Cartoon Extraction Based on Anisotropic Image Classification

Benjamin Berkels, Martin Burger*, Marc Droske, Oliver Nemitz, Martin Rumpf

Institut für Numerische Simulation,
Rheinische Friedrich-Wilhelms-Universität Bonn, Nussallee 15, 53115 Bonn, Germany
*Institut für Numerische und Angewandte Mathematik,
Westfälische Wilhelms-Universität Münster, Einsteinstraße 62, 48149 Münster, Germany

Email: {benjamin.berkels, marc.droske}@ins.uni-bonn.de,
{oliver.nemitz, martin.rumpf}@ins.uni-bonn.de,
martin.burger@jku.at

Abstract

We propose a new approach for the extraction of cartoons from 2D aerial images. Particularly in city areas, these images are mainly characterized by rectangular geometries of locally varying orientation. The presented method is based on a joint classification of the shape orientation and a rectangular structure preserving prior in the restoration of image shapes. Mathematically, an anisotropic area functional encodes the preference for edges aligned to locally preferable directions and a higher order regularization term ensures a smooth variation of these directions. The concrete model is an anisotropic version of the Rudin-Osher-Fatemi (ROF) scheme with a position dependent anisotropy. Given the knowledge of the anisotropic image structure, the restoration process can be significantly improved, in particular the round-off effect of the ROF model can be reduced. By combining the extraction of the anisotropy with the denoising method in a joint variational approach, we obtain a suitable classification method, in which a tedious direct anisotropy estimation can be avoided. The implementation is based on a finite element discretization and an energy minimization via a step-size-controlled Newton method. Instructive synthetic images are considered to demonstrate the methods performance and the approach is applied to aerial images as a prototype application.

1 Introduction

Image restoration and the decomposition of images into a cartoon (representation of the actual shapes) and a texture are nowadays extensively studied imaging tools [12, 13, 23]. An already classical approach is the Rudin-Osher-Fatemi model [20] and variants of this method [14, 7, 27]. These methods are well-suited to restore sharp edge contours. But at corners formed by edges they come along with a significant rounding artifact. In partic-
ular for images characterized by rectangular shapes this hampers the identification of structures and destroys a proper cartoon representation. Concepts for anisotropic variational approaches, such as those presented in [8, 17], and the anisotropic variant of the Rudin-Osher-Fatemi model by Esedoglu and Osher [11] point out a suitable modification, which we are developing further here. As a prototype application we consider aerial images of city zones, the technique is however also suitable for other types of images with similar morphologies. Hence, we obtain the following problem set-up: We assume, that the given possibly noisy and locally destroyed image contains primarily structures with straight edges and corners with right angles. Furthermore, we assume that the orientation of these structures varies in space. In particular we do not fix an orientation a priori. The aim is now to extract a cartoon representation of image shapes, while preserving or even enhancing edges and sharp corners. This extraction can also be regarded as an image restoration technique. Let us briefly review the state of the art. There already exists a large variety of approaches to feature preserving image restoration, as for example nonlinear diffusion methods [25] and the Rudin-Osher-Fatemi (ROF) model. The ROF-model is the fundamental basis for a wide range of image decomposition models, which separate the input signal into a cartoon part $u$ and a texture part $v$ (c. f. for instance [2, 3]). Inspired by Y. Meyer’s idea [16] to characterize textures by functions with a bounded $\| \cdot \|_\infty$ norm, i. e., the dual norm of the $BV$-norm, the key ingredient for decomposition models is the study of qualitative properties for different norms in which the fidelity $u - u_0$ is measured.

Several methods were introduced to approximate this problem by related problems, that are computationally feasible and yield qualitatively similar results [19, 12]. Recently, decomposition models based on a $L^1$-fidelity have attracted much attention due to their desirable scale decomposition properties [14, 7, 27].

It is well known, that the restored image of the ROF-model often suffers from a significant loss of contrast. An iterative procedure based on Bregman iterations leads to a sequence of decreasing scale, converging back to the original image, where the loss of contrast is compensated already in very early stages of the iteration [22, 18, 6, 5]. In the continuous setting, this process can be interpreted as an inverse scale space. The focus of this paper is the study of the classical ROF model with an anisotropic $BV$-norm. Based on the theory of anisotropically aligned microstructures [26, 1, 24], the concept of so-called Wulff shapes has been used to denoise surfaces [8] and images [11] using estimated a-priori information about the shape of the object to be denoised. In [17] the anisotropic structure of blood vessels has been determined in a first estimation step and subsequently deblurred by “cigar-like” Wulff shapes with locally volume-preserving mean-curvature flow.

In this paper, we propose a joint classification of image anisotropies and a discontinuity-preserving denoising model based on an anisotropic variant of the approach by Rudin, Osher and Fatemi [20]. It is well-known, that this model tends to round-off non-smooth parts of the boundary of the shapes to be restored. This motivated Esedoglu and Osher [11, 4] to consider the minimization of

$$E_\gamma[u] := \int_\Omega \gamma(\nabla u) \, dx + \int_\Omega \lambda (u_0 - u)^2 \, dx$$

which already generalized the original ROF model, in which $\gamma(\nabla u) = |\nabla u|$. Here, $\gamma$ encodes the anisotropic area. In this paper we further generalize this approach to tackle real applications in which the orientation of the anisotropy usually varies in space.

The joint estimation of feature anisotropies and the corresponding image cartoon approach one obtains a convenient method of reconstructing lost shape information, e. g., partially destroyed edges or corners.

## 2 A Variational Approach

Let us first state the main goals of the model. For the restored image $u$ it is desirable to preserve the functional features of the signal such as co-dimension one discontinuities and at the same time geometric features, such as the shape of the level sets of the original signal, which its co-dimension two vertex characteristics. For the non-texture part of images it can often be assumed that in many areas the anisotropic structure does not vary strongly in space. Based on this assumption, we aim not only at the preservation of geometric features but also at
the restoration in smaller areas, where strong corruption of the morphology can still be recovered by the shape information in the vicinity.

Thus we consider anisotropy functions $\gamma$ from a suitable restricted space of admissible anisotropies which are parameterized over the position. Previous models for anisotropic image or surface denoising typically rely on estimated shape classification \cite{9, 17}, which is used to specify a given anisotropy a priori. The main disadvantage of these approaches is the fact that they all need a separate classification. This two-step method is either fairly expensive or inaccurate, and hence we want to solve both problems simultaneously. Thus we consider a joint classification and smoothing approach encoded in one energy functional.

As described in \cite{11}, an anisotropic version of the total variation semi-norm on $L^1_{loc}(\Omega)$ is given by

$$\| v \|_{BV_\gamma} := \sup_{g \in C^1_c(\Omega; \mathbb{R}^d)} \int_\Omega v \text{div} g \, dx.$$  

It is crucial to note that $\| \cdot \|_{BV_\gamma}$ is topologically equivalent to $\| \cdot \|_{BV}$ on $L^1_{loc}(\mathbb{R}^d)$. For the ease of presentation we use instead the widespread formal notation $\int_\Omega \gamma(\nabla v) \, dx$. Here $\gamma$ is assumed to be positive and one-homogeneous.

The Franck diagram $F_\gamma$ and the corresponding Wulff shape $W_\gamma$ are defined by

$$F_\gamma := \left\{ z \in \mathbb{R}^d : \gamma(z) = 1 \right\},$$  

$$W_\gamma := \left\{ z \in \mathbb{R}^d : \gamma^*(z) := \sup_{n \in S^{d-1}} \frac{\langle z, n \rangle}{\gamma(n)} = 1 \right\}.$$

We essentially exploit the well-known fact, that the Wulff shape has the optimal geometry, if normal directions in $S^{d-1}$ are measured in terms of $\gamma$.

We eventually want to formulate a variational problem over admissible anisotropies $\gamma$ and images $u$, however the differentiation w.r.t. a general space of anisotropies $\gamma$ is not straightforward. We aim at posing the problem over a restricted set of anisotropies – well suited in particular for our application on aerial images – that yields a convenient differentiable structure and provides enough freedom for typical configurations in images with accentuated edges, as in aerial images of city zones. Let us first assume a fixed preferred alignment of edges, namely horizontal and vertical structures. In this case, the anisotropy would be expressed by

$$\gamma(z) = \left| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} z \right| = |z_1| + |z_2|,$$

which is the 1-norm with the unit square as the respective Wulff shape.
In order to yield an alignment for arbitrary right angles we have to rotate the Wulff shape. Consequently, we introduce a free parameter $\alpha$, which represents the angle of the rotation.

In this paper we confine on the background of our application to a rotated $l_1$-norm as a Wulff shape. Thus we are interested in structures with right angles and an orientation given by an angle $\alpha$. Therefore we introduce a vector $p = p(\alpha)$ which is collinear to the base line of the Wulff shape and a vector $q = q(\alpha)$ which is orthogonal to it (see Figure 2):

$$p(\alpha) := \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}, \quad q(\alpha) := \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}. $$

We denote by $M(\alpha) := \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$ the orthogonal matrix for a rotation by $-\alpha$. This leads to the anisotropic energy

$$E_\gamma[u, \alpha] := \frac{\lambda}{s} \int_\Omega |u - u_0|^s \, dx + \int_\Omega |M(\alpha) \nabla u|_1 \, dx,$$

where $1 \leq s < \infty$. Typical choices are $s = 2$ or $s = 1$. Furthermore, we have to control the variation of the free parameter $\alpha$. Recall, that the focus of the proposed restoration method is the treatment of corners, which are co-dimension two objects. In case of a simple Dirichlet type regularization, we would observe a lack of regularity from the Sobolev embedding theorem. Thus, we consider a higher order regularization energy, namely:

$$E_\alpha[\alpha] := \int_\Omega \frac{1}{2} \left( \mu_1 |\nabla \alpha|^2 + \mu_2 |\Delta \alpha|^2 \right) \, dx.$$

Now, the total energy to be minimized is given by

$$E[u, \alpha] = E_\gamma[u, \alpha] + E_\alpha[\alpha].$$

The first term of the energy $E_\gamma$ ensures, that the evolution does not differ too much from the original image, the second term is the rotated $1$-norm taking care of the preferred shapes. Furthermore, the energy $E_\alpha$ limits the spatial variations of the orientation parameter $\alpha$.

Let us have a closer look at the second term:

$$|M(\alpha) \nabla u|_1 = |p \cdot \nabla u| + |q \cdot \nabla u|.$$
in such a way, that the coordinate system spanned by \( p \) and \( q \) is aligned to the image edges. At corners, we will switch then from an alignment of \( p \) to an alignment of \( q \) or vice versa (cf. Figure 3). This alignment requirement together with the regularity of \( \alpha \) ensured by \( E_{\alpha} \) will lead to a smoothing of curved structures as well.

![Figure 4: The energy attains a minimum if \( p \) is collinear or orthogonal to \( \nabla u \).](image)

**3 Implementation**

**Regularization of the functional.** First of all we have to regularize the corner singularities in the anisotropy \( \gamma \). Thus, we replace the \( l_1 \)-norm by its regularized version \( |x|_{1,\delta} = |x_1|_\delta + |x_2|_\delta \) with \( |x|_\delta = \sqrt{|x|^2 + \delta^2} \) and obtain for the corresponding regularized energy

\[
E_{\delta} = \int_\Omega \frac{\lambda}{2} |u - u_0|^2 + |M(\alpha)\nabla u|_{1,\delta}^2 + \frac{1}{2} (\mu_1 |\nabla \alpha|^2 + \mu_2 |\Delta \alpha|^2) \, dx.
\]

As discussed in [10] the regularization parameter \( \delta \) has to be coupled with the grid size \( h \) of the computational grid. \( \delta \) is usually chosen proportional to \( h \).

**Postprocessing by Bregman iteration.** The coefficients have to be choosen such that we balance the fidelity energy and the anisotropic length functional. This has to be done in such a way that the sharpening of edges is indeed energetically more preferable than just keeping destroyed edges in their initial shape, thereby reducing the fidelity term. This balance with a rather small coefficient in front of the fidelity term leads to a significant loss of contrast. To compensate for this loss, we proceed iteratively for with the minimization problems resulting from the following Bregman iteration [18]:

\[
(u^{k+1}, \alpha^{k+1}) := \arg \min_{(u, \alpha)} \left\{ \int_\Omega |M(\alpha)\nabla v|_{1,\delta} \, dx + \frac{\lambda}{2} \int_\Omega (u_0 + v^k - u)^2 \, dx + E_{\alpha}[\alpha] \right\},
\]

\[
v^{k+1} := v^k + u_0 - u^{k+1}
\]

where \( v^0 := 0 \). As can be seen in Figure 5, we retain high contrast already in the early stage of the iteration. More precisely, for an image consisting of a cylinder, the Bregman iteration yields the true solution already after the first Bregman iteration, given that \( \lambda \) is large enough, due to the compensation-effect of adding the noise back onto the signal.

The Bregman iteration for ROF-type models does also have a geometric interpretation, namely the successive approximation of the normals of the input image. Employing Bregman iterations using an anisotropic \( BV \)-norm, we obtain even more precise shape approximation in the early stage of the iteration. However, we also expect the sequence of iterations to converge back to the original signal as in the isotropic case.

**Minimization Algorithm.** We employ an alternating minimization algorithm to compute the minimum of the regularized energy in each Bregman iteration, i. e., we alternately compute the optimal \( u \) for fixed \( \alpha \) and vice versa. This means we search for \( u^{k+1} \in BV(\Omega) \) and \( \alpha^{k+1} \in H^{2,2}(\Omega) \) such that

\[
\delta_\alpha E^k_{\delta}[u, \alpha] = 0 \quad \text{for fixed } \alpha \quad \text{and} \quad \delta_u E^k_{\delta}[u, \alpha] = 0 \quad \text{for fixed } u.
\]

Here \( \delta_\alpha E^k_{\delta}[u, \alpha] \) and \( \delta_u E^k_{\delta}[u, \alpha] \) denote the first variations of \( E^k_{\delta}[u, \alpha] \) (cf. Appendix), the energy to be minimized in the \( k \)-th Bregman iteration, which differs from \( E_{\delta} \) only by a different function \( u_0 \) in the fidelity term.

For this sake we use Newton’s method to find the root of a function \( F \) – in our case the gradient of the energy \( E_{\delta} \) with respect to \( u \) and \( \alpha \), respectively. For a given start-value \( u^0 = u_0 \) we have to solve
the linear system of equations
\[ F'(u^{k,i})(u^{k,i+1} - u^{k,i}) = -\tau F'(u^{k,i}) \]
in each iteration of our minimization algorithm. Thus, we also need the second variations of \( E_\delta[u, \alpha] \) (cf. Appendix). The step-size \( \tau \) of Newton’s method is controlled by the Armijo-rule (cf. [15]).

**Finite Element Discretization.** We consider a uniform rectangular mesh \( \mathcal{C} \) covering the whole image domain \( \Omega \) and use a standard bilinear Lagrange finite element space. The integrals \( \int_\Omega vw \, dx \) and \( \int_\Omega \nabla \xi \cdot \nabla \vartheta \, dx \) result in the usual mass \((M)\) and stiffness \((L)\) matrices. Since we deal with piecewise affine finite elements, we introduce a second unknown \( w = -\Delta \alpha \) and write \( \int_\Omega \Delta \alpha \Delta \vartheta = \int_\Omega \nabla w \cdot \nabla \vartheta \), which leads to the matrix \( LM^{-1}L \). We use a numerical Gauss quadrature scheme of order three (cf. [21]) to compute the integrals in the corresponding matrices and vectors. The inverse of the second variation is computed applying a conjugate gradient descent preconditioned by SSOR.

**4 Discussion & Outlook**

We have demonstrated the benefits of an anisotropic Rudin-Osher-Fatemi-model for the cartoon extraction from images whose shapes are primarily rectangular with spatially varying orientation. Degrees of freedom are the local orientation and the restored image intensity. They are computed via a minimization of a joint variational classification and cartoon extraction approach. An anisotropic shape prior reflects the preference for rectangular shapes, whereas a higher order regularization energy for the orientation controls its spatial variation. As a prototype application we have considered aerial images of city zones with predominantly right-angled structures (see the Figures on page 8 and the colorplate Figure 6, which both show the original image, the cartoon and the estimated angular structure). Furthermore, we have shown that this approach can also be used to recover blurred corners. Obviously, natural images can reveal far more complex structures. Corner singularities with opening angle different from \( \frac{\pi}{2} \) have to be tackled via a further generalized model - a focus for future studies. Furthermore, besides the improvement of anisotropic cartoon extraction, also the identification of the image texture component can benefit from an anisotropic variational treatment.

**Acknowledgment.** This project is partially supported by the Deutsche Forschungsgemeinschaft (SPP 611), the Austrian Fonds zur Förderung der Wissenschaftlichen Forschung (SFB F 013 / 08), and the Johann Radon Institute for Computational and Applied Mathematics (Austrian Academy of Sciences). We would also like to thank Gerhard Dziuk for inspirings discussions on anisotropic energies. The data is courtesy Aerowest GmbH and Vexcel Corporation.

**5 Appendix**

In this section, to enable a reimplementation we give for the readers convenience a complete list of the first and second variations of our energy in the case of \( s = 2 \). To simplify notation we introduce the following abbreviations: \( \partial_{\alpha} u = \nabla u \cdot \alpha = \nabla u \cdot (\cos \alpha, \sin \alpha)^T \) and \( \partial_{q} u = \nabla u \cdot q = \nabla u \cdot (-\sin \alpha, \cos \alpha)^T \) (see also Figures 2 and 4). Using this we get the following first and second variations with respect to \( u \):

\[ \delta_u E_\delta[u, \alpha](v) = \lambda \int_\Omega (u - u_0)v \, dx \]
\[ + \int_\Omega \frac{\partial_{\alpha} u}{|\partial_{\alpha} u|_\delta} \partial_{\alpha} v + \frac{\partial_{q} u}{|\partial_{q} u|_\delta} \partial_{q} v \, dx, \]

\[ \delta_u \delta_u E_\delta[u, \alpha](v, w) = \lambda \int_\Omega vw \, dx \]
\[ + \int_\Omega \left( \frac{1}{|\partial_{\alpha} u|_\delta} (\partial_{\alpha} u)^2 (p \otimes p) \nabla v \cdot \nabla w \, dx \right) \]
\[ + \int_\Omega \left( \frac{1}{|\partial_{q} u|_\delta} (\partial_{q} u)^2 (q \otimes q) \nabla v \cdot \nabla w \, dx. \right) \]

Here \( (a \otimes a) = (a_i a_j)_{i,j}. \) The first and second variations with respect to \( \alpha \) turn out to be:

\[ \delta_\alpha E_\delta[u, \alpha](\vartheta) = \mu_1 \int_\Omega \nabla \alpha \cdot \nabla \vartheta \, dx + \mu_2 \int_\Omega \Delta \alpha \Delta \vartheta \, dx \]
\[ + \int_\Omega \frac{\partial_{\alpha} u}{|\partial_{\alpha} u|_\delta} \partial_{\alpha} u - \frac{\partial_{q} u}{|\partial_{q} u|_\delta} \partial_{q} u \, dx, \]
\[ \delta_\alpha \delta_\alpha E_\delta[u, \alpha](\vartheta, \xi) \]
\[ = \mu_1 \int_\Omega \nabla \vartheta \cdot \nabla \xi \, dx + \mu_2 \int_\Omega \Delta \vartheta \Delta \xi \, dx \]
\[ + \int_\Omega \left( \frac{(\partial^{(\alpha)} u)^2}{|\partial p^{(\alpha)} u|^3} \right) \partial \xi \, dx \]
\[ - \int_\Omega \left( \frac{(\partial q^{(\alpha)} u)^2}{|\partial q^{(\alpha)} u|^3} \right) \partial \xi \, dx \]
\[ + \int_\Omega \left( \frac{(\partial q^{(\alpha)} u)^2}{|\partial q^{(\alpha)} u|^3} \right) \partial \xi \, dx \]
\[ - \int_\Omega \left( \frac{(\partial q^{(\alpha)} u)^2}{|\partial q^{(\alpha)} u|^3} \right) \partial \xi \, dx. \]

References


Figure 6: Application of our method to 3 different aerial images of city areas. Left: original image. Middle: result of our algorithm. Right: color-coded angle of the anisotropic structure of the image.