Fast Convex Relaxations using Graph Discretizations

Jonas Geiping[1]
jonas.geiping@uni-siegen.de
Fjedor Gaede[2]
fjedor.gaede@wwu.de
Hartmut Bauermeister[1]
hartmut.bauermeister@uni-siegen.de
Michael Moeller[1]
michael.moellerg@uni-siegen.de

[1] Department of Electrical Engineering and Computer Science
University of Siegen
Siegen, Germany

[2] Institut fur Analysis und Numerik,
Westfälische Wilhelms-Universität Münster,
Münster, Germany

Abstract

Matching and partitioning problems are fundamentals of computer vision applications with examples in multilabel segmentation, stereo estimation and optical-flow computation. These tasks can be posed as non-convex energy minimization problems and solved near-globally optimal by recent convex lifting approaches. Yet, applying these techniques comes with a significant computational effort, reducing their feasibility in practical applications. We discuss spatial discretization of continuous partitioning problems into a graph structure, generalizing discretization onto a Cartesian grid. This setup allows us to faithfully work on super-pixel graphs constructed by SLIC or Cut-Pursuit, massively decreasing the computational effort for lifted partitioning problems compared to a Cartesian grid, while optimal energy values remain similar: The global matching is still solved near-globally optimal. We discuss this methodology in detail and show examples in multi-label segmentation by minimal partitions and stereo estimation, where we demonstrate that the proposed graph discretization can reduce runtime as well as memory consumption of convex relaxations of matching problems by up to a factor of 10.

1 Introduction

Matching problems and the closely inter-related minimal partitioning problems are low-level computer vision tasks that build a backbone for a variety of applications such as multi-label segmentation [4], stereo estimation [17, 34] and optical flow estimation [7, 16]. However, posing these problems as energy minimization problems leads to non-convex objectives that are difficult to solve. In the last years, it has been demonstrated that functional lifting techniques are very well suited for solving these non-convex minimization problems via convex relaxations in a higher-dimensional space. Unfortunately, despite the precise solutions these methods provide, they incur significant costs in memory and computational effort, due to the high dimensionality of the lifted problem.

Consider the continuous minimal partitions problem, also referred to as piecewise-constant Mumford-Shah problem [26], which builds the basis of the aforementioned computer vision

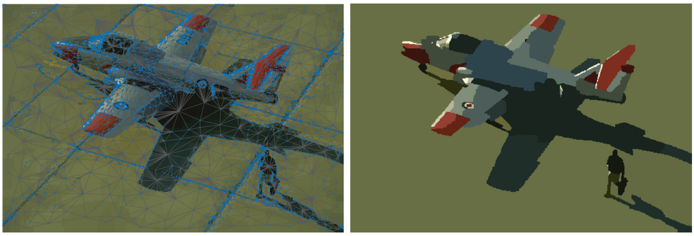

Figure 1: Graph discretization on the left and segmentation by convex relaxation with 32 labels, computed on the graph structure on the right. The segmentation is 3.03% less optimal compared to relaxation on a full Cartesian grid, but requires only around 4.8% of computation time, the original problem has 4.9 mio. variables, and the reduced problem only 120k.

applications,

$$\min_{\{P_k\}_{k=1}^L} \sum_{k=1}^{L} \int_{P_k} -f_k(x) + \frac{1}{2}\operatorname{Per}(\Omega, P_k). \tag{1}$$

Given a set of L potential functions f_k we are looking for a partitioning of the set Ω into L non-overlapping partitions $\{P_k\}_{k=1}^L$, i.e. $P_k \cap P_l = \emptyset$ if $k \neq l$ and $\bigcup_{k=1}^L P_k = \Omega$ [9]. Without the regularizing perimeter term, the solution is given by $\arg\max_k f_k(x)$ for every $x \in \Omega$, but under inclusion of the second term in (1), the perimeter of each partition $\operatorname{Per}(\Omega, P_k)$ is penalized, leading to spatially coherent solutions, where every point is "matched" to a partition, but the global surface energy stays minimal.

Minimal partitions problems are abundant in imaging. In the discrete setting, they directly relate to the Potts model [33] and MRFs [4, 5]. This variability hinges on the choice of potential functions f_k. For multi-label segmentation, for example, we can consider each f_k to give the prior likelihood that the point x should be assigned label k, with the likelihood being computed by model-based approaches [12] or returned as the output of a neural network [11]. On the other hand, considering every partition with label k to encode a displacement of k pixels between two stereo images, the same framework can be used to solve stereo matching problems [45]. Further immediate examples include optical flow[16], scene flow and multiview reconstruction [18].

To be able to solve Eq. (1), we minimize over a set of partitions. This discrete matching problem at the heart of the minimal partitions problem is in general NP-hard. In practice, one of the most powerful approaches is the solution of a suitable convex relaxation of the original minimal partitions problem. To do so, the minimization over partitions is first replaced by minimization over their characteristic functions $u_k : \Omega \to \{0,1\}$, satisfying $\sum_{k=1}^L u_k(x) = 1 \ \forall x \in \Omega$,

$$\min_{\{u_k\}_{k=1}^L} \sum_{k=1}^{L} \int_{\Omega} -f_k(x)u_k(x) + \int_{\Omega} |Du_k|, \tag{2}$$

where we have equivalently replaced the perimeter of a set by the total variation of its characteristic function, see [9]. In a next step, the functions u_k are relaxed to take values in the full interval $[0,1]$, either directly [24, 44], or by jointly deriving tighter convex reformulations of the regularizer [9]. Relaxation approaches often lead to near-optimal solutions with high fidelity [30, 39], yet the computational effort amounts to solving a non-smooth, non-strongly convex optimization problem over all functions u_k, each of which is usually discretized to be as large as the given potential functions. Especially when k is large, memory costs quickly become impractical for computer vision applications.

In this work we hence consider strategies to remediate the computational costs of functional lifting techniques, without majorly impeding their global matching capabilities. As illustrated in Figure 1, we propose to first discretize the problem on a precomputed graph structure instead of a Cartesian grid, and then solve the convex relaxation on the graph. This strategy leads to significant computational advantages while sacrificing almost no accuracy in terms of the energy of the final solution.

2 Related Work

The minimal partitions problem is a prime example of energy minimization methods in computer vision [8, 26, 36], which have found widespread use. In the context of convex relaxations of these partition problems, there have been works in as diverse applications such as stereo estimation [30, 32, 34, 35, 43], optical flow [39, 40], segmentation [9, 24, 44, 46],and optimization on manifolds [25], with algorithmic improvements such as [38].

Previous work discusses the optimal discretization of the continuous label dimension [23, 27, 28], reducing the computational effort of functional lifting in a variety of applications. In our work, we discuss an orthogonal direction of research, as we are discussing compact discretizations of the image space.

The choice of efficient discretization of the input image data is directly related to superpixel approaches, e.g. [1, 42]. Their general idea is to generically reduce the computational complexity of any (pixel-based) numerical algorithm, by locally grouping pixels of similar color to larger superpixels. The most prominent algorithm in current practice is SLIC (Simple Iterative Linear Clustering) [1]. Ideally, the superpixel setup should also be chosen by an appropriate minimization procedure that adheres object edges. However, edge adherence is often costly. An interesting exception is the Cut-Pursuit algorithm [20, 21], which solves total variation minimization and related problems in a fast sequence of binary graph cuts, making it competitive as a discretization step and leading to boundaries that better adhere with minimal partitions. This approach has been successfully applied in practice in such works as [15, 22] and we will contrapose a superpixel structure generated by Cut Pursuit with one generated by SLIC in our main comparison to a Cartesian grid.

3 Graph Discretizations for Convex Relaxations

3.1 Preliminaries

Let us first introduce our general notation. For the discrete setup we consider an undirected graph structure is defined by its vertices V, edges connecting vertices $E \subset V \times V$ and weights of these edges $w \in \mathbb{R}^{|E|}$. We refer to [13] for details.

The continuous minimal partition problem requires the definition of the total variation of a function $u = (u_1, \dots, u_L) \in L^1(\Omega, \mathbb{R}^L)$ as

$$TV(u) = \sup\left\{ -\int_\Omega \sum_{k=1}^{L} u_k(x) \operatorname{div} \mathbf{p}_k(x)\, dx,\ \ \mathbf{p} \in C_c^1(\Omega, \mathbb{R}^{d\times L}),\ \ \sum_{k=1}^{L} |\mathbf{p}_k(x)|^2 \leq 1 \right\},$$

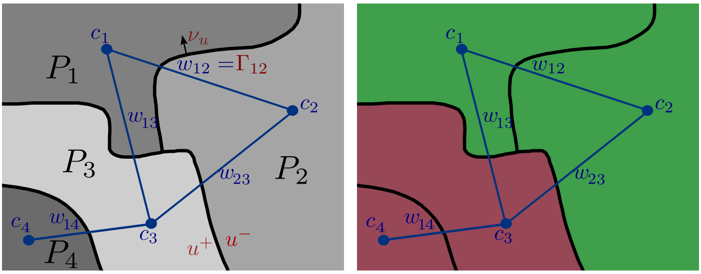

Figure 2: Sketch of the discretization process for a graph-based discretization. *Left:* The underlying continuous function $u \in BV_\Pi(\Omega, \mathbb{R}^L)$ is pictured in black, with piecewise-constant partitions $\Pi = \{P_1, P_2, P_3, P_4\}$ shown in gray as well as the discrete graph structure in blue. *Right:* A minimal partitions problem with two potentials is solved on this graph structure. Pictured is the solution u^* which now corresponds to a piecewise constant solution (in red and green).

Note that the above definition reduces to $TV(u) = \sum_{k=1}^{L} \int_\Omega |\nabla u_k(x)|\, dx$ for smooth u. We define u to be an element of the space of bounded variation $BV(\Omega, \mathbb{R}^L)$ if $TV(u)$ is finite. We can then identify this value with the mass of the distributional derivative Du, i.e. $\int_\Omega |Du| = TV(u)$. The bounded Radon measure Du can be decomposed [1, Thm. 10.4.1] into

$$Du = \nabla u\, \mathcal{L}^d + Cu + (u^+ - u^-) \otimes \nu_u \mathcal{H}^{d-1} \llcorner J_u, \tag{3}$$

where $\mathcal{L}^d$ is the Lebesgue measure, J_u is the jump set of u, where $u^+ \neq u^-$, i.e. the values at the boundary differ, ν_u the normal of the boundary and Cu a remainder Cantor part. In the following we will consider functions $u \in SBV(\Omega, \mathbb{R}^L)$, which is the space of functions for which $Cu = 0$ [2, 27]. We further define the perimeter of a measurable set $P \subset \Omega$, $\mathrm{Per}(\Omega, P)$, in turn by the total variation of its characteristic function $\chi_P : \Omega \to \mathbb{R}$ [9], the boundary of a set as $\partial S = \bar{S} \setminus \mathrm{int}(S)$, and the length of the boundary $\Gamma_{k,l} = \partial P_k \cap \partial P_l$ between two sets P_k, P_l via

$$|\Gamma_{k,l}| = \mathcal{H}^{d-1}(\Gamma_{k,l}), \tag{4}$$

where, again, $\mathcal{H}^{d-1}$ denotes the $d-1$-dimensional Hausdorff measure. These definitions allow us to examine the continuous boundary of shapes. Refer to Fig. 2, where these continuous objects are marked in red.

3.2 Graph Discretization

We are interested in solving the continuous minimal partition problems Eq. (2) numerically. To do so we need to translate the problem into the discrete setting. To take a step from the continuous definitions to a discrete problem, we make use of the fact that we expect solutions u^* to the minimal partitioning problem to be piecewise constant with a finite number of pieces. A good discretization to a finite setting that mimics this piecewise constant structure exactly. *We hence represent the discretization by a graph of candidate constant sets, the nodes of which represent each separate constant piece and where neighboring pieces are connected by edges in the graph.* After solving the matching problem on this discrete graph, the final solution u^* can be reassembled by assigning to each constant piece its matched value according to the respective value of the node that represents it. This setup is sketched in Fig. 2.

Note that this is a generalization of a classical discretization to a Cartesian grid. Placing a continuous function on a pixel grid corresponds to claiming that the function is piecewise-constant on every image pixel - hence the boundaries of the solution u^* to the minimal partitions problem will be a subset of the boundaries imposed by the image pixels.

In slight generalization of the minimal partitions problem in Eq. (2) we now discuss the type of continuous functionals $F : SBV(\Omega, \mathbb{R}^L) \to \mathbb{R}$, that we want to represent discretely. Due to the discontinuities of SBV, we define different components of F on the continuous parts and the jump parts J_u of Eq. (3):

$$F(u) = \int_{\Omega \setminus J_u} \Phi(x, u(x), \nabla u(x))\, dx + \int_{J_u} \kappa\left(|u^+ - u^-|\right) |\nu_u| d\mathcal{H}^{d-1}, \tag{5}$$

where ∇, u^+, u^- refer to the decomposition detailed in Eq. (3). $\Phi : \Omega \times \mathbb{R}^L \times \mathbb{R}^{L \times d} \to \mathbb{R}$ is a function defined away from the jump set of u, while $\kappa : \mathbb{R} \to \mathbb{R}$ is a concave function measuring the jump penalty with $\kappa(0) = 0$ [9, 27]. We can consider the first term to be a generalized data term, and the second as a (jump)-regularizer.

To connect the continuous formulation of Eq. (5) to a discrete setting we define the discretization as a finite set of candidate sets $\Pi = \{P_i \subset \Omega \mid P_i \cap P_j = \emptyset,\ \forall j \neq i\}$ with $M = |\Pi|$ partitions. The continuous function $u \in SBV(\Omega, \mathbb{R}^L)$ is assumed to be constant on every partition, so that we can denote its value on partition $P_i \in \Pi$ by a vector $c_i \in \mathbb{R}^L$. Thus, $u(x) = c_i$ for every $x \in P_i \subset \Omega$. The partition Π can be represented by a set of nodes $V = \{1, \dots, M\}$ where each node corresponds to a segment $P_i \in \Pi$. Furthermore we can describe every boundary between sets P_i and P_j as Γ_{ij} and by that define an edge set $E \subset V \times V$ as $E = \{(i,j) \in V \times V \mid |\Gamma_{ij}| > 0, i \neq j\}$. Note, that the perimeter of some partition $P_i \in \Pi$ is given by $\mathrm{Per}_{P_i} = \sum_{(i,j) \in E} |\Gamma_{ij}|$.

Let us assume that our desired solution u^*, which minimizes Eq. (5), is piecewise constant. More formally, given some partition Π let us write $u \in SBV_\Pi(\Omega, \mathbb{R}^L)$ to denote continuous functions in BV which are piecewise constant on the regions in Π, and assume $u^* \in SBV_\Pi(\Omega, \mathbb{R}^L)$. This implies that the jump set J_u is a subset of $\cup_{(i,j)\in E} \Gamma_{ij}$ and that $\Omega \setminus J_u$ is a subset $\cup_{i \in V} P_i$, or, in other words, the discrete partitioning by Π is able to represent the continuous structure of u^*.

Under the above assumption we can restrict the minimization of F over all functions $u \in SBV(\Omega, \mathbb{R}^L)$ to those in $SBV_\Pi(\Omega, \mathbb{R}^L)$ which allows to simplify Eq. (5) to a problem in which merely the values c_i inside the piecewise constant regions are the unknowns. Let us discuss the three main components of Eq. (5) separately.

Data Term:

Considering F for any $u \in SBV_\Pi(\Omega, \mathbb{R}^L)$ allows us to rewrite the first term of Eq. (5) as

$$K(u) = \int_{\Omega \setminus J_u} \Phi(x, u(x), \nabla u(x))\, dx = \sum_{i=1}^{M} \int_{P_i} \Phi(x, c_i, 0)\, dx =: K_\Pi(c) \tag{6}$$

which is the discrete representation $K_\Pi(c) : \mathbb{R}^{M \times L} \to \mathbb{R}$ of this term that mere depends on the values c_i. For linear data terms such as in Eq. (2), i.e. $\Phi(x, u(x), \nabla u(x)) = \sum_{k=1}^{L} f_k(x) u_k(x) = f_k(x)(c_k)_i$ for $x \in P_i$, this is further simplified to

$$K_\Pi(c) = \sum_{k=1}^{L} \sum_{i=1}^{M} (c_i)_k \int_{P_i} f_k(x)\, dx = \sum_{i=1}^{M} \langle c_i, \tilde{f}_k \rangle \tag{7}$$

with $\tilde{f}_k = \left(\int_{P_i} f_k(x) dx \right)_{k=1}^{L} \in \mathbb{R}^L$.

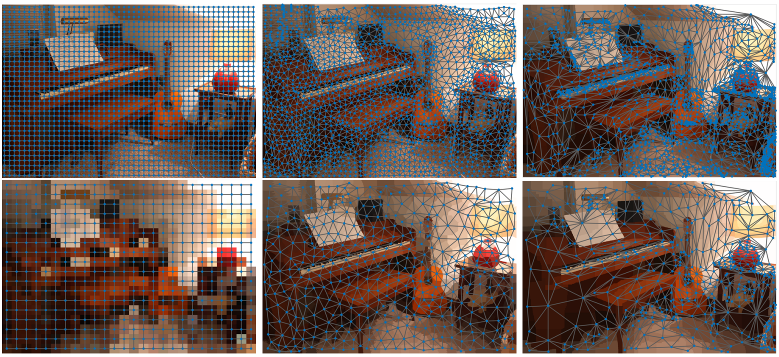

Figure 3: From left to right: Grid Sampling, SLIC Superpixels and L^0 Cut-Pursuit. Images from the Middlebury dataset [37]. The top row shows a fine discretization into the same number of nodes for every method, whereas the lower row shows a coarse discretization with the same number of nodes for every method.

Regularization Term:
For the jump regularization, we can write $R(u)$ for any $u \in SBV_\Pi(\Omega, \mathbb{R}^L)$ as

$$\begin{aligned} R(u) = \int_{J_u} \kappa\left(|u^+ - u^-|\right)\, d\mathcal{H}^{d-1} &= \sum_{(i,j)\in E} \int_{\Gamma_{ij}} \kappa\left(|c_i - c_j|\right)\, d\mathcal{H}^{d-1} \\ &= \sum_{(i,j)\in E} \kappa\left(|c_i - c_j|\right) \int_{\Gamma_{ij}} d\mathcal{H}^{d-1} = \sum_{(i,j)\in E} w_{ij}\, \kappa\left(|c_i - c_j|\right) =: R_\Pi(c) \end{aligned} \tag{8}$$

identifying the weights $w_{ij} = \int_{\Gamma_{ij}} d\mathcal{H}^{d-1} = |\Gamma_{ij}|$. With this weighting we can define the weighted finite graph $G = (V, E, w)$ as the discrete graph structure with which any continuous function $u \in SBV_\Pi(\Omega, \mathbb{R}^L)$ can be represented. Note that if κ is the identity, then $R_\Pi(c)$ is equivalent to graph total variation of c (cf. [6, 14]).

Constraint Set:
We are further carrying a constraint set when minimizing the minimal partitions problem. However, both constraints are pointwise and therefore straight forward to relate to constraints on c_i, i.e., the constraint set directly translates to

$$C_\Pi = \left\{ c \;\middle|\; (c_i)_k \in [0,1],\ \sum_{k=1}^{L} (c_i)_k = 1, \forall i \right\}.$$

Interestingly, the above restriction from the minimization of F over $SBV(\Omega, \mathbb{R}^L)$ to its minimization over $SBV_\Pi(\Omega, \mathbb{R}^L)$ (which translates into the minimization of F_Π over $c \in C_\Pi$) remains valid as long as the jump set of the true solution is a subset of the jumps in the partition Π, independent of what exactly the "super" jump-set of Π is. Let us formalize this result:

Proposition 1. *Assume a discretization Π and its assorted partitions P_i to be given. Let u^* be a minimizer to the continuous problem Eq. (2) for given potentials f_k. If the jump-set J_{u^*} of u^* is a subset of the jump set of Π given as the boundaries $\cup_{(i,j)\in E}\Gamma_{ij}$, then*

$$\min_{u\in C} F(u) = \min_{c\in C_\Pi} F_\Pi(c),$$

Figure 4: Illustrating the use of model-based segmentation methods: The user scribbles different objects to be segmented (left) from which a unary data term f is generated, e.g. by approaches like [29] or a pointwise neural network. As the unaries are insufficient for a good segmentation (illustrated by the fact that a thresholding of the unaries shown in the middle does not segment the background well), the proposed framework offers an efficient approach to obtain accurate segmentations as shown on the right.

for the discrete energy $F_\Pi = K_\Pi + R_\Pi$, i.e. the continuous minimum $F(u^*)$ is equal to the minimum $F_\Pi(c^*)$ of the discrete energy of F_Π under the constraints C_Π.

Proof. See appendix. □

Proposition 1 shows that if the jump set of u^* is contained in Pi, then the exact optimum u^* of the continuous problem can actually by found by computing a discrete solution c^* of the function F_Π numerically on a finite graph Practically however, we now need to find some partition Π that approximates (or ideally overestimates) the true jump set J_{u^*}, but consists of a limited number of segments. On a Cartesian grid, the equivalent operation is to subsample the image, result in the "superpixels" seen in Fig. 3 on the left, which are not well aligned with edges in the images. However approaches such as SLIC (middle) or Cut-Pursuit (right, in the variant of [41]) are more adept at finding a superset of candidate partitions.

4 Numerical Evaluation

This section focuses on evaluating the proposed approach. We discuss examples in segmentation and stereo estimation. For segmentation we show a practical example, where the approach is used to align the output of a pixelwise neural network. We then follow up with a detailed comparison of graphs generated by SLIC, Cut Pursuit and subsampling.

4.1 Segmentation

Multi-label segmentation is a central application of minimal partition problems, having been discussed in the continuous setting in works such as [10, 31] and widely studied in discrete methods such as [8, 19]. To apply multilabel segmentation we use the model described in Eq. (1), which can be recovered from Eq. (5) by setting $\kappa = \mathrm{Id}$ and choosing the L^1 norm for $|\cdot|$, leading to an anisotropic penalty of the jumps. We solve the discrete matching on the graph by a preconditioned primal-dual algorithm as discussed in [41]. Relaxed solutions are matched to corresponding partitions with maximal argument.

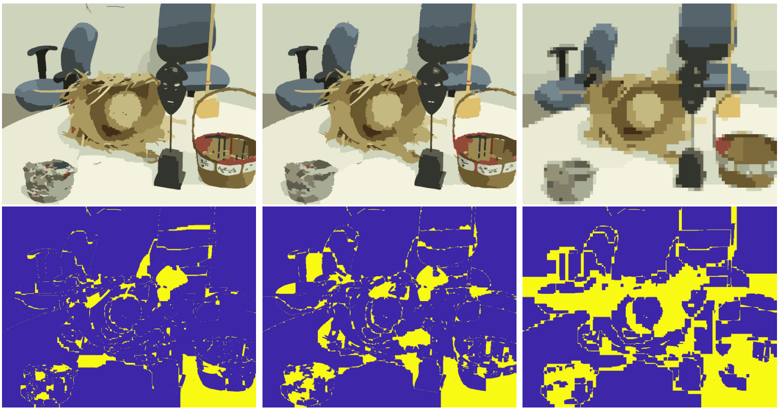

Figure 5: Qualitative comparison of matching quality for the example of image segmentation from a given set of pixel features. From left to right:L^0-CP graph, the SLIC graph, and a rectangular grid (right), all with the same number of vertices. The graphs are constructed as described in Section 3.2. The top images show the final minimal partition result. The bottom images show the errors compared to the minimal partition computed on the full pixel grid, where yellow marks regions that are matched differently compared to the ground truth matching of the full image grid.

Use Case:
As one application scenario, imagine a user wants to segment an image by marking the objects to be segmented with scribbles, see Fig. 4 on the left. Once the scribbling is complete we train a tiny pixelwise fully connected network on classifying the scribbled pixels correctly. The output of this network provides us with pixelwise features as shown in Fig. 4 in the middle. Globally matching these features with 15 labels on the full grid requires 6 GB of memory, whereas the proposed approach reduces the memory requirements to 0.2 GB due to the graph construction and needs only 13% of computation time in total to yield the segmentation shown in Fig. 4 on the right, which is precise enough to conduct various image manipulations such as inserting or removing some of the bottles or fruits.

Quantitative Analysis:
To analyze a wide range of images with canonical potentials, we turn to cartooning, i.e. multi-label segmentation with a fixed set of target colors chosen by a k-means selection. Figure 5 visualizes the result of the minimal partition problem for our graph discretization via L^0-CP, a graph constructed via SLIC, and a subsampling of the pixel grid, all with the same number of vertices. Checking the error maps on the bottom row of Fig. 5 we see that both superpixel methods lead to solutions that closely match the solution at the finest level, while the subsampling is comparatively error-prone. The L^0-CP constructed discretization outperforms the SLIC-based discretization, due to its closer adherence to image edges. Evaluating the gained efficiency in terms of time and in times of vertices in Fig. 6 shows that this behavior leads to stable improvements over a wide range of graph discretization steps, energy values can be matched very closely using the superpixel-based graph discretization. In Table 1 we compare the three methods for three different examples. We find significant savings in runtime and memory, while staying close to the original energy, observing that graph-based discretization leads to a significant improvement in accuracy, compared to computing the segmentation on a downsampled grid, and further that using the L^0-CP superpixels leads

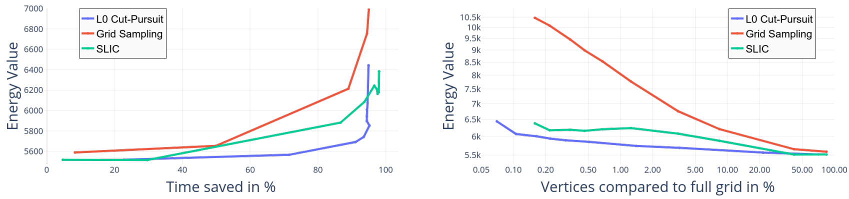

Figure 6: Computing the minimal partition on a well chosen graph discretization is much more efficient than computing it on the full grid. The y-axis in both plots denotes the energy value of the ground truth solution, which is computed on the full grid. Left: Time saved vs ground truth plotted vs the matching energy of the minimizer. Right: The number of nodes compared to the partition energy of the minimizer.

Ex.	Methods	Red. Rate	Time Saved	Mem.	Energy Offset
1	L^0-CP	**31%**	**79%**	**15MB**	**0.73%**
	SLIC	41%	41%	16MB	1.18%
	sampling	74%	22%	32MB	5.77%
2	L^0-CP	**3.6%**	**87%**	**12MB**	**3.5%**
	SLIC	8.42%	**87%**	32MB	6.29%
	sampling	8.42%	**87%**	27MB	13.33%
3	L^0-CP	**6.2%**	**83%**	**463MB**	**2.1%**
	SLIC	14%	79%	1144MB	4.43%
	sampling	14%	82%	2976MB	3.79%

Table 1: Different scores for three examples. Shown are the ratio of time saved and the ratio of energy mismatch. The baseline method (a full image grid) uses 301, 684 and 6113 MB for each experiment.

to the most efficient final result, even though the computation of these superpixels itself takes more time than SLIC (the time/memory to compute superpixels and construct the graph is factored into all measurements we consider).

4.2 Stereo Matching

For the task of estimating disparities in stereo images the problem setting is different from that of segmentation tasks. While we want to reconstruct discrete labels for segmentation the estimated disparities between images live in a continuous range and therefore need dedicated treatment. The binarized vector structure of discrete segmentation labels ideally has to be translated to a continuously metricized label space. Assuming piecewise constant disparities in natural images it is still possible to transfer the graph reduction ideas to stereo estimation as proposed in [32] and the sublabel accurate setting of [28]. As the data term for stereo matching can be expressed as a cost-vector defined for each pixel, we need to find a sensitive scalar data function where superpixels can be computed to construct the graph. A natural choice for such a function is the pixelwise minimizing argument of the data term. This is motivated by the intuition that constant regions of pointwise minimizing disparities likely induce constant regions of the original data term. The features f_k are either given directly as absolute pixelwise disparities, or as output of a stereo network and are then matched globally to combinations of candidate disparities in a given range. Figure 7 (top) gives a visual impression of the approximation behavior of the graph reduction on an exemplary

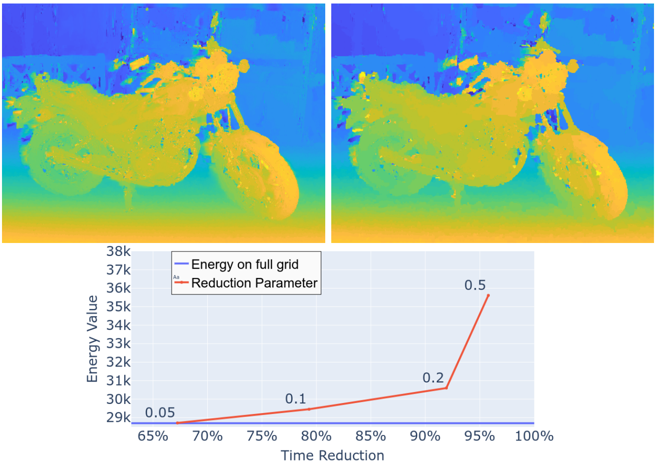

Figure 7: Results on stereo matching baselines. *Top:* Comparison of full (left) and proposed reduced (right) matching. Both methods use sublabels [28] between 32 labels. The proposed method uses Cut-Pursuit (with parameter $\alpha_c = 0.1$, a higher parameter corresponds to fewer vertices in the reduced graph) to find a reduced graph, amounting to a time reduction by a factor of 4.8, although the matching quality is near indistinguishable. *Bottom:* Time reduction and energy levels for different parameters α_c, showing the granular relationship between graph reduction and difference in energy value of the matching algorithm.

stereo image. Figure 7 (bottom) visualizes the time vs. the achieved energy values of our method compared to the full stereo matching problem. Note the scale of the x-axis. We can easily reduce the necessary time and memory costs by using the graph-based discretization. Despite of the significant speedup for stereo matching the proposed method still is capable of producing visually pleasing results, as the matching is still computed with respect to all variables, just with an optimally chosen discretization.

5 Conclusions

In this work we presented strategies for the efficient realization of convex relaxations by directly moving from geometric properties of minimal partitions solutions to a graph-based discretization. We prove that such a graph-based discretization can be constructed in adherence to the global partitioning problem and implementing it on superpixel graphs yields accurate and efficient solutions in practice. We further find that using a superpixel approach that is more faithful to minimal surface energies, as the L^0-Cut Pursuit algorithm leads to more accurate solutions compared to SLIC and can be well worth the additional effort. We believe that the proposed methodology can facilitate the use of convex relaxation methods in practical applications, especially if input data is of high-resolution, where memory and computation constraints made these approaches previously infeasible.

Acknowledgements:
This research was supported by the German Research Foundation (DFG) under grant MO 2962/2-1.

References

[1] Radhakrishna Achanta, Appu Shaji Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. *IEEE Trans. Pattern Anal. Mach. Intell.*, 34(11):2274–2282, November 2012. ISSN 0162-8828. doi: 10.1109/TPAMI.2012.120.

[2] Luigi Ambrosio, Nicola Fusco, and Diego Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford university press, Oxford, 2000.

[3] H. Attouch, Giuseppe Buttazzo, and Gérard Michaille. *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*. MPS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics : Mathematical Programming Society, Philadelphia, 2006. ISBN 978-0-89871-600-9.

[4] Yuri Boykov and Gareth Funka-Lea. Graph Cuts and Efficient N-D Image Segmentation. *Int. J. Comput. Vis.*, 70(2):109–131, November 2006. ISSN 0920-5691, 1573-1405. doi: 10.1007/s11263-006-7934-5.

[5] Yuri Boykov, Olga Veksler, and Ramin Zabih. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1222–1239, November 2001. ISSN 0162-8828. doi: 10.1109/34.969114.

[6] Xavier Bresson and Tony F. Chan. Non-local Unsupervised Variational Image Segmentation Models. UCLA CAM report, University of California, Los Angeles, 2008.

[7] Thomas Brox, Andrés Bruhn, Nils Papenberg, and Joachim Weickert. High Accuracy Optical Flow Estimation Based on a Theory for Warping. In *Computer Vision - ECCV 2004*, volume 3024, pages 25–36. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004. ISBN 978-3-540-21981-1 978-3-540-24673-2. doi: 10.1007/978-3-540-24673-2_3.

[8] Martin Burger and Stanley Osher. A Guide to the TV zoo. In *PDE Based Reconstruction Methods in Imaging*, number 2090 in Lecture Notes in Mathematics. Springer International Publishing, Switzerland, first edition, 2013.

[9] Antonin Chambolle, Daniel Cremers, and Thomas Pock. A Convex Approach to Minimal Partitions. *SIAM J. Imaging Sci.*, 5(4):1113–1158, October 2012. ISSN 1936-4954. doi: 10.1137/110856733.

[10] Tony F. Chan and Luminita A. Vese. Active contours without edges. *IEEE Trans. Image Process.*, 10(2):266–277, 2001.

[11] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L. Yuille. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *ArXiv160600915 Cs*, June 2016.

[12] Daniel Cremers, Thomas Pock, Kalin Kolev, and A. Chambolle. Convex Relaxation Techniques for Segmentation, Stereo and Multiview Reconstruction. In *Markov Random Fields for Vision and Image Processing*. MIT Press, Boston, 2011.

[13] Abderrahim Elmoataz, Olivier Lezoray, and Sébastien Bougleux. Nonlocal Discrete Regularization on Weighted Graphs: A framework for Image and Manifold Processing. *IEEE Trans. Image Process.*, 17(7):1047–1060, 2008.

[14] Guy Gilboa and Stanley Osher. Nonlocal Operators with Applications to Image Processing. *Multiscale Model. Simul.*, 7(3):1005–1028, November 2008. ISSN 1540-3459. doi: 10.1137/070698592.

[15] Stephane Guinard, Loic Landrieu, Laurent Caraffa, and Bruno Vallet. Piecewise-Planar Approximation of Large 3D Data as Graph-Structured Optimization. In *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume IV-2-W5, pages 365–372. Copernicus GmbH, May 2019. doi: 10.5194/isprs-annals-IV-2-W5-365-2019.

[16] Berthold K. P. Horn and Brian G. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1):185–203, August 1981. ISSN 0004-3702. doi: 10.1016/0004-3702(81)90024-2.

[17] Hiroshi Ishikawa and Davi Geiger. Occlusions, discontinuities, and epipolar lines in stereo. In *Computer Vision — ECCV'98*, Lecture Notes in Computer Science, pages 232–248. Springer Berlin Heidelberg, 1998. ISBN 978-3-540-69354-3.

[18] Kalin Kolev and Daniel Cremers. Integration of Multiview Stereo and Silhouettes Via Convex Functionals on Convex Domains. In *Computer Vision – ECCV 2008*, Lecture Notes in Computer Science, pages 752–765, Berlin, Heidelberg, 2008. Springer. ISBN 978-3-540-88682-2. doi: 10.1007/978-3-540-88682-2_57.

[19] Vladimir Kolmogorov and Ramin Zabin. What energy functions can be minimized via graph cuts? *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(2):147–159, February 2004. ISSN 0162-8828. doi: 10.1109/TPAMI.2004.1262177.

[20] Loic Landrieu and Guillaume Obozinski. Cut Pursuit: Fast algorithms to learn piecewise constant functions. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pages 1384–1393, Cadiz, Spain, April 2016. PMLR.

[21] Loic Landrieu and Guillaume Obozinski. Cut Pursuit: Fast Algorithms to Learn Piecewise Constant Functions on General Weighted Graphs. *SIAM J. Imaging Sci.*, 10(4):1724–1766, January 2017. doi: 10.1137/17M1113436.

[22] Loic Landrieu and Martin Simonovsky. Large-scale Point Cloud Semantic Segmentation with Superpoint Graphs. *ArXiv171109869 Cs*, November 2017.

[23] Emanuel Laude, Thomas Möllenhoff, Michael Moeller, Jan Lellmann, and Daniel Cremers. Sublabel-Accurate Convex Relaxation of Vectorial Multilabel Energies. In *Computer Vision – ECCV 2016*, Lecture Notes in Computer Science, pages 614–627. Springer, Cham, October 2016. ISBN 978-3-319-46447-3 978-3-319-46448-0. doi: 10.1007/978-3-319-46448-0_37.

[24] Jan Lellmann, Jörg Kappes, Jing Yuan, Florian Becker, and Christoph Schnörr. Convex Multi-class Image Labeling by Simplex-Constrained Total Variation. In *Scale Space and Variational Methods in Computer Vision*, Lecture Notes in Computer Science, pages 150–162. Springer Berlin Heidelberg, 2009. ISBN 978-3-642-02256-2.

[25] Jan Lellmann, Evgeny Strekalovskiy, Sabrina Koetter, and Daniel Cremers. Total Variation Regularization for Functions with Values in a Manifold. In *Proceedings of the 2013 IEEE International Conference on Computer Vision*, ICCV '13, pages 2944–2951, Washington, DC, USA, 2013. IEEE Computer Society. ISBN 978-1-4799-2840-8. doi: 10.1109/ICCV.2013.366.

[26] David Mumford and Jayant Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Commun. Pure Appl. Math.*, 42(5):577–685, 1989.

[27] Thomas Möllenhoff and Daniel Cremers. Sublabel-Accurate Discretization of Nonconvex Free-Discontinuity Problems. *Proc. IEEE Int. Conf. Comput. Vis.*, pages 1183–1191, August 2017. doi: 10.1109/ICCV.2017.134.

[28] Thomas Möllenhoff, Emanuel Laude, Michael Moeller, Jan Lellmann, and Daniel Cremers. Sublabel-Accurate Relaxation of Nonconvex Energies. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3948–3956, 2016. doi: 10.1109/CVPR.2016.428.

[29] Claudia Nieuwenhuis, Eno Töppe, and Daniel Cremers. A Survey and Comparison of Discrete and Continuous Multi-label Optimization Approaches for the Potts Model. *Int J Comput Vis*, 104(3):223–240, September 2013. ISSN 1573-1405. doi: 10.1007/s11263-013-0619-y.

[30] Thomas Pock, Thomas Schoenemann, Gottfried Graber, Horst Bischof, and Daniel Cremers. A convex formulation of continuous multi-label problems. In *European Conference on Computer Vision*, pages 792–805. Springer, 2008.

[31] Thomas Pock, Antonin Chambolle, Daniel Cremers, and Horst Bischof. A convex relaxation approach for computing minimal partitions. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference On*, pages 810–817. IEEE, 2009.

[32] Thomas Pock, Daniel Cremers, Horst Bischof, and Antonin Chambolle. Global Solutions of Variational Models with Convex Regularization. *SIAM J. Imaging Sci.*, 3(4): 1122–1145, January 2010. doi: 10.1137/090757617.

[33] Renfrey Burnard Potts. Some generalized order-disorder transformations. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 48, pages 106–109. Cambridge Univ Press, 1952.

[34] Rene Ranftl, Stefan Gehrig, Thomas Pock, and Horst Bischof. Pushing the limits of stereo using variational stereo estimation. In *Intelligent Vehicles Symposium (IV), 2012 IEEE*, pages 401–407. IEEE, 2012.

[35] Rene Ranftl, Thomas Pock, and Horst Bischof. Minimizing TGV-Based Variational Models with Non-convex Data Terms. In *Scale Space and Variational Methods in Computer Vision*, Lecture Notes in Computer Science, pages 282–293. Springer Berlin Heidelberg, June 2013. ISBN 978-3-642-38266-6 978-3-642-38267-3. doi: 10.1007/978-3-642-38267-3_24.

[36] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, November 1992. ISSN 0167-2789. doi: 10.1016/0167-2789(92)90242-F.

[37] Daniel Scharstein, Heiko Hirschmüller, York Kitajima, Greg Krathwohl, Nera Nešić, Xi Wang, and Porter Westling. High-Resolution Stereo Datasets with Subpixel-Accurate Ground Truth. In *Pattern Recognition*, Lecture Notes in Computer Science, pages 31–42. Springer, Cham, September 2014. ISBN 978-3-319-11751-5 978-3-319-11752-2. doi: 10.1007/978-3-319-11752-2_3.

[38] Mohamed Souiai, Martin R. Oswald, Youngwook Kee, Junmo Kim, Marc Pollefeys, and Daniel Cremers. Entropy Minimization for Convex Relaxation Approaches. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1778–1786, 2015.

[39] Evgeny Strekalovskiy, Bastian Goldluecke, and Daniel Cremers. Tight convex relaxations for vector-valued labeling problems. In *2011 International Conference on Computer Vision*, pages 2328–2335, November 2011. doi: 10.1109/ICCV.2011.6126514.

[40] Evgeny Strekalovskiy, Antonin Chambolle, and Daniel Cremers. Convex Relaxation of Vectorial Problems with Coupled Regularization. *SIAM J. Imaging Sci.*, 7(1):294–336, January 2014. ISSN 1936-4954. doi: 10.1137/130908348.

[41] Daniel Tenbrinck, Fjedor Gaede, and Martin Burger. Variational Graph Methods for Efficient Point Cloud Sparsification. *ArXiv190302858 Cs Math*, March 2019.

[42] Roy Uziel, Meitar Ronen, and Oren Freifeld. Bayesian Adaptive Superpixel Segmentation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 8470–8479, 2019.

[43] Manuel Werlberger, Markus Unger, Thomas Pock, and Horst Bischof. Efficient Minimization of the Non-local Potts Model. In *Scale Space and Variational Methods in Computer Vision*, Lecture Notes in Computer Science, pages 314–325. Springer Berlin Heidelberg, May 2011. ISBN 978-3-642-24784-2 978-3-642-24785-9. doi: 10.1007/978-3-642-24785-9_27.

[44] Christopher Zach, Thomas Pock, and Horst Bischof. A Duality Based Approach for Realtime TV-L1 Optical Flow. In *Pattern Recognition*, pages 214–223. Springer, Berlin, Heidelberg, September 2007. doi: 10.1007/978-3-540-74936-3_22.

[45] Christopher Zach, David Gallup, Jan-Michael Frahm, and Marc Niethammer. Fast Global Labeling for Real-Time Stereo Using Multiple Plane Sweeps. In *Vision, Modeling, and Visualization*, pages pp. 243–252, Amsterdam, The Netherlands, August 2008. IOS Press.

[46] Christopher Zach, Christian Häne, and Marc Pollefeys. What is optimized in tight convex relaxations for multi-label problems? In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1664–1671, June 2012. doi: 10.1109/CVPR.2012.6247860.