view of computational models for image segmentation (2021). arXiv:2102.05533v3
4. Sethian, J.: Level Set Methods and Fast Marching Methods. Cambridge University Press, Cambridge (1999)
5. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. 12(7), 629–639 (1990). https://doi.org/10.1109/34.56205
6. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Commun. Pure Appl. Math. 42(5), 577–685 (1989). https://doi.org/10.1002/cpa.3160420503
7. Minaee, S., Boykov, Y.Y., Porikli, F., Plaza, A.J., Kehtarnavaz, N., Terzopoulos, D.: Image segmentation using deep learning: a survey. IEEE Trans. Pattern Anal. Mach. Intell. 1–1 (2021). https://doi.org/10.1109/TPAMI.2021.3059968
8. Rashmi, R., Prasad, K., Udupa, C.B.K.: Multi-channel Chan-Vese model for unsupervised segmentation of nuclei from breast histopathological images. Comput. Biol. Med. 136, 104651 (2021). https://doi.org/10.1016/j.compbiomed.2021.104651
9. Liu, Y., Duan, Y., Zeng, T.: Learning multi-level structural information for small organ segmentation. Signal Process. 193, 108418 (2022). https://doi.org/10.1016/j.sigpro.2021.108418
10. Challa, A., Danda, S., Sagar, B.S.D., Najman, L.: Watersheds for semi-supervised classification. IEEE Signal Process. Lett. 26(5), 720–724 (2019). https://doi.org/10.1109/LSP.2019.2905155
11. Aletti, G., Benfenati, A., Naldi, G.: A semiautomatic multi-label color image segmentation coupling dirichlet problem and colour distances. J. Imaging 7(10) (2021). https://doi.org/10.3390/jimaging7100208
12. Niazi, M., Rahbar, K., Sheikhan, M., Khademi, M.: Entropy-based kernel graph cut for textural image region segmentation. Multimedia Tools Appl. 81(9), 13003–13023 (2022)
13. He, K., Wang, D., Wang, B., Feng, B., Li, C.: Foreground extraction combining graph cut and histogram shape analysis. IEEE Access 7, 176248–176256 (2019). https://doi.org/10.1109/ACCESS.2019.2957504
14. Bampis, C.G., Maragos, P., Bovik, A.C.: Graph-driven diffusion and random walk schemes for image segmentation. IEEE Trans. Image Process. 26(1), 35–50 (2017). https://doi.org/10.1109/TIP.2016.2621663
15. Wang Z., Q.J. Zhu L.: Roi extraction in dermatosis images using a method of Chan-Vese segmentation based on saliency detection. In: Kidwelly, P. (ed.) Mobile, Ubiquitous, and Intelligent Computing. Lecture Notes in Electrical Engineering, vol. 274, pp. 197–203 (2004).
16. Zhang, J., Kasturi, R.: Extraction of text objects in video documents: Recent progress. In: 2008 The Eighth IAPR International Workshop on Document Analysis Systems, pp. 5–17 (2008). https://doi.org/10.1109/DAS.2008.49
17. Minaei, S., Wang, Y.: Screen content image segmentation using sparse decomposition and total variation minimization. In: 2016 IEEE International Conference on Image Processing (ICIP), pp. 3882–3886 (2016). https://doi.org/10.1109/ICIP.2016.7533087
18. Minaei, S., Fotouhi, M., Khalaj, B.H.: A geometric approach for fully automatic chromosome segmentation (2011). arxiv:1112.4164 [cs.CV]
19. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE Trans. Pattern Anal. Mach. Intell. 23, 1222–1239 (2001). https://doi.org/10.1109/34.969114
20. Gregoretti, F., Cesarini, E., Lanzuolo, C., Oliva, G., Antonelli, L.: An automatic segmentation method combining an active contour model and a classification technique for detecting Polycomb-group proteins in high-throughput microscopy images. Methods Mol. Biol. 1480, 181–197 (2016). https://doi.org/10.1007/978-1-4939-6380-5_16
21. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. Image Process. 10(2), 266–277 (2001). https://doi.org/10.1007/s10851-019-00893-0
22. Nguyen, K.L., Tekitek, M.M., Delachartre, P., Berthier, M.: Multiple relaxation time lattice Boltzmann models for multigrid phase-field segmentation of tumors in 3D ultrasound images. SIAM J. Imaging Sci. 12(3), 1324–1346 (2019). https://doi.org/10.1137/18M123462X
23. Roberts, M., Spencer, J.: Reformulation for selective image segmentation. Math. Imaging Vis. 61, 1173–1196 (2019). https://doi.org/10.1007/s10851-019-00893-0
24. Babu, K.R., Nagajaneyulu, P.V., Prasad, K.S.: Performance analysis of cnn fusion based brain tumour detection using chan-vese and level set segmentation algorithms. Int. J. Signal Imaging Syst. Eng. 12(1–2), 62–70 (2020). https://doi.org/10.1504/IJISISE.2020.113571
25. Yousefirizi, F., Rahim, A.: Consolidating deep learning framework with active contour model for improved PET-CT segmentation. J. Nucl. Med. 62(supplement 1), 1415–1415 (2021) https://jnmm.jnmjournals.org/content
26. Zhao, W., Wang, W., Feng, X., Han, Y.: A new variational method for selective segmentation of medical images. Signal Process. 190, 108292 (2022). https://doi.org/10.1016/j.sigpro.2021.108292
27. Chan, T.F., Esedoğlu, S., Nikolova, M.: Algorithms for finding global minimizers of image segmentation and denoising models. SIAM J. Appl. Math. 66(5), 1632–1648 (2006). https://doi.org/10.1137/040645186
28. Wang, J., Chan, K.L.: Incorporating patch subspace model in Mumford-Shah type active contours. IEEE Trans. Image Process. 22(11), 4473–4485 (2013). https://doi.org/10.1109/TIP.2013.2274385
29. Antonelli, L., De Simone, V., di Serafino, D.: Spatially adaptive regularization in image segmentation. Algorithms 13(226), 100390 (2020). https://doi.org/10.3390/a13090226
30. Xu, R., Xu, Y., Quan, Y.: Structure-texture image decomposition using discriminative patch recurrence. IEEE Trans. Image Process. 30, 1542–1555 (2021). https://doi.org/10.1109/TIP.2020.3043665
31. Meyer, Y.: Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures. American Mathematical Society, New York (2001)
32. Le Guen, V.: Cartoon + Texture image decomposition by the TV-L1 model. Image Process. On Line 4, 204–219 (2014). https://doi.org/10.5201/ipol.2014.103
33. Aujol, J., Gilboa, G., Chan, T., O., S.: Structure-texture image decomposition-modeling, algorithms, and parameter selection. Int. J. Comput. Vis. 67, 111–136 (2006). https://doi.org/10.1007/s11263-006-4331-z
34. Osher, S., Solé, A., Vese, L.: Image decomposition and restoration using total variation minimization and the H1. Multiscale Model. Simul. 1(3), 349–370 (2003). https://doi.org/10.1137/S154034590241624
35. Aujol, J., Chambolle, A.: Dual norms and image decomposition models. Int. J. Comput. Vis. 63, 85–104 (2005). https://doi.org/10.1007/s11263-005-4948-3
36. Fadili, M.J., Starck, J.-L., Bobin, J., Moudden, Y.: Image decomposition and separation using sparse representations: an overview. Proc. IEEE 98(6), 983–994 (2010). https://doi.org/10.1109/JPROC.2009.2024776

37. Ono, S., Miyata, T., Yamada, I., Yamaoka, K.: Image recovery by decomposition with component-wise regularization. IEICE Trans. Fundam. Electron. Commun. Comput. Sci. 95-A(12), 2470–2478 (2012). https://doi.org/10.1587/transfun.E95.A.2470

38. Duval, V., Aujol, J.F., Vese, L.A.: Mathematical modeling of textures: application to color image decomposition with a projected gradient algorithm. J. Math. Imaging Vis. 37(3), 232–248 (2010). https://doi.org/10.1007/s10851-010-0203-9

39. Xu, R., Quan, Y., Xu, Y.: Image cartoon-texture decomposition using isotropic patch recurrence (2018). arXiv:1811.04208

40. Buades, A., Le, T.M., Morel, J., Vese, L.A.: Fast cartoon + texture image filters. IEEE Trans. Image Process. 19(8), 1978–1986 (2010). https://doi.org/10.1109/TIP.2010.2046605

41. Esedo ğ lu, S., Osher, S.J.: Decomposition of images by the anisotropic Rudin-Osher-Fatemi model. Commun. Pure Appl. Math. 57(12), 1609–1626 (2004). https://doi.org/10.1002/cpa.20045

42. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn. 3(1), 1–122 (2011). https://doi.org/10.1561/2200000016

43. Eckstein, J., Bertsekas, D.P.: On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. Math. Program. 55(1), 293–318 (1992). https://doi.org/10.1007/BF01581204

44. Goldstein, T., Osher, S.: The split Bregman method for L1-regularized problems. SIAM J. Imaging Sci. 2(2), 323–343 (2009)

45. Goldstein, T., Bresson, X., Osher, S.: Geometric applications of the split Bregman method: segmentation and surface reconstruction. J. Sci. Comput. 45(1–3), 272–293 (2010). https://doi.org/10.1007/s10915-009-9331-z

46. Parikh, N., Boyd, S.: Proximal algorithms. Found. Trends Optim. 1(3), 127–239 (2014). https://doi.org/10.1561/2400000003

47. Beck, A.: First-order Methods in Optimization/Amir Beck, Tel-Aviv University, Tel-Aviv. Israel. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, Mathematical Optimization Society, Philadelphia (2017)

48. Gulshan, V., Rother, C., Criminisi, A., Blake, A., Zisserman, A.: Geodesic star convexity for interactive image segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition (2010)

49. Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 33(5), 898–916 (2011). https://doi.org/10.1109/TPAMI.2010.161

50. Rand, W.M.: Objective criteria for the evaluation of clustering methods. J. Ame. Stat. Assoc. 66(336), 846–850 (1971)

51. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proceedings of Eighth IEEE International Conference on Computer Vision, 2001 (ICCV 2001), vol. 2, pp. 416–423 (2001). https://doi.org/10.1109/ICCV.2001.937655

52. Meilă, M.: Comparing clusterings by the variation of information. In: Schölkopf, B., Warmuth, M.K. (eds.) Learning Theory and Kernel Machines, pp. 173–187. Springer, Berlin (2003)

53. Freixenet, J., Muñoz, X., Raba, D., Martí, J., Cufí, X.: Yet another survey on image segmentation: Region and boundary information integration. In: Proceedings of the 7th European Conference on Computer Vision-Part III. ECCV '02, pp. 408–422. Springer, Berlin (2002)

54. Michelson, A.A.: Studies in Optics. The University of Chicago Press, Chicago (1927)

55. Houhou, N., Thiran, J.-P., Bresson, X.: Fast texture segmentation based on semi-local region descriptor and active contour. Numer. Math. 2(4), 445–468 (2009). https://doi.org/10.4208/nmtma.2009.m9007s

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.