Variational Methods in Computer Vision
ICCV Tutorial, 6.11.2011

Chapter 6

Vectorial Total Variation and Multilabel Problems

Daniel Cremers and Bastian Goldlücke
Computer Vision Group
Technical University of Munich

Thomas Pock
Institute for Computer Graphics and Vision
Graz University of Technology

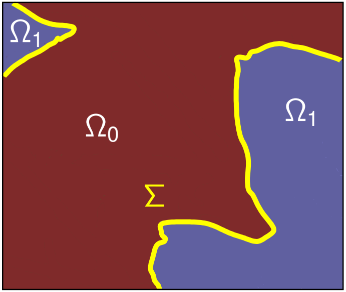
Overview

1. **Vectorial Total Variation**
   - Reminder: TV and binary segmentation
   - Generalizations of the total variation
   - Generalizations of the total variation
   - Analytic properties
   - Geometric properties
2. **Multilabel segmentation**
   - The multilabel problem
   - Regularization
3. **Product Label Spaces**
4. **Summary**

Overview

1. **Vectorial Total Variation**
   - Reminder: TV and binary segmentation
   - Generalizations of the total variation
   - Generalizations of the total variation
   - Analytic properties
   - Geometric properties
2. **Multilabel segmentation**
   - The multilabel problem
   - Regularization
3. **Product Label Spaces**
4. **Summary**

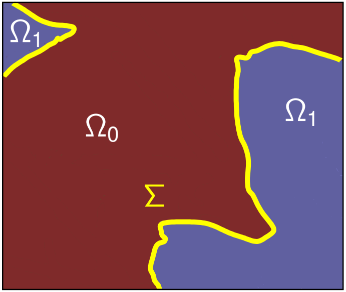
Binary Segmentation



Region Ω_0 (background)

Region Ω_1 (flower)

Binary Segmentation



Region Ω_0 (background)

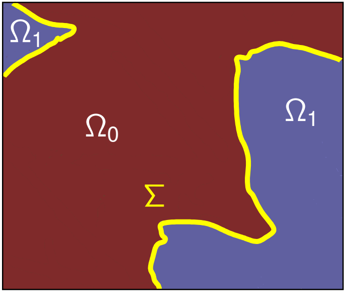
Region Ω_1 (flower)

Find binary labeling $u : \Omega \to \{0,1\}$ which minimizes

$$\underbrace{\int_\Omega |\nabla u|_2 \, dx}_{\text{length of interface } \Sigma} + \underbrace{\int_\Omega c_1 \cdot u \, dx.}_{\text{assignment cost}}$$

$c_1(x)$ = cost of assigning “1” to the point $x \in \Omega$.

Binary Segmentation



Region Ω_0 (background)

Region Ω_1 (flower)

Find binary labeling $u : \Omega \to \{0, 1\}$ which minimizes

$$\underbrace{\int_\Omega |\nabla u|_2 \, dx}_{\text{length of interface } \Sigma} + \underbrace{\int_\Omega c_1 \cdot u \, dx.}_{\text{assignment cost}}$$

$c_1(x)$ = cost of assigning "1" to the point $x \in \Omega$.

Can be minimized globally (Chan, Esedoglu and Nikolova 2006), as shown in an earlier chapter of the tutorial.

Vectorial Total Variation

First goal in this chapter: introduce **total variation for vector-valued functions** which has a similar geometric interpretation, and can be used to define a regularizer for multi-label problems.

Reminder: scalar total variation

For a greyscale image $u : \Omega \to \mathbb{R}$ on a domain $\Omega \subset \mathbb{R}^m$, the scalar total variation (TV) is defined as

$$\mathrm{TV}(u) = \underbrace{\int_\Omega |\nabla u|_2 \, dx}_{\text{primal and}} = \underbrace{\sup_{\xi \in \mathcal{C}_c^1(\Omega, \mathbb{E}(m))} \left\{ \int_\Omega u \, \mathrm{div}(\xi) \, dx \right\}}_{\text{dual formulation}},$$

where $\mathbb{E}(m)$ is the unit ball in $\mathbb{R}^m$.

Requirements for the generalization to u $: \Omega \to \mathbb{R}^n$**:**

- Definitions coincide for $n = 1$
- Dual formulation available, so that it is defined for non-differentiable functions
- Convex, closed $\implies$ efficient minimization algorithms available
- Important invariances and other properties of scalar TV still satisfied

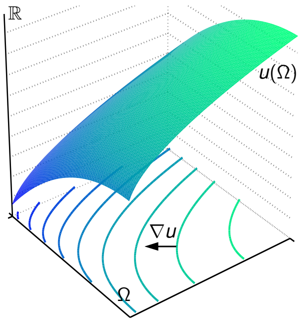
Comparisons: Goldlücke and Cremers, CVPR 2010 and Goldlücke, Strekalovskiy and Cremers, 2011 (under review)

Sapiro's general approach: image manifold

- The metric tensor of the image manifold $u(\Omega)$ is given by

$$(g_{\mu\nu}) = \underbrace{(D\mathbf{u})^T D\mathbf{u}}_{m \times m} .$$

- The Eigenvector corresponding to the largest Eigenvalue λ_1 gives the direction of the vectorial edge.
- $n = 1$: Equal to direction of the gradient ∇u, which is always orthogonal to the level lines.



Leads to familiy of possible definitions for the vectorial TV in the case $m = 2$, which is of the form

$$\mathrm{TV}_{\mathrm{SR}}(\mathbf{u}) := \int_{\Sigma} \varphi(\lambda_1, \ldots, \lambda_n)\, ds,$$

where φ is a suitable scalar-valued function (Sapiro and Ringach, 1996).

Generalizations with dual formulation

Variant	*Primal*	*Dual*
$\mathrm{TV}_S(\mathbf{u})$	$\sum_{i=1}^{n} \int_\Omega \lvert \nabla u_i \rvert_2 \, \mathrm{d}x$	$\sup_{(\xi_1,\dots,\xi_n)\in K_S} \left\{ \sum_{i=1}^{n} \int_\Omega u_i \, \mathrm{div}(\xi_i) \, \mathrm{d}x \right\}$ with $K_S = \mathcal{C}_c^1(\Omega, \mathbb{E}(m) \times \dots \times \mathbb{E}(m))$
$\mathrm{TV}_F(\mathbf{u})$	$\int_\Omega \lVert D\mathbf{u}(x) \rVert_F \, \mathrm{d}x$	$\sup_{(\xi_1,\dots,\xi_n)\in K_F} \left\{ \sum_{i=1}^{n} \int_\Omega u_i \, \mathrm{div}(\xi_i) \, \mathrm{d}x \right\}$ with $K_F = \mathcal{C}_c^1(\Omega, \mathbb{E}(n \cdot m))$
$\mathrm{TV}_J(\mathbf{u})$	$\int_\Omega \sqrt{\lambda_1} \, \mathrm{d}x$	$\sup_{(\xi_1,\dots,\xi_n)\in K_J} \left\{ \sum_{i=1}^{n} \int_\Omega u_i \, \mathrm{div}(\xi_i) \, \mathrm{d}x \right\}$ with $K_J = \mathcal{C}_c^1(\Omega, \mathrm{co}\,(\mathbb{E}(m) \otimes \mathbb{E}(n)))$

Comparison: Goldlücke and Cremers, CVPR 2010

Frobenius TV

For this tutorial chapter, we choose the Frobenius TV as the vectorial total variation. The primal definition for differentiable **u** is

$$\int_\Omega \|D\mathbf{u}(x)\|_F \, \mathrm{d}x = \int_\Omega \sqrt{\sum_{i,j} \left(\frac{\partial u_i}{\partial x_j}\right)^2} \mathrm{d}x$$

$$= \int_\Omega \sqrt{\lambda_1 + \ldots + \lambda_n} \mathrm{d}x.$$

The latter equality can be checked by substituting the SVD of D**u**. This corresponds to the dual definition

$$\sup_{(\xi_1,\ldots,\xi_n) \in K_F} \left\{ \sum_{i=1}^n \int_\Omega u_i \operatorname{div}(\xi_i) \; \mathrm{d}x \right\}$$

$$\text{with } K_F = \mathcal{C}_c^1(\Omega, \mathbb{E}(n \cdot m)).$$

Analytic properties

The analytic properties can be verified directly from the definition, as in the scalar case.

Proposition

- TV_F is a semi-norm, in particular it is convex.
- TV_F is lower semi-continuous (closed).

Geometric properties

TV_F has a geometric property similar to the scalar TV with regard to curve length. It allows to construct very general regularizers for multilabel segmentation problems.

Theorem

Let $S \subset \Omega$ and $\bar{S} := \Omega \setminus S$. Furthermore, let $a, b \in \mathbb{R}^k$. Then

$$\mathrm{TV}_F(a\,1_S + b\,1_{\bar{S}}) = |a - b|_2\,\mathrm{Per}(S).$$

Note that this is a generalization of the scalar case, since

$$\mathrm{TV}(1_S) = \mathrm{TV}(1 \cdot 1_S + 0 \cdot 1_{\bar{S}}) = |1 - 0|_2\,\mathrm{Per}(S) = \mathrm{Per}(S).$$

see Lellmann et al., ICCV 2009

Overview

1. **Vectorial Total Variation**
   - Reminder: TV and binary segmentation
   - Generalizations of the total variation
   - Generalizations of the total variation
   - Analytic properties
   - Geometric properties
2. **Multilabel segmentation**
   - The multilabel problem
   - Regularization
3. **Product Label Spaces**
4. **Summary**

The Multilabel Problem

Interpretation as segmentation problem

$$\Omega = \Omega_1 \cup \ldots \cup \Omega_N$$

- Region Ω_0 (background)
- Region Ω_1 (leaves)
- Region Ω_2 (flower 1)
- Region Ω_3 (flower 2)

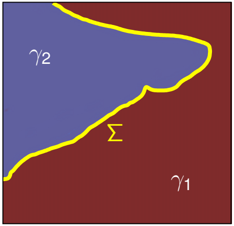
The Multilabel Problem

Interpretation as labeling problem

$$g : \Omega \to \{\gamma_1, ..., \gamma_N\}$$

- Label γ_0 (background)
- Label γ_1 (leaves)
- Label γ_2 (flower 1)
- Label γ_3 (flower 2)

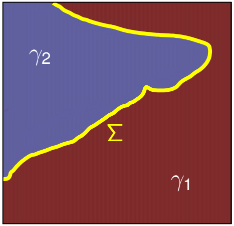
Regularization



The regularization penalty is propotional to the **label distance** times the **length of the interface**.

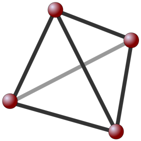
In this example $d(\gamma_1, \gamma_2) \cdot L(\Sigma)$

Regularization



The regularization penalty is propotional to the **label distance** times the **length of the interface**.

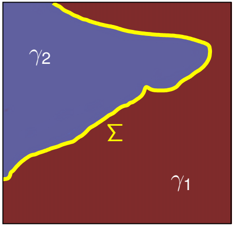
In this example $d(\gamma_1, \gamma_2) \cdot L(\Sigma)$



Euclidean representation of the label distance:

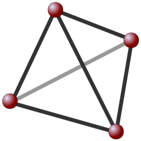
- Each label γ is *represented* by a point $a_\gamma \in \mathbb{R}^k$.
- Label distance $d(\gamma, \mu) = |a_\gamma - a_\mu|_2$.

Regularization



The regularization penalty is propotional to the **label distance** times the **length of the interface**.

In this example $d(\gamma_1, \gamma_2) \cdot L(\Sigma)$



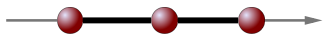
Euclidean representation of the label distance:

- Each label γ is *represented* by a point $a_\gamma \in \mathbb{R}^k$.
- Label distance $d(\gamma, \mu) = |a_\gamma - a_\mu|_2$.

Except for a special case of the embedding, the problem is **currently unsolvable** (discrete case: NP-hard).

Important special cases

Assume the labels are numbered, $\Gamma = \{1, ..., N\}$.



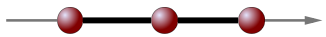
$a_\gamma = \gamma \in \mathbb{R}$

Ordered Labels

- Example: depth reconstruction
- Can be solved globally with functional lifting [Pock, Schönemann, Graber, Bischof, Cremers '08]
- Continuous version of [Ishikawa '03]

Important special cases

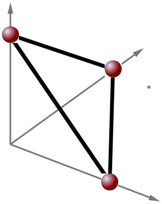
Assume the labels are numbered, $\Gamma = \{1, ..., N\}$.



$a_\gamma = \gamma \in \mathbb{R}$

Ordered Labels

- Example: depth reconstruction
- Can be solved globally with functional lifting [Pock, Schönemann, Graber, Bischof, Cremers ’08]
- Continuous version of [Ishikawa ’03]



$a_\gamma = e_\gamma \in \mathbb{R}^N$

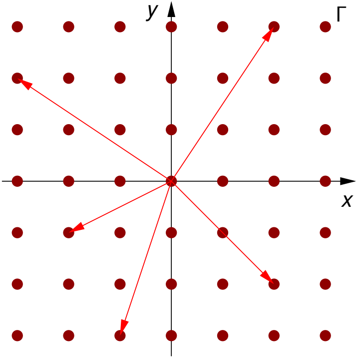
Potts model

- Example: segmentation
- No globally optimal solution possible if $N > 2$
- Continuous version of [Potts ’52]

Optic Flow

Color input images
$I_0, I_1 : \Omega \to \mathbb{R}^3$:

Label each pixel in I_0 with a flow vector in $\Gamma \subset \mathbb{R}^2$, choose representation $a_\gamma = \gamma$.

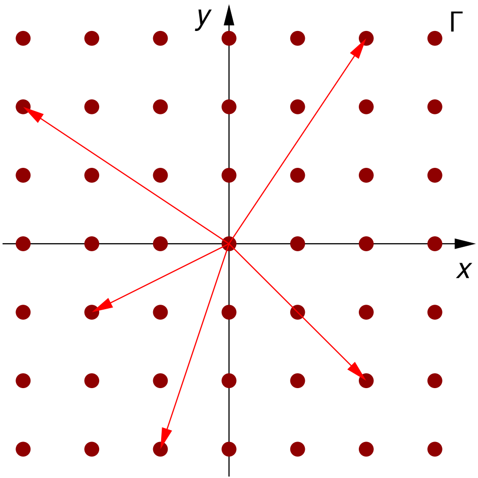


Cost function compares pointwise pixel colors in the images:

$$c_\gamma(x) = |I_0(x) - I_1(x + \gamma)|_2$$

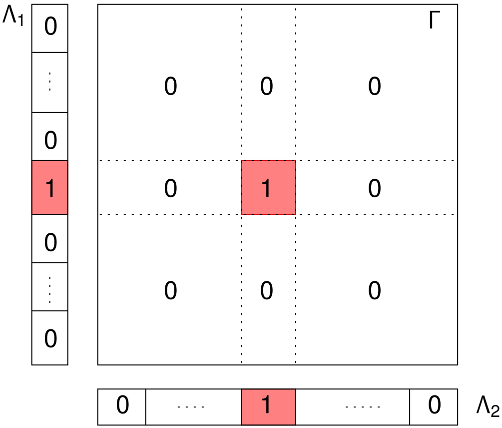
Overview

The optic flow label space is a product space

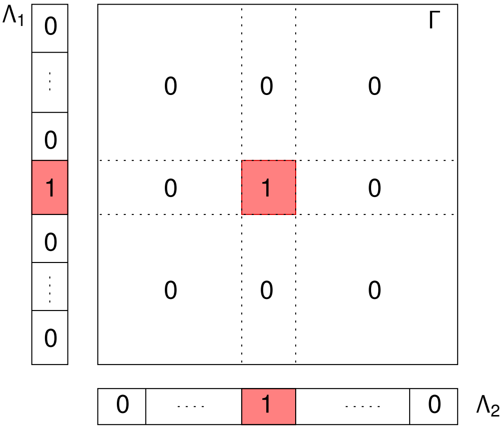


Each red dot requires one indicator function - **too many.**
Can we exploit the special structure of the label space?

Reduction idea for product label spaces



Reduction idea for product label spaces



Indicator functions are products $u_\gamma = u^1_{\lambda^1} \cdot u^2_{\lambda^2}$.
Goldluecke and Cremers, ECCV 2010

Relaxation of the products

The data term is now non-convex:

$$E(\mathbf{u}_1, \mathbf{u}_2) = \sum_{\gamma \in \Gamma} \left\langle c^\gamma, u^1_{\lambda^1} \cdot u^2_{\lambda^2} \right\rangle .$$

We show in Strekalovskiy, Goldluecke, Cremers, ICCV 2011 that a convex relaxation of the data term is given by

$$R(u) = \sup_{q^1_{\lambda^1} + q^2_{\lambda^2} \leq c^\gamma} \left\{ \sum_{\lambda_1} \left\langle q^1_{\lambda^1}, u^1_{\lambda^1} \right\rangle + \sum_{\lambda_2} \left\langle q^2_{\lambda^2}, u^2_{\lambda^2} \right\rangle \right\}$$

It has the two necessary properties:

- $R(u) = E(u)$ for binary u.
- If $\hat{u}$ is a binary minimizer of E, then $\hat{u}$ also minimizes R in the relaxed space.

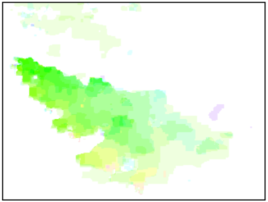
Runtime and memory requirements

# of Pixels	# Labels	Memory [Mb]		Run time [s]	
$P = P_x \times P_y$	$N_1 \times N_2$	Previous	Proposed (g/p)	Previous	Proposed (g/p)
320 × 240	8 × 8	112	112 / 102	196	26 / 140
320 × 240	16 × 16	450	337 / 168	*	80 / 488
320 × 240	32 × 32	1800	1124 / 330	*	215 / 1953
320 × 240	50 × 50	4394	2548 / 504	*	950 / 5188
320 × 240	64 × 64	7200	4050 / 657	-	1100 / 8090
640 × 480	8 × 8	448	521 / 413	789	102 / 560
640 × 480	16 × 16	1800	1351 / 676	*	295 / 1945
640 × 480	32 × 32	7200	4502 / 1327	-	1290 / 7795
640 × 480	50 × 50	17578	10197 / 2017	-	- / 32887
640 × 480	64 × 64	28800	16202 / 2627	-	- / 48583

Multilabel optic flow (1)

First image I_0

Second image I_1



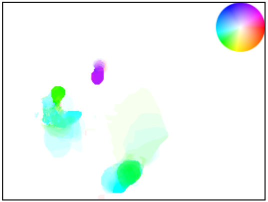
Result

32 $\times$ 32 labels, image resolution 320 $\times$ 240, TV regularity
2 minutes runtime, within 5% of global optimum.

Multilabel optic flow (2)

First image I_0

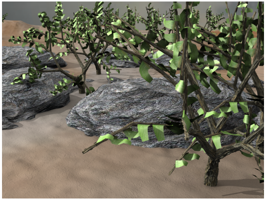
Second image I_1



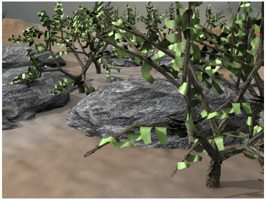
Result

32×32 labels, image resolution 320×240, TV regularity
2 minutes runtime, within 5% of global optimum.

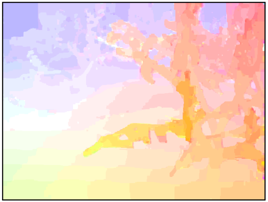
Multilabel optic flow (3)



First image I_0



Second image I_1



Result

32×32 labels, image resolution 640×480, truncated TV regularity
15 minutes runtime, within 6% of global optimum.

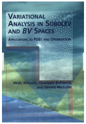
Summary

- Vectorial total variation extends the definition of TV from scalar- to vector-valued functions.
- A common and useful generalization is the Frobenius-TV. In the primal formulation, you integrate over the Frobenius norm of the derivative matrix (Jacobian).
- Frobenius-TV is closed and convex, so it can be minimized efficiently. Furthermore, it has a similar geometric property that the scalar TV with regards to jump functions.
- Vectorial TV can be used to construct functionals for multilabel problems with convex relaxations available.
- In the case of product label spaces, the memory and runtime requirements can be drastically reduced.

See our poster on Thursday,
"Tight convex relaxations for vector-valued labeling",
for continuous label spaces and more general regularizers.

References (1)

Vectorial Total Variation



Attouch, Buttazzo and Micaille,
"Variational Analysis in Sobolev and BV spaces",
SIAM 2006.

- Exhaustive introduction to variational methods and convex optimization in infinite dimensional spaces, as well as the theory of BV functions.
- Mathematically very advanced, requires solid knowledge of functional analysis.

Goldlücke and Cremers,
"An Approach to Vectorial Total Variation based on Geometric Measure Theory",
CVPR 2010.

- Classification and comparison of several extensions of TV to vector valued function.
- Evaluation of the cases with a dual formulation available.

References (2)

VTV and Multilabel Problems

Lellmann, Becker and Schnörr,
"Convex Optimization for Multi-Class Image Labeling with a Novel Family of Total Variation Based Regularizers",
ICCV 2009.


Convex Optimization for Multi-Class Image Labeling with a Novel Family of Total Variation Based Regularizers

- Introduction of a certain convex relaxation for multilabel problems
- VTV to define regularizers with Euclidean representations for the label distance.

Goldlücke and Cremers,
"Convex Relaxation for Multilabel Problems with Product Label Spaces",
ECCV 2010.


Convex Relaxation for Multilabel Problems with Product Label Spaces

Bastian Goldluecke and Daniel Cremers

Computer Vision Group, TU Munich

- Reduction technique for label spaces with product structure.
- Makes the algorithm feasible for very large problems like optic flow with thousands of labels.