

# A Convex Formulation of Continuous Multi-Label Problems

Thomas Pock<sup>1,2</sup>, Thomas Schoenemann<sup>1</sup>, Gottfried Graber<sup>2</sup>, Horst Bischof<sup>2</sup>,  
and Daniel Cremers<sup>1</sup>

<sup>1</sup> Department of Computer Science, University of Bonn

<sup>2</sup> Institute for Computer Graphics and Vision, Graz University of Technology

**Abstract.** We propose a spatially continuous formulation of Ishikawa’s discrete multi-label problem. We show that the resulting non-convex variational problem can be reformulated as a convex variational problem via embedding in a higher dimensional space. This variational problem can be interpreted as a minimal surface problem in an anisotropic Riemannian space. In several stereo experiments we show that the proposed continuous formulation is superior to its discrete counterpart in terms of computing time, memory efficiency and metrication errors.

## 1 Introduction

Many Computer Vision problems can be formulated as labeling problems. The task is to assign a label to each pixel of the image such that the label configuration is minimal with respect to a discrete energy.

A large class of binary labeling problems can be globally minimized using graph cut algorithms [1, 2]. Applications of binary labeling problems include two-region image segmentation, shape denoising and 3D reconstruction. On the other hand, multi-label problems in general cannot be globally minimized. They can only be solved approximately within a known error bound [3–6]. There exists one exception where multi-label problems can be solved exactly. Ishikawa [7] showed that, if the pairwise interactions are convex in terms of a linearly ordered label set, one can compute the exact solution of the multi-label problem. Applications of multi-label problems include image restoration, inpainting, multi-region image segmentation, motion and stereo.

The continuous counterpart to discrete labeling problems is the variational approach. Similar to the labeling problem, the aim of the variational approach is to find the minimizer of an energy functional. The major difference between the variational approach and the discrete labeling approach is that the energy functional is defined in a spatially continuous setting and the unknown functions can take continuous values. If the energy functional is convex and the minimization is carried out over a convex set, the globally optimal solution can be computed. On the other hand, it is generally hard to minimize non-convex energy functionals globally.

In this paper we present a new variational method which allows to compute the exact minimizer of an energy functional incorporating Total Variation regularization and a non-convex data term. Our method can solve problems of the same complexity as the Ishikawa’s method. Hence, our method can be seen as the continuous counterpart.

Our method comes along with several advantages compared to Ishikawa’s approach. First, our method is largely independent from grid bias, also known as metrication error. This leads to more accurate approximations of the continuous solution. Second, our method is based on variational optimization techniques which can be effectively accelerated on parallel architectures such as graphics processing units (GPUs) and third, it requires less memory. Fourth, our method allows to compute sub-pixel-accurate solutions.

The remainder of the paper is follows. In Section 2 we review the method of Ishikawa. In Section 3 we give the definition of the energy functional which can be solved with our method. We show how this non-convex energy functional can be cast by an equivalent convex optimization problem. In Section 4 we show results of our method applied to stereo. In the last Section we give some conclusions and show directions for future investigations.

## 2 Ishikawa’s Discrete Approach

Ishikawa [7] presents a method to globally solve multi-label problems of a certain class. A less general class was given independently by Veksler [4].

Given a graph with node set  $\mathcal{V}$  and edge set  $\mathcal{E}$  and a label set  $L \subset \mathbb{Z}$ , Ishikawa considers the task to compute the optimal labeling  $l \in L^{\mathcal{V}}$  for an energy of form

$$\min_l \sum_{(u,v) \in \mathcal{E}} P(l(u) - l(v)) + \sum_{v \in \mathcal{V}} D(l(v)) \quad (1)$$

Such a labeling problem combines a certain pairwise regularity term  $P(\cdot)$  with an (arbitrary) data term  $D(\cdot)$ . Many problems in Computer Vision can be stated in this form, among them are stereo estimation, image restoration and image segmentation.

Ishikawa shows that such problems can be solved in a globally optimal manner as long as the function  $P(\cdot)$  is convex in  $l(u) - l(v)$ . This is achieved by computing the minimal cut in an auxiliary graph with extended node set. For each combination of node  $v \in \mathcal{V}$  and label  $l(v) \in L$  a node in the auxiliary graph is created. For details see [7].

While this approach is able to find global optimizers of a discrete energy, in practice it suffers from several drawbacks:

- The algorithm requires a huge amount of memory. In part this is due to the large set of nodes. The true bottleneck however lies in the algorithms to find the minimal cut in the graph: All efficient solvers are based on computing the maximal flow in the graph [8]. This requires the storage of a flow value for each edge and hence an explicit storage of edges.

- Graph-based methods generally suffer from grid bias (also known as metrication errors). To remove this grid bias and get close to rotational invariance, large neighborhood systems are required. The resulting huge number of edges increases the memory consumption even further.
- Lastly the efficient parallelization of max-flow-based methods is still an open issue. While current graphics cards offer highly parallel architectures, to date this potential could not be exploited to speed up max-flow algorithms.

In this paper we deal with all of these drawbacks. We propose a sub-pixel-accurate continuous formulation which makes use of continuous optimization techniques. As a direct consequence our method does not suffer from grid bias. Moreover, it requires much less memory and is easy to parallelize.

### 3 A Continuous Approach

This work is devoted to the study of the variational problem

$$\min_u \left\{ \int_{\Omega} |\nabla u(\mathbf{x})| d\mathbf{x} + \int_{\Omega} \rho(u(\mathbf{x}), \mathbf{x}) d\mathbf{x} \right\}, \quad (2)$$

which can be seen as the continuous counterpart of (1), where we used  $P(\cdot) = |\cdot|$ . Let  $u : \Omega \rightarrow \Gamma$  be the unknown function, where  $\Omega \subseteq \mathbb{R}^2$  is the image domain,  $\Gamma = [\gamma_{min}, \gamma_{max}]$  is the range of  $u$  and  $\mathbf{x} = (x, y)^T \in \Omega$  is the pixel coordinate. We may assume homogeneous Neumann boundary conditions for  $u$  on  $\partial\Omega$ .

The left term of (2) is for regularization, i.e. to obtain smooth results. It is based on minimizing the Total Variation (TV) of  $u$ . Note that the gradient operator is understood in its distributional sense. Therefore, the TV energy is also well-defined for discontinuous functions (e.g. characteristic functions).

$$|\nabla u(\mathbf{x})| = \sqrt{\left(\frac{\partial u(\mathbf{x})}{\partial x}\right)^2 + \left(\frac{\partial u(\mathbf{x})}{\partial y}\right)^2}. \quad (3)$$

The main property of the TV term is that it allows for sharp discontinuities in the solution while still being a convex function [9]. The discontinuity preserving property is important for many Computer Vision problems, e.g. to preserve edges in the solution.

The right term of (2) is the data term. It is based on a pixel-wise defined non-negative function  $\rho(u(\mathbf{x}), \mathbf{x}) : \Omega \rightarrow \mathbb{R}^+$ , which directly relates to the data term  $D(\cdot)$  of Ishikawa’s discrete approach. Note that our model is able to handle any pixel-wise defined data term, including non-convex ones. The type of data term also defines the application domain of our variational model. For example, if  $\rho(u, f)$  measures the fidelity of  $u$  to given noisy input image  $f$ , our model could be used for image denoising. On the other hand, if  $u$  represents a disparity field and  $\rho(u, I_L, I_R)$  measures the matching quality of a rectified stereo image pair  $I_L$  and  $I_R$ , our model could be used for stereo matching.

Let us now discuss whether we can find an exact solution of (2). The regularization term is convex in  $u$ . Therefore, this term can be globally minimized. However,  $\rho(u)$  is per definition non-convex. Hence, we cannot expect that we are able to compute the global minimizer in this setting.

### 3.1 A Convex Formulation via Functional Lifting

In this section, we will develop a convex formulation of the non-convex variational model (2). The key idea is to lift the original problem formulation to a higher-dimensional space by representing  $u$  in terms of its level sets. Consequently, this will allow us to compute the exact solution of the original non-convex problem. Let us first give some definitions.

**Definition 1** *Let the characteristic function  $\mathbf{1}_{\{u>\gamma\}}(\mathbf{x}) : \Omega \rightarrow \{0,1\}$  be the indicator for the  $\gamma$  - super-levels of  $u$ :*

$$\mathbf{1}_{\{u>\gamma\}}(\mathbf{x}) = \begin{cases} 1 & \text{if } u(\mathbf{x}) > \gamma \\ 0 & \text{otherwise} \end{cases} . \quad (4)$$

Next, we make use of the above defined characteristic functions to construct a binary function  $\phi$  which resembles the graph of  $u$ .

**Definition 2** *Let  $\phi : [\Omega \times \Gamma] \rightarrow \{0,1\}$  be a binary function defined as*

$$\phi(\mathbf{x}, \gamma) = \mathbf{1}_{\{u>\gamma\}}(\mathbf{x}) . \quad (5)$$

As a direct consequence of (4) we see that  $\phi(\mathbf{x}, \gamma_{min}) = 1$  and  $\phi(\mathbf{x}, \gamma_{max}) = 0$ . Hence, the feasible set of functions  $\phi$  is given by

$$D' = \{ \phi : \Sigma \rightarrow \{0,1\} \mid \phi(\mathbf{x}, \gamma_{min}) = 1, \phi(\mathbf{x}, \gamma_{max}) = 0 \} , \quad (6)$$

where we used the short notation  $\Sigma = [\Omega \times \Gamma]$ . Note that the function  $u$  can be recovered from  $\phi$  using the following layer cake formula [10].

$$u(\mathbf{x}) = \gamma_{min} + \int_{\Gamma} \phi(\mathbf{x}, \gamma) d\gamma \quad (7)$$

Our intention is now to rewrite the variational problem (2) in terms of  $\phi$ . This can be seen as lifting the variational problem (2) to a higher-dimensional space. This is stated by the following Theorem which forms the basis of our approach.

**Theorem 1** *The variational problem (2) is equivalent to the higher dimensional variational problem*

$$\min_{\phi \in D'} \left\{ \int_{\Sigma} |\nabla \phi(\mathbf{x}, \gamma)| + \rho(\mathbf{x}, \gamma) |\partial_{\gamma} \phi(\mathbf{x}, \gamma)| d\Sigma \right\} , \quad (8)$$

*in the sense that the minimizer of (8) is related to the minimizer of (2) via the layer cake formula (7).*

**Proof:** First, the TV term of (2) can be easily rewritten in terms of  $\phi$ , making use of the generalized co-area formula of Fleming and Rishel [11].

$$\int_{\Omega} |\nabla u(\mathbf{x})| d\mathbf{x} = \int_{\Omega} \left\{ \int_{\Gamma} |\nabla \phi(\mathbf{x}, \gamma)| d\gamma \right\} d\mathbf{x}, \quad (9)$$

where  $|\nabla \phi(\mathbf{x}, \gamma)|$  denotes the Total Variation of the characteristic function of the  $\gamma$  - super-levels of  $u$ . The co-area formula essentially states that the TV norm can be decomposed into a sum of the length of the level - sets of  $u$ .

Second, we have to rewrite the data term of (2) by means of  $\phi$ . From (5) we observe that

$$|\partial_{\gamma} \phi(\mathbf{x}, \gamma)| \equiv \delta(u(\mathbf{x}) - \gamma) \quad (10)$$

where  $\delta(\cdot)$  is the Dirac Delta function. As a direct consequence, the data term can be rewritten as

$$\begin{aligned} \int_{\Omega} \rho(u(\mathbf{x}), \mathbf{x}) d\mathbf{x} &= \int_{\Omega} \left\{ \int_{\Gamma} \rho(\gamma, \mathbf{x}) \delta(u(\mathbf{x}) - \gamma) d\gamma \right\} d\mathbf{x} \\ &= \int_{\Omega} \left\{ \int_{\Gamma} \rho(\gamma, \mathbf{x}) |\partial_{\gamma} \phi(\mathbf{x}, \gamma)| d\gamma \right\} d\mathbf{x} \end{aligned} \quad (11)$$

By substitution of the terms (9) and (11) into (2), we arrive at the higher dimensional variational model (8).  $\square$

Although (8) is convex in  $\phi$ , the variational problem is still non-convex since the minimization is carried out over  $D'$  which is a non-convex set. The idea is now to relax the variational problem (8) by allowing  $\phi$  to vary smoothly in the interval  $[0, 1]$ . This leads to the following convex set of feasible solutions of  $\phi$ .

$$D = \{ \phi : \Sigma \rightarrow [0, 1] \mid \phi(\mathbf{x}, \gamma_{min}) = 1, \phi(\mathbf{x}, \gamma_{max}) = 0 \}. \quad (12)$$

The associated variational problem is now given by

$$\min_{\phi \in D} \left\{ \int_{\Sigma} |\nabla \phi(\mathbf{x}, \gamma)| + \rho(\mathbf{x}, \gamma) |\partial_{\gamma} \phi(\mathbf{x}, \gamma)| d\Sigma \right\}. \quad (13)$$

Since (13) is convex in  $\phi$  and minimization is carried over  $D$ , which is a convex set, the overall variational problem is convex. This means that we are able to compute its global minimizer.

Our intention, however, is still to solve the binary problem (8). Fortunately, minimizers of the relaxed problem can be transformed to minimizers of the binary problem. Based on [10] we state the following thresholding theorem.

**Theorem 2** *Let  $\phi^* \in D$  be the solution of the relaxed variational problem (13). Then for almost any threshold  $\mu \in [0, 1]$  the characteristic function  $\mathbf{1}_{\{\phi^* \geq \mu\}} \in D'$  is also a minimizer of the binary variational problem (8).*

**Proof:** (Proof by Contradiction.) Since (13) is homogeneous of degree one, we can make use of the generalized co-area formula to decompose (13) by means of the level sets of  $\phi$ .

$$\begin{aligned} E(\phi) &= \int_{\Sigma} |\nabla\phi(\mathbf{x}, \gamma)| + \rho(\mathbf{x}, \gamma)|\partial_{\gamma}\phi(\mathbf{x}, \gamma)| d\Sigma \\ &= \int_0^1 \left\{ \int_{\Sigma} |\nabla\mathbf{1}_{\{\phi \geq \mu\}}| + \rho(\mathbf{x}, \gamma)|\partial_{\gamma}\mathbf{1}_{\{\phi \geq \mu\}}| d\Sigma \right\} d\mu \\ &= \int_0^1 E(\mathbf{1}_{\{\phi \geq \mu\}}) d\mu \end{aligned} \quad (14)$$

Assume to the contrary that  $\mathbf{1}_{\{\phi^* \geq \mu\}} \in D'$  is not a global minimizer of the binary problem, i.e. there exists a binary function  $\phi' \in D'$  with  $E(\phi') < E(\mathbf{1}_{\{\phi^* \geq \mu\}})$  for a measurable set of  $\mu \in [0, 1]$ . This directly implies that

$$E(\phi') = \int_0^1 E(\phi') d\mu < \int_0^1 E(\mathbf{1}_{\{\phi^* \geq \mu\}}) d\mu = E(\phi^*), \quad (15)$$

which means that  $\phi^*$  is not a global minimizer of  $E(\cdot)$ , contradicting our assumption.  $\square$

We have seen that solving the non-convex variational problem (2) amounts to solving the convex variational problem (13). In the following section we will develop a simple but efficient numerical algorithm to compute the solution of (13).

### 3.2 Computing the Solution of the Relaxed Functional

The fundamental approach to minimize (13) is to solve its associated Euler-Lagrange differential equation.

$$-\operatorname{div} \left( \frac{\nabla\phi}{|\nabla\phi|} \right) - \partial_{\gamma} \left( \rho \frac{\partial_{\gamma}\phi}{|\partial_{\gamma}\phi|} \right) = 0, \quad \text{s.t. } \phi \in D. \quad (16)$$

It is easy to see that these equations are not defined either as  $|\nabla\phi| \rightarrow 0$  or  $|\partial_{\gamma}\phi| \rightarrow 0$ . In order to resolve these discontinuities, one could use regularized variants of these terms, e.g.  $|\nabla\phi|_{\varepsilon} = \sqrt{|\nabla\phi|^2 + \varepsilon^2}$  and  $|\partial_{\gamma}\phi|_{\varepsilon} = \sqrt{|\partial_{\gamma}\phi|^2 + \varepsilon^2}$ , for some small constant  $\varepsilon$ . See [12] for more details. However, for small values of  $\varepsilon$  the equations are still nearly degenerate and for larger values the properties of the model get lost.

To overcome the non-differentiability of the term  $|\nabla\phi| + \rho|\partial_{\gamma}\phi|$  we employ its dual formulation [13–16]:

$$|\nabla\phi| + \rho|\partial_{\gamma}\phi| \equiv \max_{\mathbf{p}} \{\mathbf{p} \cdot \nabla_3\phi\} \quad \text{s.t.} \quad \sqrt{p_1^2 + p_2^2} \leq 1, \quad |p_3| \leq \rho, \quad (17)$$

where  $\mathbf{p} = (p^1, p^2, p^3)^T$  is the dual variable and  $\nabla_3$  is the full (three dimensional) gradient operator. This, in turn, leads us to the following primal-dual formulation

of the functional (13).

$$\min_{\phi \in D} \left\{ \max_{\mathbf{p} \in C} \left\{ \int_{\Sigma} \nabla_3 \phi \cdot \mathbf{p} \, d\Sigma \right\} \right\}, \quad (18)$$

where

$$C = \{\mathbf{p} : \Sigma \rightarrow \mathbb{R}^3 \mid \sqrt{p_1(\mathbf{x}, \gamma)^2 + p_2(\mathbf{x}, \gamma)^2} \leq 1, |p_3(\mathbf{x}, \gamma)| \leq \rho(\mathbf{x}, \gamma)\}. \quad (19)$$

Note that the primal-dual formulation is now continuously differentiable in both  $\phi$  and  $\mathbf{p}$ . In order to solve (18) we exploit a primal-dual proximal point method [17]. The idea of the proximal point method is to generate a sequence of approximate solutions by augmenting the functional by quadratic proximal terms for both the primal and dual variables. We first minimize the functional with respect to the primal variable and then maximize the functional with respect to the dual variable.

1. **Primal Step:** For fixed  $\mathbf{p}$ , compute a proximal primal step for  $\phi$ .

$$\phi^{k+1} = \min_{\phi \in D} \left\{ \int_{\Sigma} \nabla_3 \phi \cdot \mathbf{p}^k \, d\Sigma + \frac{1}{2\tau_p} \int_{\Sigma} (\phi - \phi^k)^2 \, d\Sigma \right\}. \quad (20)$$

2. **Dual Step:** For fixed  $\phi$ , compute a proximal dual step for  $\mathbf{p}$ .

$$\mathbf{p}^{k+1} = \max_{\mathbf{p} \in C} \left\{ \int_{\Sigma} \nabla_3 \phi^{k+1} \cdot \mathbf{p} \, d\Sigma - \frac{1}{2\tau_d} \int_{\Sigma} (\mathbf{p} - \mathbf{p}^k)^2 \, d\Sigma \right\}. \quad (21)$$

The parameters  $\tau_p$  and  $\tau_d$  denote the stepsizes of the primal and dual updates. We will now characterize the solutions of the alternating minimization scheme by the following two Propositions.

**Proposition 1** *The solution of (20) is given by*

$$\phi^{k+1} = \mathcal{P}_D (\phi^k + \tau_p \operatorname{div}_3 \mathbf{p}^k), \quad (22)$$

where  $\mathcal{P}_D$  denotes the projection onto the set  $D$ .

**Proof:** We compute the Euler-Lagrange equation of (20) which provides a necessary optimality condition for  $\phi$ .

$$-\operatorname{div}_3 \mathbf{p}^k + \frac{1}{\tau_p} (\phi - \phi^k) = 0. \quad (23)$$

Solving this equation for  $\phi$  directly leads to the presented scheme. Note, that the scheme does not ensure that  $\phi^{k+1} \in D$ . Therefore we have to reproject  $\phi^{k+1}$  onto  $D$  using the following Euclidean projector.

$$\mathcal{P}_D(\phi^{k+1}) = \min_{x \in D} \|\phi^{k+1} - x\|, \quad (24)$$

which can be computed by a simple truncation of  $\phi^{k+1}$  to the interval  $[0, 1]$  and setting  $\phi(\mathbf{x}, \gamma_{min}) = 1$  and  $\phi(\mathbf{x}, \gamma_{max}) = 0$ .  $\square$

**Proposition 2** *The solution of (21) is given by*

$$\mathbf{p}^{k+1} = \mathcal{P}_C(\mathbf{p}^k + \tau_d \nabla_3 \phi^{k+1}) , \quad (25)$$

where  $\mathcal{P}_C$  denotes the projection onto the set  $C$ .

**Proof:** The optimality condition for  $\mathbf{p}$  is given by

$$\nabla_3 \phi^{k+1} - \frac{1}{\tau_d} (\mathbf{p} - \mathbf{p}^k) = 0 . \quad (26)$$

We solve this equation for  $\mathbf{p}$  which results in the presented scheme. Since we need to ensure that  $\mathbf{p}^{k+1} \in C$  we reproject  $\mathbf{p}^{k+1}$  onto  $C$  using the Euclidean projector

$$\mathcal{P}_C(\mathbf{p}^{k+1}) = \min_{\mathbf{y} \in C} \|\mathbf{p}^{k+1} - \mathbf{y}\| , \quad (27)$$

which can be computed via

$$\begin{aligned} p_1^{k+1} &= \frac{p_1^{k+1}}{\max\{1, \sqrt{p_1^2 + p_2^2}\}} , \\ p_2^{k+1} &= \frac{p_2^{k+1}}{\max\{1, \sqrt{p_1^2 + p_2^2}\}} , \\ p_3^{k+1} &= \frac{p_3^{k+1}}{\max\{1, \frac{|p_3|}{\rho}\}} . \end{aligned} \quad (28)$$

□

### 3.3 Discretization

In our numerical implementation we are using a three-dimensional regular Cartesian grid

$$\{(i, j, k) \mid 1 \leq i \leq M, 1 \leq j \leq N, 1 \leq k \leq O\} , \quad (29)$$

where  $(i, j, k)$  is used to index the discrete locations on the grid,  $M$ ,  $N$  and  $O$  denote the size of the grid. We use standard forward differences to approximate the gradient operator

$$(\nabla_3 \phi)_{i,j,k} = \left( \frac{\phi_{i+1,j,k} - \phi_{i,j,k}}{\Delta x}, \frac{\phi_{i,j+1,k} - \phi_{i,j,k}}{\Delta y}, \frac{\phi_{i,j,k+1} - \phi_{i,j,k}}{\Delta \gamma} \right)^T , \quad (30)$$

and suitable backward differences to approximate the divergence operator

$$(\operatorname{div}_3 \mathbf{p})_{i,j,k} = \frac{p_{i,j,k}^1 - p_{i-1,j,k}^1}{\Delta x} + \frac{p_{i,j,k}^2 - p_{i,j-1,k}^2}{\Delta y} + \frac{p_{i,j,k}^3 - p_{i,j,k-1}^3}{\Delta \gamma} , \quad (31)$$

where  $\Delta x$ ,  $\Delta y$  denote the width of spatial discretization and  $\Delta \gamma$  denotes the width of the disparity discretization.



### 3.4 Convergence of the Algorithm

Currently we cannot prove explicit values for  $\tau_p$  and  $\tau_d$  which ensure convergence of the proposed algorithm. Empirically we observed that the algorithm converges as long as the product  $\tau_p\tau_d \leq 1/3$ . We therefore choose  $\tau_p = \tau_d = 1/\sqrt{3}$ .

### 3.5 Interpretation as Anisotropic Minimal Surfaces

In Section 3.1 we showed that the non-convex continuous multi-label problem (2) can be cast as a convex problem (13) by rewriting it in a higher dimensional space. We will now show that this higher-dimensional problem is that of a minimal surface problem in an anisotropic Riemannian space. Specifically, if we replace the anisotropic TV-like term  $|\nabla\phi| + \rho|\partial_\gamma\phi|$  by a weighted TV term  $\rho|\nabla_3\phi|$  we obtain a variational model whose minimizer is the minimal surface with respect to an isotropic Riemannian metric  $\rho$ .

$$\min_{\phi \in D} \left\{ \int_{\Sigma} \rho |\nabla_3 \phi| d\Sigma \right\}. \quad (32)$$

This problem has been studied in the context of Total Variation minimization by Bresson et al. in [18] and in the context of Continuous Maximal Flows by Appleton and Talbot in [19]. Note that the isotropic Riemannian problem does not allow for discontinuities in the solution whereas the anisotropic does.

### 3.6 Implementation

Numerical methods working on regular grids, can be effectively accelerated by state-of-the-art graphics processing units (GPUs). We employ the huge computational power and the parallel processing capabilities of GPUs to obtain a parallel implementation of our algorithm. The algorithm was implemented on a standard desktop PC equipped with a recent Quadcore 2.66 GHz CPU, 4 GB of main memory and a NVidia GeForce GTX 280 graphics card. The computer is running a 64-bit Linux system. With this GPU implementation we achieved a speedup factor of approximately **33** compared to an optimized C++ implementation executed on the same computer.

## 4 Experimental Results

In this section we provide experimental results of our algorithm applied to standard stereo benchmark problems. First, we compare our continuous formulation to the discrete approach of Ishikawa. Second we evaluate our method on the standard Middlebury stereo database [20]. Finally, we show results from a real world stereo example.

For stereo computation we need a data term measuring the matching quality of a rectified stereo image pair  $I_L$  and  $I_R$  for a certain disparity value  $\gamma$ . We

use the absolute differences summed over the three color channels of the input images.

$$\rho(\mathbf{x}, \gamma) = \lambda \sum_{i \in \{r, g, b\}} \left| I_L^{(i)}(\mathbf{x}) - I_R^{(i)}(\mathbf{x} + (\gamma, 1)^T) \right|. \quad (33)$$

#### 4.1 Comparison to Ishikawa’s approach

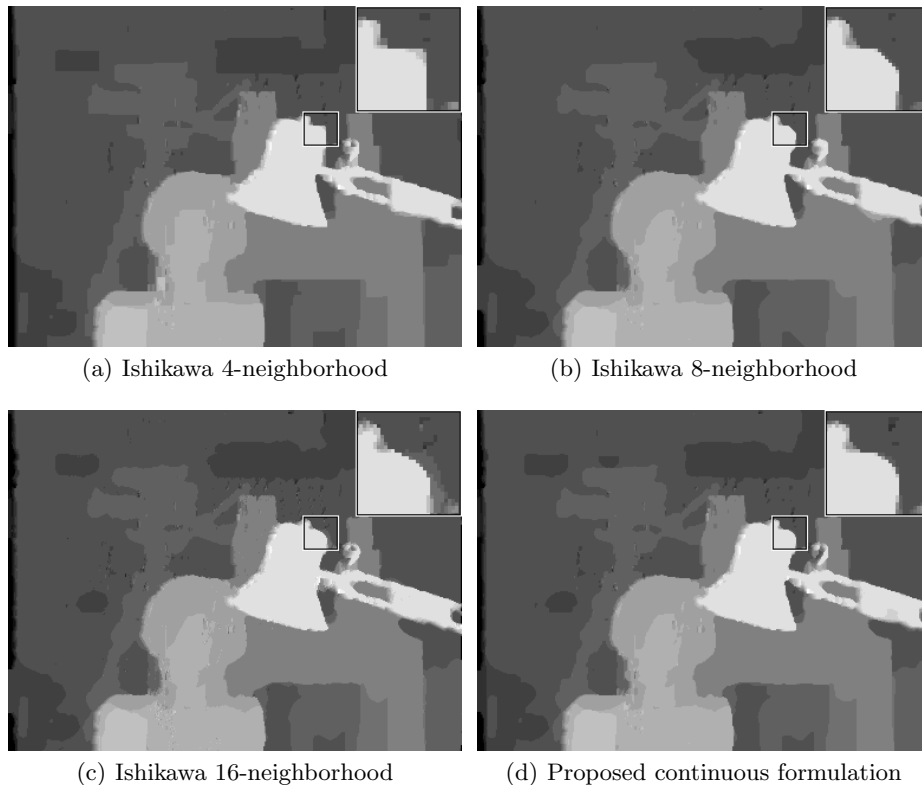
In our first experiment, we do a comparison of our continuous method to the discrete approach of Ishikawa using the Tsukuba data set [20]. According to [20], we used  $\gamma_{min} = 0$ ,  $\gamma_{max} = 16$  and  $\lambda = 50$ . The spatial domain and the disparity space was discretized using  $\Delta x = \Delta y = \Delta \gamma = 1.0$ . We ran our numerical scheme until the decrease of the energy was below a certain threshold. We also set up Ishikawa’s algorithm for different neighborhood connectivities. Since different neighborhood systems result in different weights of the smoothness term, we had to adjust the value of  $\lambda$  for the larger neighborhoods.

Fig. 1 shows a qualitative comparison of our continuous algorithm to Ishikawa’s discrete algorithm. In case of a 4-connected neighborhood, one can clearly see blocky structures in the solution. This effect, also known as metrication error, has its origin in the coarse approximation of the smoothness term when using a 4-connected or 8-connected neighborhood. We also provide a zoom in of the upper right corner of the lamp for the different results. In this region the metrication error of the discrete approach is clearly visible. In case of a 16-connected neighborhood, the result of the discrete approach is comparable to the result of our continuous approach.

Table 1 gives a quantitative comparison of our continuous algorithm to the discrete approach of Ishikawa using an error threshold of 1 for wrong pixels. It shows that the proposed continuous formulation provides error statistics slightly superior to its discrete counterpart. One can also see that both the runtime and the memory consumptions of Ishikawa’s discrete approach significantly increase with larger neighborhoods. Comparing our continuous approach to the 16-connected discrete approach of Ishikawa, we see that our GPU-based algorithm is about 20 times faster while requiring only 3.6% of its memory. This enables our method to compute the solution of stereo problems of much larger size in much shorter time.

**Table 1.** Quantitative comparison of the proposed continuous approach to Ishikawa’s discrete 16-connected approach. It shows that our GPU-based algorithm is about 20 times faster while requiring only 3.6% of its memory.

Algorithm	error (%)	Runtime CPU/GPU (sec)	Memory (MB)
Ishikawa 4-neighborhood	2.90	2.9 / -	450
Ishikawa 8-neighborhood	2.63	4.9 / -	630
Ishikawa 16-neighborhood	2.71	14.9 / -	1500
<b>Continuous formulation</b>	<b>2.57</b>	<b>25 / 0.75</b>	<b>54</b>

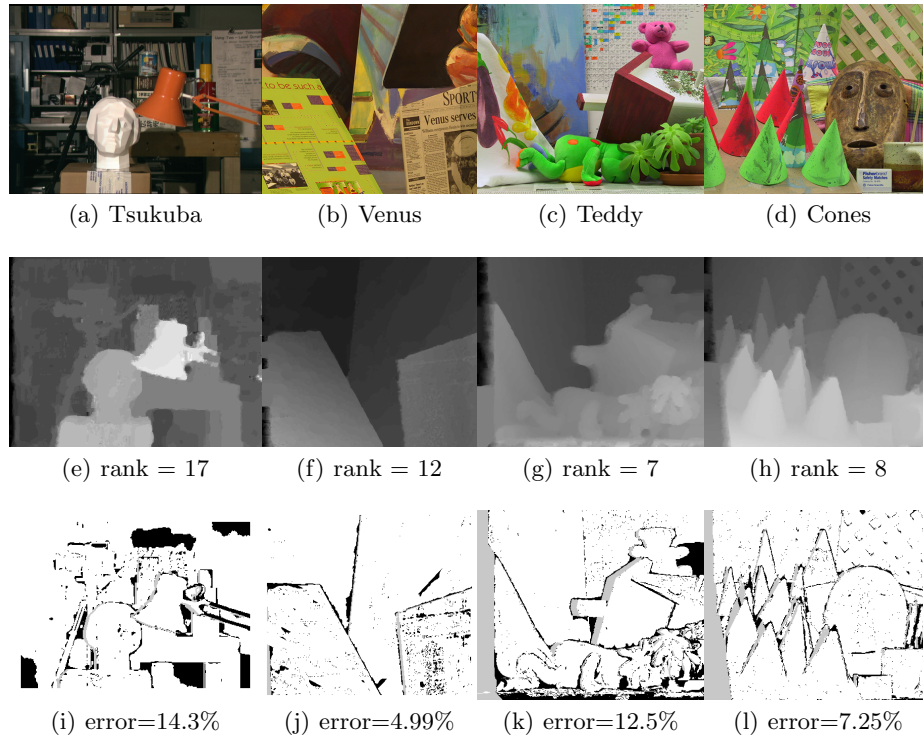


**Fig. 1.** Qualitative comparison of the proposed continuous approach to Ishikawa’s discrete approach. It clearly shows the metrication error in case of 4-connected and 8-connected neighborhoods, favoring 90 degree and 45 degree edges.

## 4.2 Evaluation on the Middlebury Stereo Database

In this section we provide a full evaluation of our algorithm on the standard Middlebury stereo database [20]. In order to be more insensitive to brightness changes in the input images we applied a high-pass filter to the input images before computing the data term. We ran our algorithm with the following constant parameter settings for the entire data base:  $\lambda = 30$ ,  $\Delta x = \Delta y = 1.0$  and  $\Delta\gamma = 0.5$ . The disparity range given by  $\gamma_{min}$  and  $\gamma_{max}$  was set according to [20]. The computing time in this setting varies between 15 seconds for the Tsukuba data set and 60 seconds for the Cones and Teddy data sets.

Fig. 2 shows the results of the stereo images. For a sub-pixel accurate threshold of  $th=0.5$ , our algorithm is currently ranked as number 15 out of 39 stereo algorithms. Note that Ishikawa’s algorithm failed in this setting due to its immense memory requirements.



**Fig. 2.** Quantitative results from the Middlebury stereo evaluation data base.

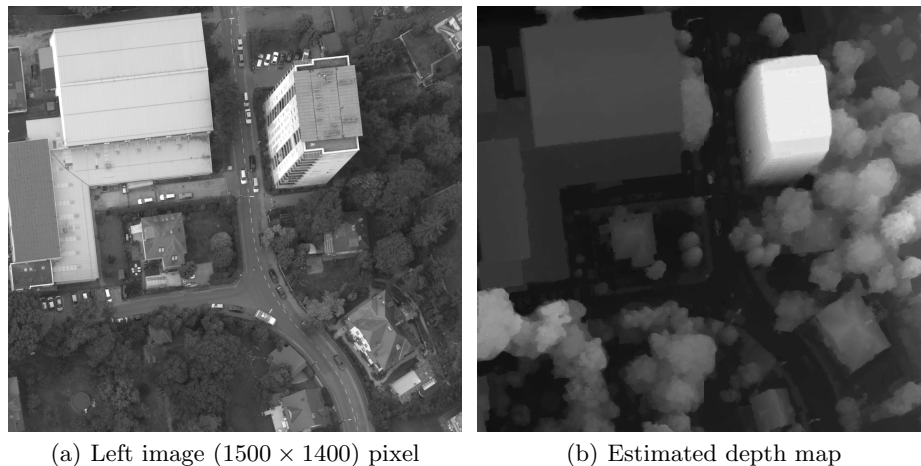
one should keep in mind that more sophisticated algorithms may provide better quantitative results for the stereo problems. However, our variational model is very simple and does not take into account additional information from image segmentation, plane-fitting and consistency checks. More importantly, our model can be exactly solved, which is not the case for the more sophisticated approaches.

### 4.3 Real World Example

Finally we give results of our algorithm applied to a real world stereo problem. Fig. 3 shows the estimated depth map from a large aerial stereo pair from Graz. We ran our algorithm with the following parameter settings:  $\lambda = 50$ ,  $\gamma_{min} = -30$ ,  $\gamma_{max} = 30$ ,  $\Delta x = \Delta y = 1.0$  and  $\Delta\gamma = 0.5$ . This example shows that the proposed algorithm yields promising results for large practical problems.

## 5 Conclusion

In this paper we proposed a continuous formulation to the discrete multi-label problem of Ishikawa. We showed that the original non-convex problem can be



**Fig. 3.** Estimated depth map of the proposed algorithm applied to a large aerial stereo data set of Graz.

reformulated as a convex problem via embedding into a higher dimensional space. Our formulation removes several shortcomings of Ishikawa’s discrete approach. First, our algorithm is defined in a spatially continuous setting and is therefore free from grid bias. Second, our algorithm is based on variational optimization techniques which can be easily parallelized. Finally our algorithm needs less memory enabling us to compute much larger problems. Results from practical stereo examples emphasize the advantages of our approach over the discrete approach of Ishikawa.

For future work we see mainly two directions. One direction is to investigate more sophisticated optimization schemes to achieve an additional speedup in computing the minimizer of the convex formulation. The other direction is to improve the variational model for stereo estimation. Specifically, we plan to incorporate additional cues such as edges into our variational model while still allowing to compute its exact solution.

## 6 Acknowledgements

This work was supported by the Hausdorff Center for Mathematics and the Austrian Research Promotion Agency within the VM-GPU project (no. 813396). We would also like to thank Microsoft Photogrammetry for providing us the aerial stereo images.

## References

1. Greig, D., Porteous, B., Seheult, A.: Exact maximum a posteriori estimation for binary images. *J. Royal Statistics Soc.* **51**(Series B) (1989) 271–279

2. Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(2) (2004) 147–159
3. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(11) (2001) 1222–1239
4. Veksler, O.: Efficient Graph-based Energy Minimization Methods in Computer Vision. PhD thesis, Cornell University (July 1999)
5. Schlesinger, D., Flach, B.: Transforming an arbitrary minsum problem into a binary one. Technical Report TUD-FI06-01, Dresden University of Technology (2006)
6. Werner, T.: A linear programming approach to max-sum problem: A review. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(7) (2007) 1165–1179
7. Ishikawa, H.: Exact optimization for markov random fields with convex priors. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(10) (2003) 1333–1336
8. Ford, L., Fulkerson, D.: *Flows in Networks*. Princeton University Press, Princeton, New Jersey (1962)
9. Rudin, L., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Physica D* **60** (1992) 259–268
10. Chan, T., Esedoglu, S., Nikolova, M.: Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM Journal of Applied Mathematics* **66**(5) (2006) 1632–1648
11. Fleming, W., Rishel, R.: An integral formula for total gradient variation. *Arch. Math.* **11** (1960) 218–222
12. Vogel, C., Oman, M.: Iteration methods for total variation denoising. *SIAM J. Sci. Comp.* **17** (1996) 227–238
13. Chan, T., Golub, G., Mulet, P.: A nonlinear primal-dual method for total variation-based image restoration. *SIAM J. Sci. Comp.* **20**(6) (1999) 1964–1977
14. Carter, J.: *Dual Methods for Total Variation-based Image Restoration*. PhD thesis, UCLA, Los Angeles, CA (2001)
15. Chambolle, A.: An algorithm for total variation minimizations and applications. *J. Math. Imaging Vis.* (2004)
16. Chambolle, A.: Total variation minimization and a class of binary MRF models. In: *Energy Minimization Methods in Computer Vision and Pattern Recognition*. (2005) 136–152
17. Rockafellar, R.: Augmented lagrangians and applications of the proximal point algorithm in convex programming. *Math. of Oper Res* **1** (1976) 97–116
18. Bresson, X., Esedoglu, S., Vandergheynst, P., Thiran, J., Osher, S.: Fast global minimization of the active contour/snake model. *J. Math. Imaging Vis.* **28**(2) (2007) 151–167
19. Appleton, B., Talbot, H.: Globally minimal surfaces by continuous maximal flows. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(1) (2006) 106–118
20. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comp. Vis.* **47**(1-3) (2002) 7–42