

Contents

1	Introduction	3
2	Background	8
2.1	Gibbs Random Fields	8
2.2	Lie Groups and the Noether Theorem	12
2.2.1	Motivation 1, the problem	12
2.2.2	Motivation 2, the solution	13
2.3	Lie Groups	16
2.3.1	The Group $\mathbb{G} = \mathbb{T} \times SO(2)$	21
2.4	Noether's First Theorem	22
2.4.1	Noethers Theorems	23
2.4.2	Noether's First Theorem: A Modern Version	26
2.5	Total Variation	29
2.5.1	The Mean Curvature of Total Variation	31
2.6	Optical Flow	33
2.6.1	Uni-Modal Optical Flow	36
2.6.2	Multi-Modal Optical Flow	36
2.7	Image Fusion	41
3	Linearized Priors	45
3.1	The Linear Structure Tensor	45
3.2	Structure Tensor Based Prior	47
3.3	Geometrical Optical Flow Model	48
3.4	Disparity	48
3.5	Localization	51
3.6	The solution algorithm	54
3.7	Results	55
3.8	Uni-Modal Data	55
3.8.1	Structure Tensor Prior	56
3.8.2	Total Variation Prior	56
3.9	Eigenvalue analysis and the stabilization parameter λ_2	67
3.10	Multimodal Optical Flow	68
3.10.1	Estimation of the resolution parameter σ	68
3.10.2	Structure Tensor Prior	68

3.10.3 Total Variation Prior	68
4 The Generalized Newton Algorithm	71
4.1 Motivation	71
4.2 The Generalized Newton Algorithm	76
4.2.1 Image De-noising	81
5 Conclusions	88
Bibliography	89

1 Introduction

The main topic of this thesis is concerned about symmetries in the mathematical modeling of computer vision problems. Many objects in nature possess among others the notable characteristic of symmetry regarding their attributes such as their form and color. A symmetry of an object \mathcal{O} is such that if \mathcal{O} undergoes a specific transformation g , then it appears for an observer to be unchanged. Say we have a computer vision problem involving the object \mathcal{O} , modeled with a mathematical model M . It is natural to reflect the symmetry of the object \mathcal{O} within the model M , such that M is invariant in some sense under the transformation g . The goal of this thesis is to analyze the structure of the symmetries of a mathematical model M . We will prove that knowledge of the symmetries of M may lead to significant speed ups of any algorithm using M .

Symmetries generally fall into two categories: global and local symmetries. A ball of uniform color for instance does not change its appearance to an observer upon rotation around an arbitrary axis through the center of the ball. This example is one of *global* symmetry since the ball as a whole is transformed (rotated). We can formally describe the global symmetry of the object \mathcal{O} in the following way: If the surface of the object is described by the functional relationship $\phi_{\mathcal{O}}(\mathbf{x}) = \text{const}$ (e.g. $\phi_{\mathcal{O}}(\mathbf{x}) = x^2 + y^2 + z^2 = 1$ for a ball of unit radius) then our intuition of global symmetry is equal to the statement that $\phi_{\mathcal{O}}(\mathbf{x}) = \text{const}$ is invariant under the global transformation $\mathbf{x}' = g \circ \mathbf{x}$

$$\phi_{\mathcal{O}}(g \circ \mathbf{x}) = \phi_{\mathcal{O}}(\mathbf{x}) \tag{1.1}$$

Not all objects in nature are symmetric with respect to global transformations. For example in figure 1.1 an image of a leaf is shown. Since the leaf is not symmetric with respect to any global transformation g , its projection onto the image plane Ω is not symmetric with respect to any global transformation g^{Ω} on Ω . However if we inspect local regions of the leaf, that is we zoom into those regions at various locations on the leaf, we see that the features of the leaf within the regions do possess symmetries. Figure 1.1b shows a close up of the region highlighted in figure 1.1a through which a vein of the leaf runs. The vein appears to be linear and thus symmetric towards translations along its tangential direction. This symmetry is reflected by the vectors at each position of the vein. They indicate

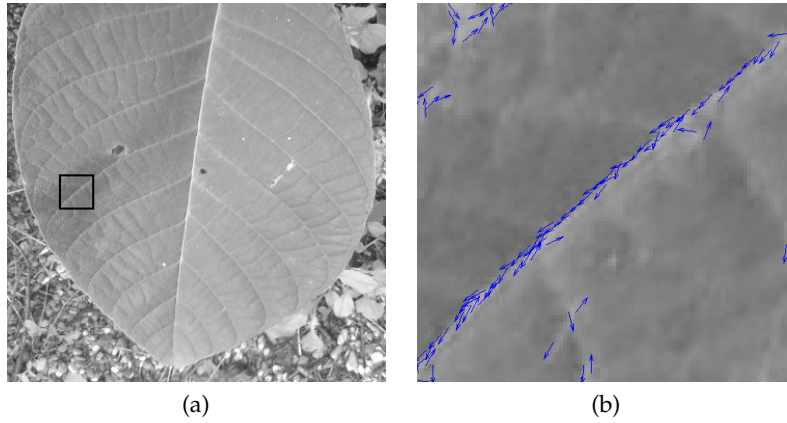


Figure 1.1: Figure 1.1a shows an image of a leaf. The leaf clearly has no global symmetry. Figure 1.1b shows a close-up of the region around a vein of the leaf, indicated by the box in figure 1.1a. The vectors in figure 1.1b along the vein indicate local translations which leave the vein invariant.

local translations, which leave the vein invariant. A local transformation as indicated by the vectors in figure 1.1b may be represented by the vector field $\omega^\Omega(\mathbf{x})$ such that the local transformation $g^\Omega(\mathbf{x}) = \mathbf{x} + \omega^\Omega(\mathbf{x})$ leaves the image ϕ invariant

$$\phi(\mathbf{x} + \omega^\Omega(\mathbf{x})) = \phi(\mathbf{x}) \quad (1.2)$$

In general we cannot assume that $g^\Omega(\mathbf{x})$ in eq. (1.2) is unique since there can always exist a vector field $\omega^{\Omega'}(\mathbf{x}) \neq \omega^\Omega(\mathbf{x})$ which satisfies eq. (1.2). On the other side any transformation g^Ω satisfying eq. (1.2) uniquely determines the geometry of ϕ for if we were to draw lines along the tangential vectors $\omega^\Omega(\mathbf{x})$ by connecting \mathbf{x} with $\mathbf{x} + \omega^\Omega(\mathbf{x})$ we would reconstruct the object \mathcal{O} from $g^\Omega(\mathbf{x})$.

The process of acquiring information from our physical reality is problematic itself in many ways. For one, the information which we may wish to gather may lay hidden in the data we can possibly acquire from a physical system. On such problem is called stereography ([?]), depicted in figure 1.2. The statement of the problem goes as follows: given two images y and I (figures 1.2b and 1.2c) of an object \mathcal{O} (the box in figure 1.2a) how can we infer the 3-dimensional structure of \mathcal{O} (the width, height and depth of the box)? This problem has already been solved by nature since the human brain capable of reconstructing a 3-dimensional image given the 2-dimensional images obtained by the left and the right eye.

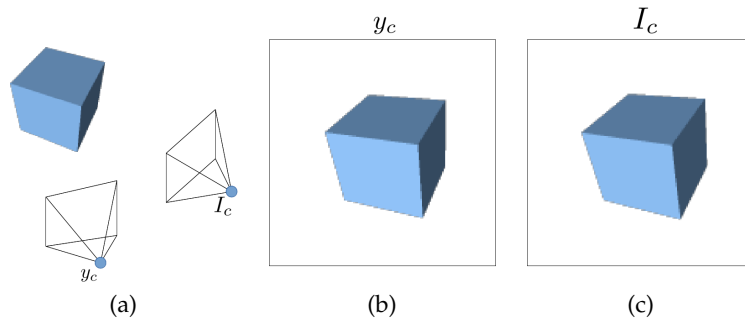


Figure 1.2: Figure 1.2a: Two cameras are shown recording a scene from different positions. The scene could be a rigid scene or a dynamic scene with moving objects. Figure 1.2b shows the image y captured from the camera y_c and figure 1.2c the image I from the camera I_c . One possible question is: How can the pixels of the image I be mapped to those of the image y ? Such a mapping can be used to deduce the 3-dimensional structure of the box similar to how the human brain constructs a 3-dimensional image given the 2-dimensional images obtained by the left and the right eye.

Besides the problem of hidden information described above there is another problem in the process of information acquisition. The means we use to acquire the data have technical limitations. For instance the cameras y_c and I_c in figure 1.2a in general produce images of limited resolutions which may also be subject to noise.

Both problems in the process of information acquisition may be sub-summed as the problem of *inference*: Given some possibly corrupted data Y of a physical system we wish to infer some information stored in the unknown latent variable ϕ . In general Y and ϕ may be discrete variables, continuous functions over some domain Ω or a combination of both. In this thesis we will only handle problems for which Y and ϕ are continuous functions over Ω

$$Y, \phi: \Omega \rightarrow \mathbb{R}^n \quad (1.3)$$

The inference problem then becomes the problem of mapping Y to ϕ

$$Y(\mathbf{x}) \xrightarrow{T_Y} \phi(\mathbf{x}) \quad (1.4)$$

where T_Y denotes a process or an algorithm which is parametrized by the data Y . Since the variable ϕ is unknown we have to for one make assumptions on its geometric properties and furthermore model how it is linked to the data Y .

These aspects of ϕ are then embedded in the inference process T_Y . For now we want to motivate how the geometrical properties of ϕ can be taken into account by T_Y . Consider a local transformation g such that the variable ϕ is transformed to the new variable ϕ'

$$\phi'(\mathbf{x}) = \phi(g \circ \mathbf{x}) \quad (1.5)$$

We can regard ϕ' as being inferred from the data Y via the inference process T_Y' similar to eq. (1.4). If ϕ is symmetric under g in the sense of eq. (1.2) then this implies that the two inference processes T_Y and T_Y' are equal and thus the inference process T_Y is itself symmetric under the action of g . We conclude that knowledge of the set of local transformations $\{g\}$ which satisfy eq. (1.2) allows us to identify those inference processes T_Y which are equal to each other upon action of $\{g\}$. This has two consequences. The first is that we can *design* an inference process T_Y which is invariant upon the action of the set $\{g\}$. As a result this guarantees the invariance of ϕ upon the action of $\{g\}$. The second consequence is more subtle. If we split the inference process T_Y into n intermediate steps

$$Y \xrightarrow{T_Y} \phi = Y \xrightarrow{T_Y^1} \phi^1 \xrightarrow{T_Y^2} \phi^2 \dots \xrightarrow{T_Y^{n-1}} \phi^{n-1} \xrightarrow{T_Y^n} \phi \quad (1.6)$$

the intermediate steps T_Y^i and ϕ^i need *not* be invariant under the set $\{g\}$. However for particularly well chosen $g' \in \{g\}$ such that

$$g' \circ T_Y^i = T_Y^{i+k} \quad (1.7)$$

we may minimize the number steps in eq. (1.6) and thus obtain the *shortest path* in the inference problem.

The overall structure of this thesis is as follows: In section 2.1 we introduce the latent variable ϕ as a Gibbs Random Field (GRF). The main property of GRFs is that they are associated with an energy functional $E_Y(\phi)$. The inference process T_Y is explicitly formulated as the minimization problem

$$\phi^* = \operatorname{argmin}_{\phi} E_Y(\phi) \quad \leftrightarrow \quad Y \xrightarrow{T_Y} \phi^* \quad (1.8)$$

In section 2.3 we will introduce the definition of an r -dimensional Lie group \mathbb{G} and its corresponding Lie algebra \mathcal{G} . This facilitates the formally correct definition of the local symmetry in eq. (1.2) in the form of the level-set equation

$$X\phi = 0 \quad \text{if} \quad \phi(g \circ \mathbf{x}) = \phi(\mathbf{x}), \quad g = \exp(tX) \in \mathbb{G}, \quad X \in \mathcal{G} \quad (1.9)$$

Sections 2.1 and 2.3 prepare the stage for the introduction of Emmy Noethers celebrated first theorem in section 2.4. In a nutshell this theorem states that if

an energy functional $E_Y(\phi)$ is invariant upon the action of an r -dimensional Lie group \mathbb{G} , then there exists r *divergence-free* vector fields \mathbf{W}_m

$$g \circ E_Y(\phi) = E_Y(\phi) \quad \forall g \in \mathbb{G} \quad \leftrightarrow \quad \exists \mathbf{W}_m, \quad \text{div}(\mathbf{W}_m) = 0 \quad \forall \quad 1 \leq m \leq r \quad (1.10)$$

Since its first publication in 1918, Noether's first theorem has had far reaching implications in our understanding of the fundamental laws of motion in physics as well as the deep connection between the symmetries of a physical system and its conservation laws. For instance the time invariance of the laws of motion in the universe reveals the conservation of energy. In layman words: It does not matter **we** carry out an experiment now or next week, the results will be the same since the energy of the universe does not vanish! Building on section 2.4 we demonstrate in **section 3** the construction of a prior energy functional $E^{prior}(\phi)$ which is invariant under the Lie group $\mathbb{T} \times SO(2)$ which is the group of local translations and rotations. In section 3.3 we will use the prior developed in section 3 in the context of optical flow [1]. In section 4 we will introduce a generalization of the Newton approach for solving the inference problem in eq. (1.8) which takes local transformations of the spatial coordinates \mathbf{x} in Ω (see eq. (1.2)) into account to facilitate the search for the shortest path in the inference problem in eq. (1.6).

2 Background

2.1 Gibbs Random Fields

A physical system C is a dynamical composite of elements which interact with each other as well as with the environment the system C is embedded in. The elements are described by a vector of parameters $\phi = (\phi_1, \dots, \phi_n)$. The physical system C relates a specific value ϕ^* of the vector ϕ to a set of observables $Y = \{Y_1, \dots, Y_k\}$

$$Y = C(\phi^*) \quad (2.1)$$

In the case that the elements of the system C are continuously distributed over a finite space Ω , the parameter vector ϕ is a function on Ω

$$\phi(\mathbf{x}) \in \mathbb{R}^n \quad \mathbf{x} \in \Omega \quad (2.2)$$

called a Gibbs-Random-Field (GRF) [2]. The interactions of the elements of the system C with the environment are characterized by an energy functional $E_Y^{data}(\phi)$ called the data term, which couples the GRF $\phi(\mathbf{x})$ to the observables Y . There is another energy form $E^{prior}(\phi, \partial_j \phi)$ within the system C called the prior. $E^{prior}(\phi, \partial_j \phi)$ describes how the elements of C interact with each other. Together both energy functionals form the total energy of the system C

$$E_Y(\phi) = E_Y^{data}(\phi) + E^{prior}(\phi, \partial_j \phi) \quad (2.3)$$

which is related to the probability distribution

$$p(\phi|Y) = p(Y|\phi) \cdot p(\phi) \sim \exp(-E_Y(\phi)) \quad (2.4)$$

$$p(Y|\phi) = \exp(-E_Y^{data}(\phi)) \quad (2.5)$$

$$p(\phi) = \exp(-E^{prior}(\phi)) \quad (2.6)$$

The value of the probability distribution $p(\phi|Y)$ evaluated at the values $\hat{\phi}(\mathbf{x})$ describes the probability that the GRF $\phi(\mathbf{x})$ assumes the values $\hat{\phi}(\mathbf{x})$ at each

point $\mathbf{x} \in \Omega$. The set of values $\hat{\phi}(\mathbf{x})$ is what is called a *configuration* of the GRF ϕ .

$E_Y(\phi)$ is designed such that it is minimal once the GRF $\phi(\mathbf{x})$ fulfills the forward problem in eq. (2.1)

$$\phi^* = \operatorname{argmin}_{\phi} (E_Y(\phi)) \quad (2.7)$$

The particular value $\phi^*(\mathbf{x})$ of the GRF ϕ is the most probable configuration of the distribution $p(\phi|Y)$ due to eq. (2.4) and the solution to the inverse problem

$$\phi^* = C^{-1}(Y) \quad (2.8)$$

An example of a physical system containing a GRF is a camera C recording an object O . The domain $\Omega \subset \mathbb{R}^2$ is the focal plane of the camera C and the object O is naturally projected onto the focal plane Ω producing the projection I_O . In theory the projection I_O is a continuous function in the coordinate frame of the plane O where the particular function value $I_O(\mathbf{x})$ is the light intensity the object O reflects to the point \mathbf{x} on the focal plane Ω . At the heart of the image acquisition process of basically all modern camera systems lies the concept of a CCD collecting the photons of the light at discrete positions $\mathbf{x}_{i,j}$ called pixels

$$I_{ij}^c \in \mathbb{R}, \quad \mathbf{x}_{i,j} \in \Omega \quad 1 < i < n, 1 < j < m \quad (2.9)$$

The observables Y are the recorded intensities I_{ij}^c at the pixels $\mathbf{x}_{i,j}$. In this sense the camera C is a function which maps the continuous projection $I_O(\mathbf{x})$ to the discretely sampled intensities I_{ij}^c

$$I_{ij}^c = C_{ij}(I_O) \quad (2.10)$$

The intensity I_{ij}^c is basically a function of the number of photons collected by the CCD at the pixel $\mathbf{x}_{i,j}$. This number cannot be acquired deterministically, it is rather the result of a stochastic process described as independently identically distributed (iid) noise

$$\hat{I}_{ij}^c = I_O(\mathbf{x}_{i,j}) + n \quad n \sim p(I_{ij}^c | I_O(\mathbf{x}_{i,j})) \quad (2.11)$$

$p(I_{ij}^c | I_O(\mathbf{x}_{i,j}))$ is the likelihood that I_{ij}^c assumes the value \hat{I}_{ij}^c given the incoming intensity $I_O(\mathbf{x}_{i,j})$ at the pixel $\mathbf{x}_{i,j}$. Like in eq. (2.5) it is mapped to the data term energy $E_{I^c}(I_O)$.

In order to infer the values of $I_O(\mathbf{x}_{i,j})$ at the pixels $\mathbf{x}_{i,j}$ from the noisy data I_{ij}^c we need to pose some form of regularity on the values $I_O(\mathbf{x})$ to counter the pixel-wise noise imposed by the CCD in eq. (2.11). Such regularity can be achieved by

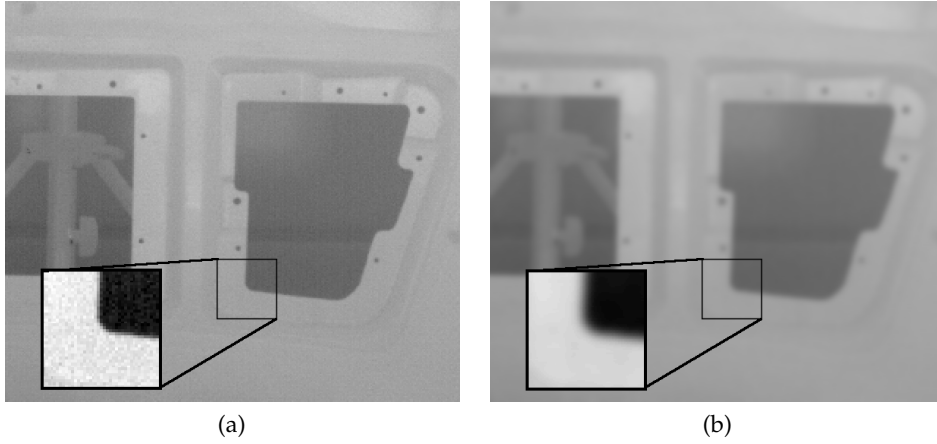


Figure 2.1: Figure 2.1a shows an image I^c taken of an object O with a thermographic camera. A region of interest is shown where the contrast was enhanced to visualize the noise corruption. Figure 2.1b shows the result I_O^* of the minimization problem eq. (2.14) with the prior in eq. (2.15). The noise is removed but the boundaries of O are over smoothed

correlating the intensities $I_O(\mathbf{x})$ at all pixels with each other in the prior

$$p(I_O) = \exp(-E^{prior}(I_O)) \quad (2.12)$$

$$E^{prior}(I_O) = \int_{\Omega} \mathcal{E}(I_O(\mathbf{x}), I_O(\Omega/\{\mathbf{x}\})) dx \quad (2.13)$$

where the integrand correlates the intensity $I_O(\mathbf{x})$ at the point $\mathbf{x} \in \Omega$ with the intensities at all other points $\Omega/\{\mathbf{x}\}$ so that the problem of inferring I_O from the data I^c becomes the minimization problem

$$I_O^* = \operatorname{argmin}_{I_O} (E_{I^c}(I_O)), \quad E_{I^c}(I_O) = E_{I^c}^{data}(I_O) + E^{prior}(\nabla I_O) \quad (2.14)$$

However in practice for a $n \times n$ dimensional image I^c the minimization in eq. (2.14) achieves a complexity of the order $\mathcal{O}(n^4)$ since every pixel is correlated to $n^2 - 1$ other pixels. Even for medium sized images with $n = 500$ the computations involved in eq. (2.14) are practically infeasible.

To reduce the complexity we want the integrand \mathcal{E} in eq. (2.13) only to correlate the values $I_O(\mathbf{x})$ within a neighborhood $U_{x_i,j} \subset \Omega$ with each other. One possible and very simple way to implement \mathcal{E} is to have it penalize the L_2 norm of the

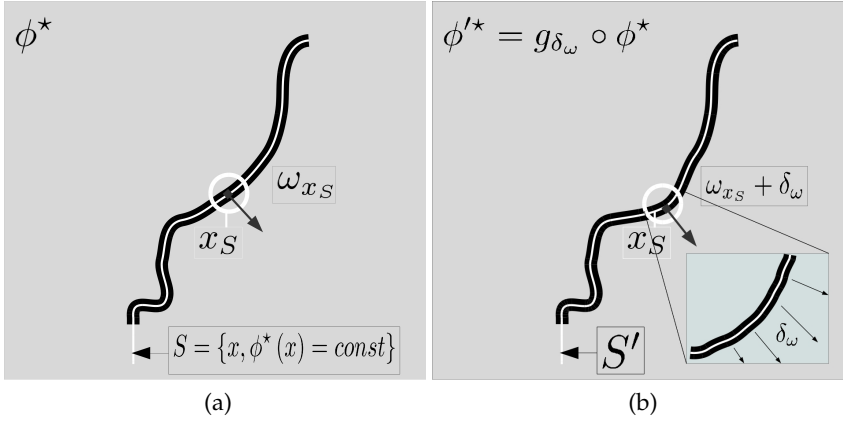


Figure 2.2: Local transformation of an image ϕ with a level-set S . Figure 2.2a shows an image $\phi(x)$ with a line S along which the intensity values are constant. At each point x_S the vector ω_S is the normal vector on S . Figure 2.2b shows the result of the local distortion of S under the action of the operator g_{δ_ω} . g_{δ_ω} acts on S by adding to ω_S a spacial dependent vector $\delta_\omega(x)$

gradient $\nabla I_O(x)$

$$E_{L_2}^{prior}(\nabla I_O) = \int_{\Omega} \|\nabla I_O(x)\|^2 dx \quad (2.15)$$

where the gradient operation ∇ can be realized by finite differences. While the prior in eq. (2.15) can be implemented in a very efficient manner, it has an important drawback. It isotropically smooths the GRF I_O regardless of the underlying geometry of the object O being recorded. In figure 2.1a the image I^c of an object O recorded by a thermographic camera is shown. A region of interest with enhanced contrast is shown to visualize the noise corruption due to the image measuring process in eq. (2.11). Figure 2.1b shows the result of the minimization in eq. (2.14) with the L_2 prior in eq. (2.13). $E_{L_2}^{prior}$ reduces the noise in I_O but due to its isotropic nature it over-smooths the boundaries of O . In section 2.2 and following we will introduce a methodology aimed at designing prior energies E^{prior} which incorporate information about the geometry of the objects recorded in order to avoid the over-smoothing across their boundaries.

2.2 Lie Groups and the Noether Theorem

2.2.1 Motivation 1, the problem

In section 2.1 we had claimed that the problem with the L_2 prior

$$E_{L_2}(\phi) = \int_{\Omega} \|\nabla\phi\|^2 \quad (2.16)$$

over-smooths the GRF ϕ over the boundaries of the object recorded by the camera C . In general the minimizers ϕ^* of the energy E_{L_2} are the constant functions $\phi = \text{const}$

$$A_c = \left\{ \phi_c^* \mid \phi_c^* = \operatorname{argmin}_{\phi} (E_{L_2}(\nabla\phi)) = c, \quad c \in \mathbb{R} \right\} \quad (2.17)$$

In the context of the minimization problem in eq. (2.14) the minimizer set A_c in eq. (2.17) emphasizes that the prior E_{L_2} does not allow for the solution I_O^* (eq. (2.14)) to have discontinuities. Thus E_{L_2} is completely unaware of the geometry in the data I^c (figure 2.1a). However E_{L_2} has a advantageous property. Consider the set of rotations $SO(2)$ of the coordinate frame Ω

$$x' = \mathbf{R}_{\theta} x, \quad \mathbf{R}_{\theta} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \in SO(2) \quad (2.18)$$

The gradient $\nabla\phi$ transforms under the rotation in eq. (2.18) like a vector, $\nabla'\phi = \mathbf{R}_{\theta}\nabla\phi$ and the matrix \mathbf{R}_{θ} satisfies $\mathbf{R}_{\theta}^T \mathbf{R}_{\theta} = \mathbf{1}$. Thus the L_2 energy is also invariant towards the rotations in eq. (2.18)

$$E'_{L_2} = \int_{\Omega} \nabla^T \phi \mathbf{R}_{\theta}^T \mathbf{R}_{\theta} \nabla \phi d^2x = \int_{\Omega} \|\nabla\phi\|^2 d^2x \quad (2.19)$$

In general the invariance of the prior energy $E^{prior}(\nabla\phi)$ of a GRF ϕ with respect to the rotations in eq. (2.18) is a favorable feature since the gradient $\nabla\phi$ should not be penalized to a specific orientation. In the context of the minimization problem in eq. (2.14) rotational invariance of the prior $E^{prior}(\nabla I_O)$ ensures the gradient ∇I_O^* is not affected by the orientation of the camera system C .

Several methods have been introduced which allow for the construction of anisotropic priors. These methods include TV-Regularization [3, 4] which will be introduced in section 2.5, anisotropic diffusion guided by directional operators like the structure tensor [5, 6, 7] and level set methods of the Mumford-Shah type [8, 9, 10]. Among the earliest attempts for anisotropic regularization was

the work of Nagel et. al. [11]. They introduced a quadratic prior

$$E_D^{prior}(\nabla\phi) = \int (\nabla\phi(\mathbf{x}))^T D(\mathbf{x}) (\nabla\phi(\mathbf{x})) d^2x \quad (2.20)$$

The operator $D(\mathbf{x})$ is a local 2×2 symmetric valued matrix estimated within a local window around each point \mathbf{x} . $D(\mathbf{x})$ is precomputed and assumed to be fixed under variation of ϕ . Thus it's eigenvectors function as a guide for the gradient ϕ . For instance in eq. (2.14) we can insert eq. (2.20) for E^{prior} . Computing D such that it has only one non-zero eigenvalue λ and an eigenvector \mathbf{b} oriented perpendicular to the weighted gradient of the data I^c

$$D(\mathbf{x}) = \lambda \mathbf{b}(\mathbf{x}) \mathbf{b}(\mathbf{x})^T, \quad \mathbf{b}(\mathbf{x}_0) \perp \langle \nabla I^c(\mathbf{x}) \rangle(\mathbf{x}_0) \quad (2.21)$$

$$\langle \nabla I^c(\mathbf{x}) \rangle(\mathbf{x}_0) = \int_A w(\|\mathbf{x} - \mathbf{x}_0\|) \nabla I^c(\mathbf{x}) d^2x \quad (2.22)$$

the prior E_D^{prior} penalizes the tangential component of ∇I_O along \mathbf{b} in the minimization in eq. (2.14). Thus the solution I_O^* can have discontinuities perpendicular to \mathbf{b} . The drawback of E_D^{prior} is that we do not know if \mathbf{b} is the true tangential vector in the unbiased projection of the object O . And since D is fixed E_D^{prior} can not be invariant under the rotations in eq. (2.18). Thus the minimization in eq. (2.14) can produce a solution I_O^* in eq. (2.14) that has discontinuities which do not reflect the true boundaries of the object O . We conclude that priors $p(\nabla)$ with energies E_D^{prior} which are *not* rotation invariant are a source of error for the orientation of ∇I_O^* in eq. (2.14). The other source of error of the orientation is the data term E^{data} in eq. (2.14).

On the other side a potential anisotropic prior E^{prior} which is rotation invariant would lead to a solution I_O^* in eq. (2.14) for which the orientation of its structures is only determined by the data term E^{data} .

In the following we will introduce a methodology which allows us to characterize prior distributions $p(\nabla\phi)$ and their energies $E^{prior}(\nabla\phi)$ which allow for discontinuities in their minimizers $\phi^* = \operatorname{argmin}_\phi E^{prior}(\nabla\phi)$ while remaining invariant to a specified but more general set of spacial transformations \mathbb{G}_Ω .

2.2.2 Motivation 2, the solution

Another way to state the problem that the prior energy E_{L_2} only allows for constant minimizers $\phi^* = \text{const}$ (eq. (2.17)) goes as follows. The energy $E_{L_2}(\nabla\phi)$ is invariant upon the transformation $\phi'(\mathbf{x}) = \phi(\mathbf{x}) + d$ where d is a constant over Ω . Thus if $\phi_0^* = c'$ is a minimizer of E_{L_2} , $c' \in A_c$ then so is $\phi'^* = c' + d$

since $c' + d \in \mathbb{R}$ and by the definition of A_c in eq. (2.17) we have $\phi'^* \in A_c$. We would like to think of the operation of addition with constants d as a set \mathbb{G}_{const} of operators g_d

$$g : \mathbb{R} \rightarrow \mathbb{G}_{const}, \quad g_d = \cdot + d, \quad g_d \in \mathbb{G}_{const} \quad (2.23)$$

With the help of the construction in eq. (2.23) we can restate the invariance of E_{L_2} in the following way

$$g_d \circ E_{L_2}(\nabla\phi) = E_{L_2}(\nabla(\phi + d)) = E_{L_2}(\nabla\phi) \quad (2.24)$$

and A_c in eq. (2.17) can be viewed as being spun by one constant function $\phi_0^\phi(\mathbf{x}) = c$ and the set \mathbb{G}_{const}

$$A_c = \{\phi^* \mid \phi^* = g_d \circ \phi_0^\phi, \quad g_d \in \mathbb{G}_{const}\} \quad (2.25)$$

With the constructions in eq. (2.23) and eq. (2.25) the problem statement that the prior E_{L_2} only allows for constant minimizers is transferred to the statement that the set \mathbb{G}_{const} under which E_{L_2} is invariant is *too small* in some sense.

A more flexible prior $p(\nabla\phi)$ with prior energy E^{prior} should be invariant to a more general set of transformations \mathbb{G}^ϕ . At the same time $p(\nabla\phi)$ and E^{prior} should also be invariant to a spacial set of transformations \mathbb{G}^Ω in order for it not to impede the orientation of the gradient $\nabla\phi$ as motivated in section 2.2.1. Hence $p(\nabla\phi)$ is assumed to be invariant to the set $\mathbb{G} = \mathbb{G}^\phi \times \mathbb{G}^\Omega$ with the actions

$$g_{\omega^\phi} \circ \phi(\mathbf{x}) = \phi(\mathbf{x}) + \omega^\phi(\mathbf{x}), \quad g_{\omega^\phi} \in \mathbb{G}^\phi \quad (2.26)$$

$$g_{\omega^\Omega} \circ \mathbf{x} = \mathbf{x} + \omega^\Omega(\mathbf{x}), \quad g_{\omega^\Omega} \in \mathbb{G}^\Omega \quad (2.27)$$

The transformations in eq. (2.26) and eq. (2.27) formally capture all the possible transformations the prior $p(\nabla\phi)$ with energy E^{prior} is invariant to. In this sense \mathbb{G} is maximal and $p(\nabla\phi)$ is *conditioned* on \mathbb{G} but *not* on any particular operator $g \in \mathbb{G}$

$$p(\nabla\phi) = p(\nabla\phi \mid \mathbb{G}) \quad (2.28)$$

We say that $p(\nabla\phi)$ is *conditionally invariant* to the set \mathbb{G} . For instance the prior p_{L_2} with its energy E_{L_2} is conditionally invariant to the set $\mathbb{G} = \mathbb{G} \times SO(2)$, the set of addition of the variable ϕ with constants and the set of rotations in Ω (see eq. (2.18)).

Similar to the definition of A_c in eq. (2.25) we can describe the maximizers of

$p(\nabla\phi)$ as being related to each other by the elements of \mathbb{G}

$$A = \{\phi^* | \phi^* = g \circ \phi_0^* \quad g \in \mathbb{G}\} \quad (2.29)$$

The set \mathbb{G}_Ω contains operators which are purely geometric. The idea is to show that A may be split into sub sets $A_\Omega(\phi_c^*)$ whose elements are related to each other by the elements $g_{\omega\Omega} \in \mathbb{G}_\Omega$

$$A_\Omega(\phi_c^*) = \{\phi^* | \phi^*(\mathbf{x}) = \phi_c^*(g_{\omega\Omega} \circ \mathbf{x}), \quad g_{\omega\Omega} \in \mathbb{G}_\Omega\} \quad (2.30)$$

$$A = \{A_\Omega(\phi_c^*) | \phi_c^* = g_{\omega\phi} \circ \phi_0^*, \quad g_{\omega\phi} \in \mathbb{G}_i\} \quad (2.31)$$

This is significant for the following reason: knowledge of the geometric set of transformations \mathbb{G}_Ω under which $p(\nabla\phi)$ is conditionally invariant allows for a reduction of the set of maximizers A to a set A_{red} such that the elements $\phi_c^* \in A_{red}$ are not related to each other by \mathbb{G}_Ω

$$A_{red} = \{\phi_c^* | \phi_c^* = g_{\omega\phi} \circ \phi_0^*, \quad g_{\omega\phi} \in \mathbb{G}_i\} \quad (2.32)$$

$$\phi_d^*(\mathbf{x}) \neq \phi_c^*(g_{\omega\Omega}\mathbf{x}) \quad \forall g_{\omega\Omega} \in \mathbb{G}_\Omega, \phi_{c,d}^* \in A_{red} \quad (2.33)$$

We may also turn the argument around: we could specify the geometric set of transformations \mathbb{G}_Ω and design a prior $p(\nabla\phi)$ which is conditionally invariant under \mathbb{G}_Ω , thus having a reduced maximizer set A_{red} . To give hint of how the prior $p(\nabla\phi)$ could be designed we need the definition of a level-set. A level-set of an image ϕ_0^* is a sub set $S_c \subset \Omega$ defined by

$$S_c = \{\mathbf{x} | \phi_0^*(\mathbf{x}) = c\} \quad (2.34)$$

The action of an element $g \in \mathbb{G}_\Omega \times \mathbb{G}_i$ on an image $\phi(\mathbf{x})$ may be written as

$$g \circ \phi(\mathbf{x}) = g_{\omega\phi} \phi(g_{\omega\Omega} \circ \mathbf{x}) \quad (2.35)$$

where we have split g into its components $g_{\omega\phi} \in \mathbb{G}_i$ and $g_{\omega\Omega} \in \mathbb{G}_\Omega$. By the definition of the action of $g_{\omega\Omega}$ in eq. (2.27) we see that $g_{\omega\Omega}$ is a geometrical transformation that deforms the level-sets S_c (see figure 2.2). We are free to define $g_{\omega\phi}$ so that it is orthogonal to $g_{\omega\Omega}$ in the sense that the level-sets S_c are invariant under $g_{\omega\phi}$

$$S_{c'} = g_{\omega\phi} \circ S_c = S_c \quad (2.36)$$

since a transformation of S_c is purely geometrical. Now the level-set S_c may alternatively be defined with the help of the vector-field $\omega_\delta(\mathbf{x})$ which (see figure 2.2) is the set of vectors *tangent* to S_c

$$S_c = \{\mathbf{x} | \omega_\delta(\mathbf{x}) \cdot \nabla\phi_0^*(\mathbf{x}) = 0\} \quad (2.37)$$

In figure 2.2b we show an example of a level-set S which is distorted by the operator $g_{\omega_\delta} \in \mathbb{G}_\Omega$. The resulting level-set S' has the vector-field $\omega'_\delta(\mathbf{x}) = \omega_\delta(\mathbf{x}) + \delta(\mathbf{x})$ as tangent vectors.

$$S'_c = \{\mathbf{x} | (\omega_\delta(\mathbf{x}) + t\delta(\mathbf{x})) \cdot \nabla\phi_0^*(\mathbf{x}) = 0\} \quad (2.38)$$

However it also possible to represent S'_c with the help of a deformation of the gradient operator ∇ itself

$$S'_c = \{\mathbf{x}' | \omega_\delta(\mathbf{x}') \cdot \nabla_{t\delta}\phi_0^*(\mathbf{x}') = 0\} \quad (2.39)$$

The operator $\nabla_{t\delta}$ loosely speaking encodes a reversal of the action of g_{ω_δ} on \mathbf{x} so that S'_c can be represented with the same tangential vector-field as S_c but in the new frame $\mathbf{x}' = g_{\omega_\delta} \circ \mathbf{x}$. The operator $\nabla_{t\delta}$ is called a *pull-back* of the gradient ∇ . With the help of the pull-backs $\nabla_{t\delta}$ it is possible to translate the notion of conditional invariance with respect to \mathbb{G}_Ω to the requirement that $p(\nabla_{t\delta}\phi)$ must be constant with respect to variations of the vector-field $\delta(\mathbf{x})$

$$\left. \frac{d}{dt} p(\nabla_{t\delta}\phi) \right|_{t=0} = 0 \quad (2.40)$$

Given a specific form of the operators in \mathbb{G}_Ω , eq. (2.40) poses constraints on the form of the differential operators in the prior $p(\nabla_{t\delta}\phi)$. Eq. (2.40) also ensures that $p(\nabla_{t\delta}\phi)$ is indifferent to a large class of level-sets $\{S\}$, which are generated by \mathbb{G}_Ω acting on S (see eq. (2.39)).

2.3 Lie Groups

In this section the set of operators \mathbb{G} is taken to act on a vector space \mathcal{M} . The set \mathbb{G} is called a *group* if there exists an operation \cdot so that \mathbb{G} contains

- the neutral element $e \in \mathbb{G}$: $e \cdot g = g$ for all $g \in \mathbb{G}$
- the inverse $g^{-1} \in \mathbb{G}$ if $g \in \mathbb{G}$

The group \mathbb{G} is called a *Lie group* [12, 13, 14] if the group operation

$$\mathbb{G} \times \mathbb{G} \mapsto \mathbb{G} : (x, y) \rightarrow x \cdot y^{-1}$$

is smooth in both x and y . The group operation \cdot can also be used to define the *left action* l_g on \mathbb{G}

$$l_g : \mathbb{G} \rightarrow \mathbb{G} \quad l_g(h) = g \cdot h \quad g, h \in \mathbb{G} \quad (2.41)$$

l_g is a smooth isomorphism in \mathbb{G} . The elements of \mathbb{G} may themselves be smooth mappings defined on an r -dimensional space \mathcal{A}

$$g : \mathcal{A} \rightarrow \mathbb{G}, \quad (a_1, \dots, a_r) \rightarrow g_{a_1, \dots, a_r} \quad (2.42)$$

In this case we say \mathbb{G} is an r -dimensional Lie group. A classical example of a Lie group is the group of invertible n -dimensional Matrices $GL(\mathbb{R}, n)$ over the vector space $\mathcal{M} = \mathbb{R}^n$. The dimension of $GL(\mathbb{R}, n)$ is n^2 and the group operation '·' is the matrix multiplication. In section 2.2.1 we argue that the set \mathbb{G} acts in a two-fold manner on the functions $\phi(\mathbf{x}) \in \mathcal{C}^\infty(\Omega)$, namely by acting on the spacial coordinates $\mathbf{x} \in \Omega$ in eq. (??) and on the function values $\phi(\mathbf{x})$ themselves in eq. (??). The spaces Ω and $\mathcal{C}^\infty(\Omega)$ are both vector spaces, that is the addition operation '+' and multiplication with a factor $\lambda \in \mathbb{R}$ are defined in both spaces. It is thus natural to combine both Ω and $\mathcal{C}^\infty(\Omega)$ to one single vector space $\mathcal{M} = \Omega \times \mathcal{C}^\infty(\Omega)$. However since the functions $\phi(\mathbf{x})$ are unknown and we would also like to place constraints on their derivatives $\phi_{,K}$ (K is a multi-index), we combine Ω together with the Jet space $J^k(\mathcal{C}^\infty(\Omega))$ [13], $\mathcal{M} = \Omega \times J^k(\mathcal{C}^\infty(\Omega))$. $J^k(\mathcal{C}^\infty(\Omega))$ is the set of smooth differentiable functions with compact support in Ω and their derivatives up to order k . The points $z \in \mathcal{M}$ are vectors of the independent variables \mathbf{x} , the dependent variable $\phi(\mathbf{x})$ and its derivatives $\phi_{,K}$

$$z = (\mathbf{x}, \phi(\mathbf{x}), \phi_{,K}(\mathbf{x})) \quad (2.43)$$

For this work we will focus only on first order derivatives, $k = 1$ so that the vectors z have the form

$$z = (\mathbf{x}, \phi(\mathbf{x}), \nabla\phi(\mathbf{x})) \quad (2.44)$$

The action of \mathbb{G} on \mathcal{M} is straightforward

$$\tilde{z} = (\tilde{\mathbf{x}}, \tilde{\phi}(\tilde{\mathbf{x}}), \tilde{\nabla}\tilde{\phi}(\tilde{\mathbf{x}})) \quad (2.45)$$

$$\tilde{\mathbf{x}} = g_{a_1 \dots a_r} \circ \mathbf{x} \quad (2.46)$$

$$\tilde{\phi} = g_{a_1 \dots a_r} \circ \phi \quad (2.47)$$

$$\tilde{\nabla} = J^{-1}\nabla, \quad J_{\mu\nu} = \frac{d\tilde{x}_\mu}{dx_\nu} \quad (2.48)$$

Since the elements $g_{a_1 \dots a_r}$ are continuous in the parameters a_i we are free to define to a smooth path γ in the parameter space \mathcal{A}

$$\gamma : t \rightarrow (a_1(t) \dots a_r(t)) \quad (2.49)$$

$$g_{\gamma(0)} = e \quad (2.50)$$

The derivative of $g_{\gamma(t)}$ with respect to t at $t = 0$ is an element of the tangential space of \mathbb{G} at the neutral element $e \in \mathbb{G}$, $T_e\mathbb{G}$

$$\left. \frac{d}{dt} g_{\gamma(t)} \right|_{t=0} = X_e \in T_e\mathbb{G} \quad (2.51)$$

The subscript on the vector X_e denotes that it belongs to $T_e\mathbb{G}$. The coordinates of X_e relative to the space \mathcal{M} can be computed when we look at the derivative of the induced action of $g_{\gamma(t)}$ on the space of smooth functions with support on \mathcal{M} , $\mathcal{F}(\mathcal{M})$. The action of X on $\mathcal{F}(\mathcal{M})$ can be computed by evaluating $F \in \mathcal{F}(\mathcal{M})$ on the transformed vector $\tilde{z} = g_{\gamma(t)} \circ z$ and the taking the derivative with respect to t at the neutral element e

$$X_e F(z) = \left. \frac{d}{dt} F(\tilde{z}) \right|_{t=0} = \sum_{i=1}^r (\omega_{\mu}^i \partial_{\mu} F(z) + \omega_i^{\phi} \frac{d}{d\phi} F(z) + D\phi_i^{\nu} \frac{d}{d\partial_{\nu}\phi} F(z)) \alpha_i \quad (2.52)$$

where we have

$$\omega_{\mu}^i(\mathbf{x}) = \left. \frac{d\tilde{x}_{\mu}}{da_i} \right|_{t=0} \quad \omega_i^{\phi}(\mathbf{x}, \phi) = \left. \frac{d\tilde{\phi}}{da_i} \right|_{t=0} \quad \alpha_i = \left. \frac{da_i}{dt} \right|_{t=0} \quad (2.53)$$

$$D\phi_i^{\nu} = \frac{d\omega_i^{\phi}}{dx_{\nu}} - \sum_{\mu} \frac{d\omega_{\nu}^i}{dx_{\mu}} \partial_{\mu}\phi \quad (2.54)$$

The function $D\phi_i^{\nu}$ is called the prolonged action of $g_{\gamma(t)}$ on the gradient operator ∇ (refer to appendix for derivation). Notice that while ω_{μ}^i and ω_i^{ϕ} are functions defined on \mathcal{M} , the coefficients α_i are independent of \mathcal{M} . They are the components of the vector X_e with respect to the r basis operators

$$X_{e,i} = X_e^{\Omega,i} + \omega_i^{\phi} \frac{d}{d\phi} + D\phi_i^{\nu} \frac{d}{d\partial_{\nu}\phi}, \quad X_e^{\Omega,i} = \omega_{\mu}^i \partial_{\mu} \quad (2.55)$$

so that X_e has the operator form

$$X_e = \sum_i \alpha_i X_{e,i} \quad (2.56)$$

An important point about X_e is that it is an *operator valued* function over Ω since the coefficients ω^i and ω^{ϕ} in eq. (2.53) are functions over Ω . We will refer to X_e as a vector at the unit element e , to ω^i as a vector valued function (VVF) and to ω^{Ω} as a scalar valued function (SVF). The vector X_e only exists in the tangential space at $e \in \mathbb{G}$, $X_e \in T_e\mathbb{G}$. However it is possible to construct a vector Y_h at a location $h \in \mathbb{G}$ by relating it to X_e with a map l_{h*} called the *push-forward*

$$Y_h F(z) = (l_{h^*} X_e) F(z) = \left. \frac{d}{dt} F(l_h(g_{\gamma(t)}) \circ z) \right|_{t=0} \quad (2.57)$$

The vector X_e operates on the function F in eq. (2.52) as a differential operator at the point $e \circ z = z$. The effect of l_{h^*} is that it transports the vector X_e to the vector Y_h which operates on F at the point $l_h(e) \circ z = h \circ z$. As Y_h is a smooth function with respect to h which is defined everywhere in \mathbb{G} it is called a *vector field*. The set of vector fields is the union of all the tangential spaces over \mathbb{G}

$$T\mathbb{G} = \bigcup_{h \in \mathbb{G}} T_h \mathbb{G} \quad (2.58)$$

It is important to keep in mind that the coordinates of the vector field Y_h are the operators $h \in \mathbb{G}$ and *not* the points $z \in \mathcal{M}$. Similar to X_e in eq. (2.56) the vector Y_h has a coordinate representation with respect to the tangential space $T_h \mathbb{G}$

$$Y_h F(z) = \sum_i \alpha'_i Y_{h,i} \quad (2.59)$$

$$Y_{h,i} = \omega_\mu^i \partial_\mu + \omega_i^\phi \frac{d}{d\phi} + D' \phi_i^\nu \frac{d}{d\partial_\nu \phi} \quad (2.60)$$

There exists a unique sub set $\mathcal{G} \subset T\mathbb{G}$ called the *Lie algebra*. It defined as the set of all vector fields $X_h \in T\mathbb{G}$ which are invariant under the left action l_g for any $g \in \mathbb{G}$

$$l_{g^*} X_h = X_{g \cdot h} = \sum_i \alpha_i X_{g \cdot h}^i \quad \forall g \in \mathbb{G}, X_h \in \mathcal{G} \quad (2.61)$$

From eq. (2.61) we see that a consequence of left invariance is that the coordinate vector α is constant under the transformation l_g . This is what is referred to as the *parallel transport* of α along the transformation l_g . The Lie algebra \mathcal{G} has the property that it is closed under the antisymmetric commutator $[\cdot, \cdot]$

$$[Y_h, X_h] = Z_h \in \mathcal{G} \quad \forall X_h, Y_h \in \mathcal{G} \quad (2.62)$$

Eq. (2.62) also implies that the commutator $[Y_h, X_h]$ is also left invariant [12]. The commutator in eq. (2.62) has a geometric meaning. Suppose Y_e is the vector in the tangent space $T_e \mathcal{G}_t^Y$ of the one parameter group g_t^Y in the sense of eq. (2.51). It is easy to show that the rate of change of the vector field $X_{g_t^Y}$ at the unit element e is equal to the commutator between X_e and Y_e

$$\left. \frac{d}{dt} X_{g_t^Y} \right|_{t=0} = [Y_e, X_e] \quad (2.63)$$

Since $[Y_h, X_h]$ is left invariant, eq. (2.63) may be translated to any point $g \in \mathbb{G}$

$$h \circ \frac{d}{dt} X_{g_t^Y} \Big|_{t=0} = \frac{d}{dt} X_{h \cdot g_t^Y} \Big|_{t=0} = [Y_h, X_h] \quad (2.64)$$

Essentially eq. (2.64) tells us that once we can tell how the vector field X_g changes along any path g_t^Y which goes through unity, $g_t^Y|_{t=0} = e$, we can compute its rate of change along any other path in \mathbb{G} .

A note on the evolution of functions on one parameter subgroups of \mathbb{G}

We want to clarify the main problem with the actual computation of the one parameter sub groups $g_t^Y \in \mathbb{G}$ and the central meaning of the unit element $e \in \mathbb{G}$ for the solution of that problem. The one parameter sub groups $g_t^Y \in \mathbb{G}$ with $Y_h \in \mathcal{G}$ define for any function F on the jet space \mathcal{M} an evolution

$$F(z(t)), \quad z(t) = g_t^Y \circ z_0 \quad (2.65)$$

The vector $z_0 = (\mathbf{x}_0, \phi_0(\mathbf{x}_0), \nabla_0 \phi_0)$ is an arbitrary chosen point in \mathcal{M} . The eq. (2.65) maps the path g_t^Y which lives in \mathbb{G} to a path in the epigraph of F (the combination of \mathcal{M} and F to $(\mathcal{M}, F(\mathcal{M}))$). The problem with eq. (2.65) is that due to the interdependency of \mathbf{x} , $\phi(\mathbf{x})$ and $\nabla \phi(\mathbf{x})$ it is generally not possible to write the path $F(z(t))$ in an analytical closed form. It is possible however to calculate in closed form the derivative of $F(g_t^Y \circ z_n)$ at time $t = 0$ if we are given an estimate z_n

$$\frac{d}{dt} F(g_t^Y \circ z_n) \Big|_{t=0} = Y_e(F(z_n)) \quad (2.66)$$

We remember that the vector Y_e is an operator valued function on the current estimate z_n

$$Y_e = y^\mu(z_n) \partial_\mu + y^\phi(z_n) \frac{\delta}{\delta \phi} \quad (2.67)$$

and that by left invariance Y_h is computed by evaluating the coefficient functions of Y_e at the point $h \circ z_n$.

$$Y_h = y^\mu(h \circ z_n) \partial_\mu + y^\phi(h \circ z_n) \frac{\delta}{\delta \phi} \quad (2.68)$$

It is in this sense that the unit element e must always be understood as *equivalent* to the current estimate z_n , $e \circ z_n = z_n$. We can use eq. (2.66) to construct a

piecewise linear approximation to the path $z(t)$ and thus for the path $F(z(t))$

$$z_{n+1} = z_n + \tau Y_e(z_n) \quad (2.69)$$

The estimates z_n from eq. (2.69) can be viewed as discrete samples from the path $z(t)$ at the discrete time steps t_n

$$z_n = z(t_n) \quad (2.70)$$

From eq. (2.67) we can see that it is the functional form of the coefficients $y^\Omega(z)$ and $y^\phi(z)$ that determine the paths $z(t)$ and $F(z(t))$ entirely via the process in eq. (2.69).

2.3.1 The Group $\mathbb{G} = \mathbb{T} \times SO(2)$

The group $\mathbb{G} = \mathbb{T} \times SO(2)$ is the group of translations and rotations in the plane \mathbb{R}^2 . Its algebra is the algebra $\mathcal{G} = \mathfrak{t} \times \mathfrak{so}(2)$ which has the basis $\{X_e^{\Omega,x}, X_e^{\Omega,y}, X_e^{\Omega,\theta}\}$. The subset $\mathfrak{t} = \{X_e^{\Omega,x}, X_e^{\Omega,y}\}$ is the set generators of infinitesimal translations

$$X_e^{\Omega,x} = \partial_x, \quad X_e^{\Omega,y} = \partial_y \quad (2.71)$$

\mathfrak{t} is a commutative basis since $[\partial_x, \partial_y] = 0$. The basis for $\mathfrak{so}(2)$ is the single operator $X_e^{\Omega,\theta}$ which is the generator of infinitesimal rotations. With respect to the Cartesian coordinate frame ∂_θ it has the following representation

$$X_e^{\Omega,\theta} = -y\partial_x + x\partial_y \quad (2.72)$$

From eq. (2.72) we can see that ∂_θ does not commute with \mathfrak{t} and the commutators for the basis $\{X_e^{\Omega,x}, X_e^{\Omega,y}, X_e^{\Omega,\theta}\}$ are easily computed

$$[X_e^{\Omega,\theta}, X_e^{\Omega,x}] = -X_e^{\Omega,y} \quad [X_e^{\Omega,\theta}, X_e^{\Omega,y}] = X_e^{\Omega,x} \quad [X_e^{\Omega,x}, X_e^{\Omega,y}] = 0 \quad (2.73)$$

We note that the group $SO(2)$ generates the unit circle S^1 by rotating the point $\mathbf{x}_0 = (x, y)$

$$\mathbf{x}(\theta) = g_\theta \circ \mathbf{x}_0 = R_\theta \mathbf{x}_0, \quad R_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \quad (2.74)$$

The meaning of the first two commutators in eq. (2.73) is that the gradient operator ∇ is rotated by 90° counter clockwise under the action of $X_e^{\Omega, \theta}$

$$\frac{d}{d\theta} \nabla_{x(\theta)} \Big|_{\theta=0} = [X_e^{\Omega, \theta}, \nabla] = M_\theta \cdot \nabla, \quad M_\theta = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (2.75)$$

The matrix M_θ is one of the Pauli matrices [15]. The Pauli matrices are the basis for the Lorentz group of special relativity which is an important symmetry for many quantum field theories for instance quantum electrodynamics [15, 16, 17].

2.4 Noether's First Theorem

In section 2.2.1 argued that in order for a prior $p(\nabla\phi)$ needs to be conditionally invariant to a large group of transformations \mathbb{G} in order for it's minimizers

$$A = \left\{ \phi^* \mid \phi^* = \operatorname{argmin}_\phi \left(-E^{prior}(\nabla\phi) \right) \right\} \quad (2.76)$$

to be non trivial, that is $\phi^* \neq \text{const.}$ Conditional invariance was linked to the requirement that the minimizer set A in eq. (2.76) be generated by the group \mathbb{G}

$$A = \{ \phi^* \mid \phi^* = g_\omega \circ \phi_0^* \quad g_\omega \in \mathbb{G} \} \quad (2.77)$$

In eq. (2.31) we explained that a transformation $g_\omega \in \mathbb{G}$ may partition the set A into subsets A_Ω whose elements are related to each other through geometrical transformations $g_{\omega\Omega}$ on the coordinate frame Ω . We motivated the introduction of deformations to the gradient operator ∇ such that the level-sets S' in eq. (2.39) have the same functional form in the transformed coordinates $\mathbf{x}' = g_{\omega\Omega} \circ \mathbf{x}$ as in the original coordinates (see eq. (2.37)). With the help of the machinery introduced in section 2.3 we can express a level-set S_X of ϕ in terms of a left invariant vector field X_h operating on ϕ at the identity $e \in \mathbb{G}$

$$S_X = \left\{ \mathbf{x} \mid X_e^\Omega \phi(\mathbf{x}) = 0 \right\} \quad (2.78)$$

The operator X_e^Ω is the spacial components of the vector X_e (see eq. (2.55)). Under the action of $g_{\omega\delta} \in \mathbb{G}$ the level-set S_X transforms the following way

$$S'_X = g_{\omega\delta} \circ S_X = \left\{ \mathbf{x} \mid X_{g_{\omega\delta}}^\Omega \phi(\mathbf{x}) = 0 \right\} \quad (2.79)$$

The requirement of conditional invariance of the prior $p(\nabla\phi)$ can be implemented by requiring that $p(\nabla\phi)$ be invariant with respect to transformations of the level-sets S_X like in eq. (2.79). Such a requirement effectively imposes constraint on

the form of the differential operators in $p(\nabla\phi)$ namely that they be expressed in terms of a basis of the Lie algebra \mathcal{G} , the left-invariant vector fields X_h^i . It is these elements of the Lie algebra \mathcal{G} which we will use as differential operators in the prior $p(\nabla\phi)$. Conditional invariance is then expressed by the equation

$$p\left(X_{g_{\gamma(t)}}^{\Omega,1}\phi^*, \dots, X_{g_{\gamma(t)}}^{\Omega,r}\phi^*\right) = \text{const}, \quad \text{with respect to } t \quad (2.80)$$

Since the basis X_e^i are the left invariant vector fields X_h^i evaluated at the identity e the eq. (2.80) would hold for a push-forwarded basis $X_{hg_{\gamma(t)}}^i$ for any $h \in \mathbb{G}$ since $h \circ \phi^* \in A$ if $\phi^* \in A$. This is why eq. (2.80) only needs to hold for small regions $U_e \subset \mathbb{G}$ around the identity

$$\left. \frac{d}{dt} p\left(X_{g_{\gamma(t)}}^{\Omega,1}\phi^*, \dots, X_{g_{\gamma(t)}}^{\Omega,r}\phi^*\right) \right|_{t=0} = V_e p\left(X_e^{\Omega,1}\phi^*, \dots, X_e^{\Omega,r}\phi^*\right) = 0 \quad (2.81)$$

where the vector V_e has a representation with respect to the basis X_e^i

$$V_e = \sum_{i=1}^r \alpha_i \cdot X_e^i \quad (2.82)$$

The eq. (2.81) would guaranty the independence of the solution space A with respect to the one parameter sub group $g_{\gamma(t)}$ where $g_{\gamma(t)}$ is related to a vector field $V = \sum_i \alpha_i X_e^i$ in the sense that V is the tangential vector of $g_{\gamma(t)}$ at the identity (see eq. (2.51)). However for eq. (2.81) to hold for all one parameter sub groups and thus for all $g \in \mathbb{G}$, it has to hold for all coefficient vectors $\alpha = (\alpha_1, \dots, \alpha_r)^T$.

2.4.1 Noethers Theorems

In her original paper [18, 19] Emmy Noether handles the question: Given a model of a physical system, encoded in an action

$$E = \int_{\Omega} (\mathcal{E}(\mathbf{x}, \{\phi_{\rho}\}, \{\nabla_K \phi_{\rho}\})) d^n x \quad (2.83)$$

which depends on ρ fields $\phi_1 \dots \phi_{\rho}$ and their derivatives to order K , and knowledge of a set of smooth transformations \mathbb{G} under which the action E is invariant

$$E' = g_{\gamma} \circ E = E \quad \forall g_{\gamma} \in \mathbb{G} \quad (2.84)$$

what are the special properties hidden in the model that invoke the symmetry?

To answer this question she deals with two cases:

- Finite dimensional Lie groups \mathbb{G} , which we introduced in section 2.3. For now it is sufficient to think of \mathbb{G} as the set of smooth functions g_γ defined on an r dimensional space, $\gamma = (\alpha_1, \dots, \alpha_r)$.
- Infinite dimensional Lie groups \mathbb{G}_∞ , which are generalizations of the finite dimensional groups in the sense that the r parameters $\alpha_1, \dots, \alpha_r$ are functions over the Cartesian coordinate frame Ω . We will not handle this case.

In the case of the finite dimensional group Emmy Noether took g_ω to be the smooth *infinitesimal* transformation, encoding both variations of the fields and of the coordinates

$$\phi'_\rho(\mathbf{x}) = \phi_\rho(\mathbf{x}) + \sum_{m=1}^r \alpha_m \omega_m^{\phi_\rho}(\mathbf{x}) \quad \mathbf{x}' = \mathbf{x} + \sum_{m=1}^r \alpha_m \omega_m^\Omega(\mathbf{x}) \quad (2.85)$$

The functions $\omega_m^{\phi_\rho}$ and ω_m^Ω can be seen as a basis for ω^ϕ and ω^Ω in eqs. ?? and ?. She proved that if the action E is invariant under g_ω eq. (2.84), then there exists r vectors \mathbf{W}_m such the integral relationship

$$E - E' = \int_\Omega \sum_{m=1}^r \alpha_m \left[\sum_\rho \bar{\omega}_m^{\phi_\rho} [\mathcal{E}]_\rho + \text{div}(\mathbf{W}_m) \right] = 0 \quad (2.86)$$

$$\bar{\omega}_m^{\phi_\rho} = \left(\omega_m^{\phi_\rho} - \omega_m^{\mu\Omega} \partial_\mu \phi_\rho \right), \quad [\mathcal{E}]_\rho = \frac{\delta \mathcal{E}}{\delta \phi_\rho} - \frac{d}{dx^\mu} \left(\frac{\delta \mathcal{E}}{\delta \phi_{\rho,\mu}} \right) \quad (2.87)$$

where $[\mathcal{E}]_\rho$ are the Euler-Lagrange differentials of the fields ϕ_ρ and the divergences $\text{div}(\mathbf{W}_m)$ appear by carefully collecting all terms which occur as a result of the integral product rule

$$\int f \cdot \partial_\mu g d^n x = \int \partial_\mu (f \cdot g) d^n x - \int \partial_\mu f \cdot g d^n x \quad (2.88)$$

when computing the symbolic form of $[\mathcal{E}]_\rho$. The main result is the argument that since the α_m , $\omega_m^{\phi_\rho}$ and the ω_m^μ are assumed to linearly independent, the r equations

$$\sum_\rho \bar{\omega}_m^{\phi_\rho} [\mathcal{E}]_\rho + \text{div}(\mathbf{W}_m) = 0 \quad m = 1, \dots, r \quad (2.89)$$

relate r of the ρ Euler-Lagrange equations $[\mathcal{E}]_\rho$ so that the physical system only has $\rho - r$ independent Euler-Lagrange equations $[\mathcal{E}]_\rho$ and thus only $\rho - r$ independent fields ϕ_ρ . In the case $\rho \leq r$ the system of equations in eq. (2.89) is overdetermined, eq. (2.86) can only hold if *all* the divergences and *all* the

Euler-Lagrange equations vanish

$$[\mathcal{E}]_\rho (\phi_1^*, \dots, \phi_\rho^*) = 0, \quad \text{div}(\mathbf{W}_m) (\phi_1^*, \dots, \phi_\rho^*) = 0, \quad \rho \leq r \quad (2.90)$$

Eq. (2.90) implies that only at the minima of the fields, ϕ_ρ^* the r vectors \mathbf{W}_m are conserved.

Kepler's Two Body Problem

Kepler's two body problem is the problem of calculating the problem of estimating the trajectory of a body of mass m_e (the earth) which is moving within the vicinity of another body with mass m_s (the sun). According to Newton there exists a gravitational force between the masses coming from the energy $V(r)$ of the gravitational field surrounding the mass m_s at the origin in \mathbb{R}^3

$$V(\mathbf{r}_e(t)) = -\frac{m_e \cdot m_s}{r} \quad r = \|\mathbf{r}_e - \mathbf{r}_s\| \quad (2.91)$$

The kinetic energy of the mass m_e is $\frac{1}{2}m_e \dot{r}^2$ so that the Lagrangian of the path $\mathbf{r}_e(t)$ is

$$L(\mathbf{r}_e(t)) = \frac{1}{2}m_e \dot{r}_e^2 + \frac{1}{2}m_e \dot{r}_s^2 - V(\mathbf{r}_e(t)) \quad (2.92)$$

The Euler-Lagrange equations are easily computed

$$\ddot{r}_e + \frac{m_s + m_e}{r^2} = 0 \quad (2.93)$$

The parameter t is the time parameter of the two body system. The Kepler Lagrangian in eq. (2.92) exhibits a symmetry under four different one parameter Lie group actions, namely the action of time shift and rotations around the three spacial axis (the group $SO(3) \times \mathbb{R}$)

$$t' = t + \delta t \quad (2.94)$$

$$\mathbf{r}' = \mathbf{r} + \partial_{\theta_i} \mathbf{r}' \delta \theta_i \quad i = x, y \text{ or } z \quad (2.95)$$

where θ_i are rotation around the x -, y - or z -axis. From Noether's theorem there

exist four corresponding conserved quantities:

$$W_t = \mathcal{H} = \frac{1}{2}m_e \dot{r}^2 + V(\mathbf{r}_e(t)) \quad \text{time shift} \quad (2.96)$$

$$W_x = z\dot{y} - y\dot{z} \quad \text{Rotation around } x\text{-axis} \quad (2.97)$$

$$W_y = z\dot{x} - x\dot{z} \quad \text{Rotation around } y\text{-axis} \quad (2.98)$$

$$W_z = x\dot{y} - y\dot{x} \quad \text{Rotation around } z\text{-axis} \quad (2.99)$$

The conserved quantity \mathcal{H} in eq. (2.96) is the *Hamiltonian Energy* of the two body system. It constant time and thus manifests that the total energy of the two body system does not dissipate away since there are no external forces interacting with the two masses m_e and m_s , that is the two body system is a *closed system*. The vector $\mathbf{W} = (W_x, W_y, W_z)$ (Eqs. eq. (2.97) to eq. (2.99)) is the total *angular momentum* the masses m_e and m_s have as they rotate around each other. The solutions to the Euler-Lagrange equations in eq. (2.93) are elliptic curves in the surface $S_{\mathbf{W}}$ orthogonal to \mathbf{W} . The constancy of \mathbf{W} with respect to the special orthogonal group $SO(3)$ comes the fact that $S_{\mathbf{W}}$ is actually a flat Euclidean plane embedded in a 3-dimensional Euclidean space.

2.4.2 Noether's First Theorem: A Modern Version

In this section we explicitly derive Noether's first theorem for models with one field ϕ and its first derivatives $X_e^{\Omega, i} \phi$ using the Lie algebra introduced in section 2.3. We consider the negative log energy

$$E = -\ln p(\phi, X_e^1 \phi, \dots, X_e^q \phi) = \int_{\Omega} \mathcal{E}(\phi, \{X_e^{\Omega, i} \phi\}) \mathcal{N} d^2 x \quad (2.100)$$

$$= \int_{\Omega} \mathcal{E}^{data}(\phi) \mathcal{N} d^2 x + \int_{\Omega} \mathcal{E}^{prior}(\{X_e^{\Omega, i} \phi\}) \mathcal{N} d^2 x \quad (2.101)$$

The explanation of the constant \mathcal{N} will shortly follow. We apply a one parameter group $g_{\gamma(t)}$ to E^{prior} and according to eq. (2.51) we can compute the vector V_e in the tangent space of $g_{\gamma(t)}$ at $t = 0$.

$$\left. \frac{d}{dt} g_{\gamma t} \circ E \right|_{t=0} = \int_{\Omega} \left(\sum_{i=1}^q P_i [V_e^{\Omega}, X_e^{i, \Omega}] \phi + v^{\phi} [\mathcal{E}] \right) \mathcal{N} d^2 x \quad (2.102)$$

$$P_i = \frac{\delta \mathcal{E}}{\delta X_e^{i, \Omega} \phi}, \quad [\mathcal{E}] = \frac{\delta \mathcal{E}}{\delta \phi} - \sum_i \frac{d}{dx^{\mu}} \left(w_i^{\mu} \frac{\delta \mathcal{E}}{\delta (X_e^{i, \Omega} \phi)} \right) \quad (2.103)$$

The differentials $[\mathcal{E}]$ are called the Euler-Lagrange differentials [19, 13] and the vector \mathbf{P} is called the canonical momentum. Eq. (2.102) is the most general form of variation. It contains two components, namely one component proportional to intensity variations of the field ϕ , v^ϕ and one component proportional to variations of the coordinate frame Ω encoded in V_e^Ω . The integral volume d^2x also transforms under the action of g_t^V

$$\frac{d}{dt} \left(g_t^V \circ d^2x \right) \Big|_{t=0} = \frac{dv^\mu}{dx^\mu} d^2x \quad (2.104)$$

However we are only interested in the Euler-Lagrange differentials $[\mathcal{E}]$ and the canonical momentum vector \mathbf{P} since only they depend on the particular Lagrangian \mathcal{E} . Thus the normalization factor \mathcal{N} in the integral in eq. (2.100) is chosen such that the volume element $\mathcal{N}d^2x$ is invariant under the transformation $g_{\gamma(t)}$

$$\frac{d}{dt} \left(g_t^V \circ \mathcal{N}d^2x \right) \Big|_{t=0} = \left(\mathcal{N} \frac{dv^\mu}{dx^\mu} + \frac{d\mathcal{N}}{dt} \Big|_{t=0} \right) d^2x = 0 \quad (2.105)$$

In the following we will drop the normalization \mathcal{N} and assume d^2x to invariant under any transformation g_t^V . Since V_e is an element of the Lie algebra \mathcal{G} we can expand it in terms of the r basis elements X_e^i

$$V_e = \sum_{i=1}^r \alpha_i X_e^i \quad (2.106)$$

Under the expansion in eq. (2.106) the eq. (2.102) becomes

$$\frac{d}{dt} g_{\gamma t} \circ E \Big|_{t=0} = \sum_{m=1}^r \alpha_m \int_{\Omega} \left(\sum_{i=1}^q P_i \left[X_e^{m,\Omega}, X_e^{i,\Omega} \right] \phi + \omega_m^\phi [\mathcal{E}] \right) d^2x \quad (2.107)$$

We can transform eq. (2.107) in to the original version in eq. (2.86) by introducing the vector valued functions (VVF) \mathbf{W}_m

$$\mathbf{W}_m^\mu = \omega_m^\mu \mathcal{E} + \sum_i \omega_i^\mu \left(\omega_m^\phi - X_e^m(\phi) \right) P_i \quad (2.108)$$

Computing the divergence of the VVFs \mathbf{W}_m we can prove the following relation

$$\text{div}(\mathbf{W}_m) - X_e^{\Omega,m}(\phi) [\mathcal{E}] = \sum_i \left[X_e^{\Omega,m}, X_e^{\Omega,i} \right] (\phi) \cdot P_i \quad (2.109)$$

and substitute the commutator in eq. (2.107) with the left hand side of eq. (2.109)

$$\left. \frac{d}{dt} g_{\gamma_t} \circ E^{prior} \right|_{t=0} = \int_{\Omega} \sum_m \alpha_m \left(\operatorname{div}(\mathbf{W}_m) + \tilde{\omega}_m^{\phi}[\mathcal{E}] \right) d^2x \quad (2.110)$$

$$\tilde{\omega}_m^{\phi} = \omega_m^{\phi} - X_e^{m,\Omega}(\phi) \quad (2.111)$$

If the energy E is assumed to be invariant with respect to any one parameter group $g_t^V \subset \mathbb{G}$

$$\left. \frac{d}{dt} g_{\gamma_t} \circ E \right|_{t=0} = 0 \quad (2.112)$$

then by the argumentation in section 2.4.1 the divergences of the vectors \mathbf{W}_m in eq. (2.108) and the Euler-Lagrange differentials must vanish

$$[\mathcal{E}](\phi^*) = 0, \quad \operatorname{div} \mathbf{W}_m = 0 \quad \forall 1 \leq m \leq r \quad (2.113)$$

and by eq. (2.109) the equations in eq. (2.113) imply

$$\sum_i \left[X_e^{\Omega,i}, X_e^{\Omega,m} \right] (\phi^*) \cdot P_i = 0 \quad (2.114)$$

There are three cases to consider such that eq. (2.114) can hold:

- Case a: The Lie algebra \mathcal{G} is commutative, $[X_e^{\Omega,i}, X_e^{\Omega,m}] = 0$ for all $1 \leq i, m \leq r$
- Case b: $P_i = 0$ for all $1 \leq i \leq r$
- Case c: If we have $[X_e^{\Omega,i}, X_e^{\Omega,m}] \neq 0$ for some i and m the functional derivative \mathbf{P} if non-vanishing must be orthogonal to the vector \mathbf{M}_m , which is a vector for fixed m defined as $(\mathbf{M}_m)_i = [X_e^{\Omega,i}, X_e^{\Omega,m}] (\phi)$ over Ω

We call cases a and b trivial symmetries and case c a non-trivial symmetry.

Pure Spatial Symmetries

A stronger constraint than the invariance of the energy E with respect to arbitrary one dimensional sub groups g_t^V as in eq. (2.112) is the case of invariance with respect to pure spatial one dimensional sub groups $g_t^{V\Omega}$

$$\left. \frac{d}{dt} \left(g_t^{V\Omega} \circ E \right) \right|_{t=0} = 0 \quad (2.115)$$

The Lie algebra element V_e^Ω corresponding to $g_t^{V^\Omega}$ does not contain any variations to the field ϕ thus we can obtain an expression for eq. (2.115) simply by setting $v^\phi = 0$ and $\omega_i^\phi = 0$ in eqs. (2.102) and (2.107)

$$\left. \frac{d}{dt} (g_t^{V^\Omega} \circ E) \right|_{t=0} = \int_\Omega \sum_{m=1}^r \alpha_m \left(\sum_i [X_e^{\Omega,m}, X_e^{\Omega,i}] (\phi) \cdot P_i \right) = 0 \quad (2.116)$$

It follows that if eq. (2.116) holds for any one parameter sub group $g_t^{V^\Omega} \subset \mathbb{G}$ (any coefficient vector α) then

$$\sum_i [X_e^{\Omega,i}, X_e^{\Omega,m}] (\phi) \cdot P_i = 0 \quad (2.117)$$

must hold for any field configuration ϕ . Eq. (2.117) is specifically a constraint on the prior energy E^{prior} since the data term E^{data} does not contain any derivatives $X_e^{\Omega,i} \phi$ and thus the canonical momentum \mathbf{P} only depends on the prior energy density \mathcal{E}^{prior} . In chapter 3 we will introduce a prior E^{prior} which is conditionally invariant to the group $\mathbb{G}_\Omega = \mathbb{T} \times SO(2)$ which is the group of local translations and rotations. Its algebra $\mathcal{G} = \mathfrak{t} \times \mathfrak{so}(2)$ is 3-dimensional and although it is not a commutative algebra we will show that eq. (2.117) still holds for any field ϕ .

2.5 Total Variation

In this section we will introduce a widely spread method for anisotropic regularization of the GRF ϕ called Total Variation (TV) [3, 4, 20, 21, 22]. In the context of shock-filtering ([3, 23, 24]) it was shown that the functional

$$E_{L_1}(\phi) = \int |\nabla \phi| dx \quad (2.118)$$

has the appealing property that it does not penalize large discontinuities However its functional derivative with respect to ϕ is ill conditioned in the case $\nabla \phi \approx 0$. To alleviate the case, [3] chose the approximative prior

$$E_{L_1,approx}(\phi) = \int \sqrt{|\nabla \phi|^2 + \epsilon} dx \quad (2.119)$$

which is well behaved for $\epsilon > 0$. They were able to achieve good results with relatively sharp preserved discontinuities with data ϕ^0 having low SNRs. Never the less in the limit $\epsilon \rightarrow 0$ the Euler-Lagrange equations become more and more computationally instable. A theoretically more well conditioned form of TV is needed which we will outline, following ([22, 25]). To do this we need to explore

the function-space the minimizers of Eq. eq. (2.118) might belong to. Smooth functions ϕ_{smooth} are functions for which $\nabla\phi$ exists everywhere, thus they may be minimizers of Eq. eq. (2.118). But functions $\phi_{discont}$ containing discontinuities do *not* have finite L_1 norm of their gradients, $E_{L_1}(\phi_{discont}) = \infty$ since the gradient $\nabla\phi_{discont}$ does not exist at the discontinuities. A generalization of Eq. eq. (2.118) is possible if one assumes $\nabla\phi$ to be a distribution, more precisely a radon measure in the space $\mathcal{M}(\Omega)$. If there exists a radon measure $\mu \in \mathcal{M}(\Omega)$, such that for every $\mathbf{p} \in \mathcal{C}_0(\Omega)$ with compact domain, the following equality holds

$$\int_{\Omega} \phi \cdot \text{Div} \mathbf{p} dx = - \int \mathbf{p}^T d\mu < \infty \quad (2.120)$$

then μ is called the weak derivative of ϕ and we can identify $\nabla\phi = \mu$. It is then possible to define the function-space of bounded variation

$$BV = \{\phi \in L_1(\Omega) \mid \nabla\phi \in \mathcal{M}(\Omega)\} \quad (2.121)$$

Now it is possible to define a norm on BV . By virtue of the Hölder relation there exists a scalar C for which we can determine the upper bound of Eq. eq. (2.120)

$$\int_{\Omega} \phi \cdot \text{Div} \mathbf{p} dx \leq C \|\phi\|_{\infty} \quad (2.122)$$

The scalar C is the norm of the radon measure $\nabla\phi$ and is called the total variation of ϕ

$$TV(\phi) = \sup \left\{ \int_{\Omega} \phi \cdot \text{Div} \mathbf{p} dx \mid \|\mathbf{p}\|_{\infty} \leq 1 \right\} \quad (2.123)$$

As was discussed in [22] the functions ϕ are geometrically piecewise smooth, meaning there exists a partitioning $\{\Omega_k\}$ of Ω such that $(\nabla\phi)_{\Omega_k}$ are L_1 integrable. If dl_{mk} is a line segment in the intersection $\Omega_m \cap \Omega_k$ then $TV(\phi)$ can be written in the form

$$TV(\phi) = \sum_k \|\nabla\phi_{\Omega_k}\|_{L_1} + \sum_{k < m} L_{lm} \quad (2.124)$$

$$L_{lm} = \int_{\Omega_l \cap \Omega_m} |\phi_l - \phi_m| dl_{lm} \quad (2.125)$$

where ϕ_l the value of ϕ on the portion of $\partial\Omega_k$ which is interfacing with Ω_m and vice versa for ϕ_m . The first term in eq. (2.124) penalizes the smooth parts of ϕ (the gradients $(\nabla\phi)_{\Omega_k}$). Similar to eq. (??) $\|\nabla\phi_{\Omega_k}\|_{L_1}$ is invariant to shifts of ϕ_{Ω_k} by constants c'_k

$$\|\nabla(\phi_{\Omega_k} + c'_k)\|_{L_1} = \|\nabla\phi_{\Omega_k}\|_{L_1} \quad (2.126)$$

Thus due to eq. (2.126) we can view the smooth functions ϕ_{Ω_k} as being centered around constants c_k

$$\phi_{\Omega_k}(\mathbf{x}) = c_k + \tilde{\phi}_{\Omega_k}(\mathbf{x}) \quad (2.127)$$

where the c_k are determined by the data ϕ_d . For instance if we combined the TV functional with a data term $E_{data} = \sum_k \int_{\Omega_k} (\phi_d - \phi_{\Omega_k})^2 dx$

$$E(\phi, \nabla\phi) = \sum_k \int_{\Omega_k} (\phi_d - \phi_{\Omega_k})^2 dx + \lambda TV(\phi) \quad (2.128)$$

then in [8, 10] it was shown that the c_k can be computed to be the *mean* of the data ϕ_d within the area Ω_k , $c_k = \int_{\Omega_k} \phi_d d^2x$ given that the deviations $\tilde{\phi}_{\Omega_k}$ are penalized by the first term in eq. (2.124).

The second term penalizes the length of the section $\Omega_m \cap \Omega_k$ while maintaining the values $\phi_{k,m}$ and thus the *jump* $|\phi_k - \phi_m|$. It essentially penalizes the curvature of the line interfacing with both Ω_k and Ω_m . We will make this point clear in the following section. For now we remark that if we set $\phi_{\Omega_k} = c_k$ in the data term in eq. (2.128) then we obtain

$$\tilde{E}(\phi, \nabla\phi) = \sum_k \int_{\Omega_k} (\phi_d - c_k)^2 dx + \lambda TV(\phi) \quad (2.129)$$

which is of course only an approximation to eq. (2.128). The data term in eq. (2.129) is a measure for the variance of ϕ_d in Ω_k . The two terms in eq. (2.124) together with the data term $\sum_k \int_{\Omega_k} (\phi_d - c_k)^2 dx$ in eq. (2.129) balance the size of the partitions Ω_k since the boundaries of *small* partitions Ω_k have *high* curvature and thus high TV values, but *low* variances. On the other side *large* partitions Ω_k have boundaries of *low* curvature and thus low TV values, but *high* variances. The parameter λ in eqs. 2.128 and 2.129 marks the trade-off between the TV term and the data term in eqs. 2.128 and 2.129 and thus it determines the size of the partitions Ω_k .

2.5.1 The Mean Curvature of Total Variation

In eq. (2.124) we had argued that the TV measure can be split into a smooth part $\|\nabla\phi_{\Omega_k}\|_{L^1}$ measuring the deviation of the smooth functions ϕ_{Ω_k} from the constants c_k . We had claimed that the second term in eq. (2.124), the boundary term L_{lm} measures the curvature of the boundary between Ω_l and Ω_l . The line integral in L_{lm} in eq. (2.125) can be rewritten essentially as a measure for the

length of the level-set S_{lm} interfacing Ω_l and Ω_m

$$L_{lm} = |\phi_k - \phi_m| \|S\|_{lm}, \quad \|S\|_{lm} = \int_0^T \left\| \frac{d}{dt} (\mathbf{x}(t)) \right\| dt \quad (2.130)$$

The path $\mathbf{x}(t)$ can be considered as being generated by a one parameter Lie group $g_t^{V^\Omega}$ acting on the point \mathbf{x}_0 which is on the interfacing boundary between Ω_l and Ω_m

$$\mathbf{x}(t) = g_t^{V^\Omega} \circ \mathbf{x}_0, \quad \mathbf{x}_0 \in \Omega_l \cap \Omega_m \quad (2.131)$$

so that the length $\|S\|_{lm}$ is controlled by the Lie algebra element $V_e^\Omega = v(\mathbf{x})^\mu \partial_\mu$

$$\|S\|_{lm} = s(T), \quad s(t) = \int_0^t \|v(\mathbf{x}(t'))\| dt' \quad (2.132)$$

The function $s(t)$ in eq. (2.132) is called the *arc length* of the curve $\mathbf{x}(t)$. By virtue of the definition of the arc length $s(t)$ in eq. (2.132) we can express derivatives with respect to s by

$$\frac{d}{ds} = \frac{1}{\|v\|} \frac{d}{dt} \quad (2.133)$$

The *curvature* of $\mathbf{x}_{lm}(t)$, $\kappa(\mathbf{x}_{lm}(t))$ is obtained by re-parameterizing $\mathbf{x}(t)$ in terms of its arc length s , $\mathbf{x}(t) \rightarrow \mathbf{x}(t(s))$ and taking the second derivative of $\mathbf{x}(s)$ using eq. (2.133)

$$\kappa(\mathbf{x}_0) = \left\| \frac{d^2}{ds^2} \mathbf{x}(s) \right\|_{s=0} \quad (2.134)$$

$$= \frac{1}{\|v\|^3} \left(v_x \cdot \frac{dv_y}{dt} - v_y \cdot \frac{dv_x}{dt} \right) \Big|_{t=0} \quad (2.135)$$

In [22] it is shown that the expression for the curvature $\kappa(\mathbf{x}_0)$ at the point \mathbf{x}_0 in eq. (2.135) is equivalent to the *mean curvature* [26, 27, 28]

$$\kappa = \text{Div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \quad (2.136)$$

which is the functional derivative of the TV norm in eq. (2.118) with respect to ϕ

$$\kappa = -\partial TV(\phi) \quad (2.137)$$

For a thorough derivation of the mean curvature κ in terms of weak derivatives in BV spaces see [25]. The Euler-Lagrange equations of any energy function

$E = \int \mathcal{E} d^2x$ including the TV functional in eq. (2.118) as a prior

$$[\mathcal{E}](\mathbf{x}) = \frac{\delta \mathcal{E}}{\delta \phi}(\mathbf{x}) + \lambda \kappa(\mathbf{x}) = 0 \quad (2.138)$$

pose a bound on the value of the curvature $\kappa(\mathbf{x})$. Thus the TV functional penalizes the curvature κ of the S_{lm} interfacing Ω_l and Ω_m . As κ is an invariant of the Lie group $SE(2)$, the group of rotations and translations, TV is also an invariant of that group.

2.6 Optical Flow

In section 2.1 we had introduced the notion of an inverse problem, namely that given some data Y and a model C we would like to find the GRF ϕ which is mapped to the data Y by the model C , see eq. (2.8). A prime example of an inverse problem in computer vision is optical flow [29, 30, 1, 31, 32, 33, 20]. Optical Flow labels the task of densely measuring the motion between two or more frames captured by a camera, or the dense registration of two or more cameras on a pixel-by-pixel basis. Optical flow is a crucial step in many areas of computer vision. For instance optical flow estimation is a part of video compression [34, 35] used to detect areas of the video in which the rate brightness change is small. For example during the recording of a rigid scene optical flow can be used to determine when the camera motion stalls. During such periods the frames of the video can be stored in a memory efficient manner. In recent years structure from stereography and structure from motion (video from a single camera) have gained popularity as a means to capture 3D models for film productions and also due to the availability of low cost 3D printing [30, 36, 37, 38, 39, 40, 41, 42]. In both the stereography and the structure from motion pipelines optical flow is used for the triangulation of the dense point cloud, prior to generation of the final 3D mesh. In the case of a dual-modal setup both cameras may be of different types. For instance in medical imaging multi-modal dense image registration is used to fuse image information from CT and MR modalities of the human brain [43] and of the human spine [44].

In optical flow modeling the task at hand is to estimate the disparity between two images y and I recorded by two cameras y_c and I_c (see figure 2.3). Each image is a map between the coordinate space $\Omega \subset \mathbb{R}^2$ and the real numbers \mathbb{R} . Thus $y(\mathbf{x})$ is the intensity recorded by the camera y_c at the pixel location $\mathbf{x} \in \Omega$ while $I(\mathbf{x}')$ is the intensity recorded by I_c at the location $\mathbf{x}' \in \Omega$. In figure 2.3a we have depicted a multi-modal setup in which the two cameras y_c and I_c are recording images (figures 2.3c and 2.3d) from different angles. In this context the

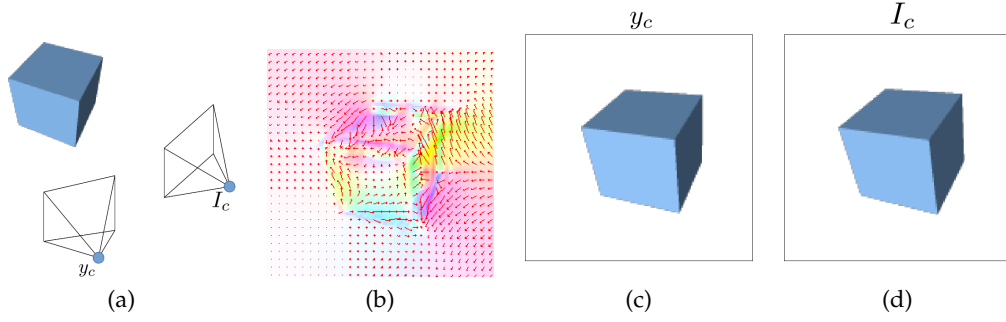


Figure 2.3: Figure 2.3a: Two cameras are shown recording a scene from different positions. The scene could be a rigid scene or a dynamic scene with moving objects. Figure 2.3c shows the image y captured from the camera y_c and figure 2.3d the image I from the camera I_c . Figure 2.3b shows the optical flow d . The vectors in figure 2.3b indicate which pixels x' in I and x in y are mapped to each other.

optical flow field is the unknown variable d which maps the location x' in the image I to the location x in the image y

$$x' = x + d(x) \quad (2.139)$$

The optical field d is shown in figure 2.3b as a set of vectors at every pixel $x' \in \Omega$, whose magnitude and orientation reflect the motion of the pixel x' .

The standard methodology [29, 30, 45, 46] for the estimation of the optical flow d is to model d as a GRF with a given data term $E_{y,I}^{data}(d)$. Without further information of the mapping between y and I from another source (e.g. sparse feature mapping), the data term $E_{y,I}^{data}(d)$ cannot depend directly on d but can only be defined as a *similarity measure* between the image $y(x)$ and the *warped* image $I_d(x) = I(x + d(x))$ [45, 46]

$$E_{y,I}^{data}(d) = F(y, I_d), \quad I_d(x) = I(x + d(x)) \quad (2.140)$$

In general the mapping of $y(x)$ and $I_d(x)$ via $d(x)$ is ill-determined: According to [33] the regions $\mathcal{A} \subset \Omega$ of an image ϕ can be given an *intrinsic dimension* [47, 48] iD which depends on their content

- $iD = 0$ if the image patch $\phi_{\mathcal{A}}$ is homogeneous
- $iD = 1$ if the image patch $\phi_{\mathcal{A}}$ contains an edge
- $iD = 2$ otherwise (e.g corners and/or textures)

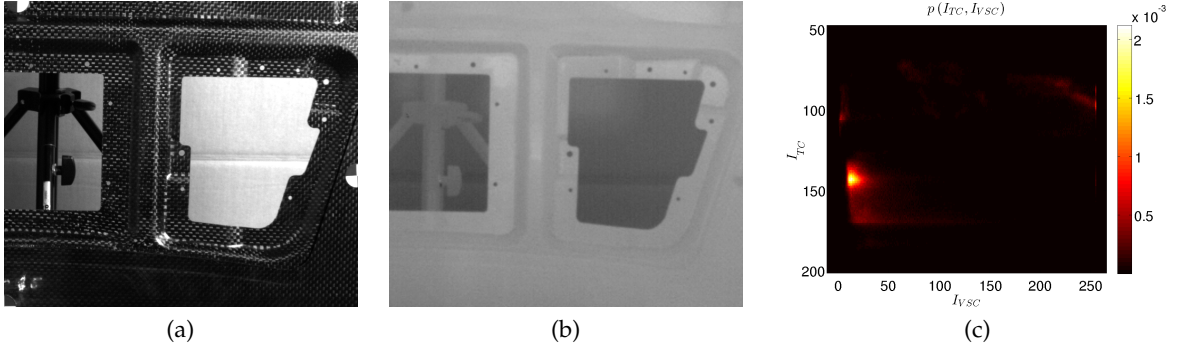


Figure 2.4: figure 2.4a shows an image from a visual spectrum camera (VSC). The object recorded is a carbon-fiber reinforced polymer (CFRP). Figure 2.4b shows an image of the same CFRP recorded with a thermographic camera (TC). The TC is sensitive in the infra-red domain, thus higher intensities in figure 2.4b correspond to warmer objects (the CFRP) and lower intensities to colder objects (the background). As in figure 2.3 the optical centers of the VSC and the TC are physically separated so the problem that is being addressed is that of finding the optical flow field $\mathbf{d}(\mathbf{x})$ (see eq. (2.139)) which maps every pixel in the TC image to the corresponding pixel in the VSC image. Figure 2.4c shows the joint probability distribution of the VSC and TC image. It shows a complex mapping of the intensities of figure 2.4a to those of figure 2.4b indicating that a linearity assumption between the TC and the VSC is not valid

If we consider two image patches y_A and I_A with equal intrinsic dimension iD then the number of components of the optical flow \mathbf{d} between y_A and I_A which can be *uniquely* determined by $E_{y,I}^{data}(\mathbf{d})$ is identical to iD . For instance if y_A and I_A both display a corner of an object ($iD = 2$), the optical flow in \mathcal{A} , \mathbf{d}_A can be uniquely determined by the similarity measure $E_{y,I}^{data}(\mathbf{d})$. However for edges ($iD = 1$) only one component of the optical flow \mathbf{d}_A can be determined and for homogeneous patches \mathbf{d}_A is completely undefined. Thus globally, that is over all $\mathbf{x} \in \Omega$ the similarity measure $E_{y,I}^{data}(\mathbf{d})$ cannot determine $\mathbf{d}(\mathbf{x})$ uniquely. For this reason optical flow models deploy a prior energy $E^{prior}(\nabla \mathbf{d})$ which smooths $\mathbf{d}(\mathbf{x})$ such that information of the components of \mathbf{d} in regions with $iD = 2$ is carried on to neighboring regions with $iD \neq 2$ such that $\mathbf{d}(\mathbf{x})$ is well defined over all Ω . The total energy

$$E_{y,I}(\mathbf{d}) = E_{y,I}^{data}(\mathbf{d}) + \lambda E^{prior}(\mathbf{d}) \quad (2.141)$$

is then a trade-off between the similarity of y and I_d and the smoothness of the optical flow $\mathbf{d}(\mathbf{x})$.

2.6.1 Uni-Modal Optical Flow

Among the earliest methods for optical flow estimation are the methods described in the seminal papers of Horn and Schunck [29] and Lukas and Kanade [30]. In [29] the following model for computing the flow between two frames of a video was proposed

$$E_{y,I}(\mathbf{d}) = E_{y,I}^{data}(\mathbf{d}) + \lambda E^{prior}(\mathbf{d}) \quad (2.142)$$

$$E_{y,I}^{data}(\mathbf{d}) = \int_{\Omega} (y(\mathbf{x}) - I_d(\mathbf{x}))^2 dx, \quad I_d(\mathbf{x}) = I(\mathbf{x} + \mathbf{d}(\mathbf{x})) \quad (2.143)$$

$$E^{prior}(\mathbf{d}) = \lambda \int_{\Omega} \sum_i \|\nabla d_i\|^2 dx \quad (2.144)$$

In eq. (2.143) the frame I is warped back to the frame y by the field $\mathbf{d}(\mathbf{x})$. The prior energy $E^{prior}(\mathbf{d})$ in eq. (2.144) imposes an isotropic smoothness constraint on the flow field \mathbf{d} . As we discussed in section 2.2.1 the main limitation of the L_2 prior in eq. (2.142) is that it does not preserve edges in the flow field $\mathbf{d}(\mathbf{x})$. To overcome this limitation [49] and [50] used the TV prior in eq. (2.118) as a smoothing term for each of the components of \mathbf{d}

$$E_{y,I}(\mathbf{d}) = E_{y,I}^{data}(\mathbf{d}) + \lambda \int_{\Omega} \sum_i \|\nabla d_i\| dx \quad (2.145)$$

According to section 2.5 the level-sets of each component $d_i(\mathbf{x})$ are smoothed while the discontinuities are preserved.

2.6.2 Multi-Modal Optical Flow

The next issue with the model in eq. (2.142) is that the likelihood $E_{y,I}^{data}(\mathbf{d})$ in eq. (2.143) makes the assumption that the cameras y_c and I_c are sensitive to the same physical light spectrum. For instance in figure 2.3 the image y recorded by the camera y_c in figure 2.3c has the same intensity spectrum as the image I recorded by the camera I_c (figure 2.3d) and we say that y and I are *equal by distribution*

$$y \stackrel{d}{\approx} I \quad (2.146)$$

Thus it is possible to find an optical flow field \mathbf{d}^* such that for each pixel $\mathbf{x} \in \Omega$ the warped image I_d approximates the image y , $y(\mathbf{x}) \approx I_d(\mathbf{x})$. However there

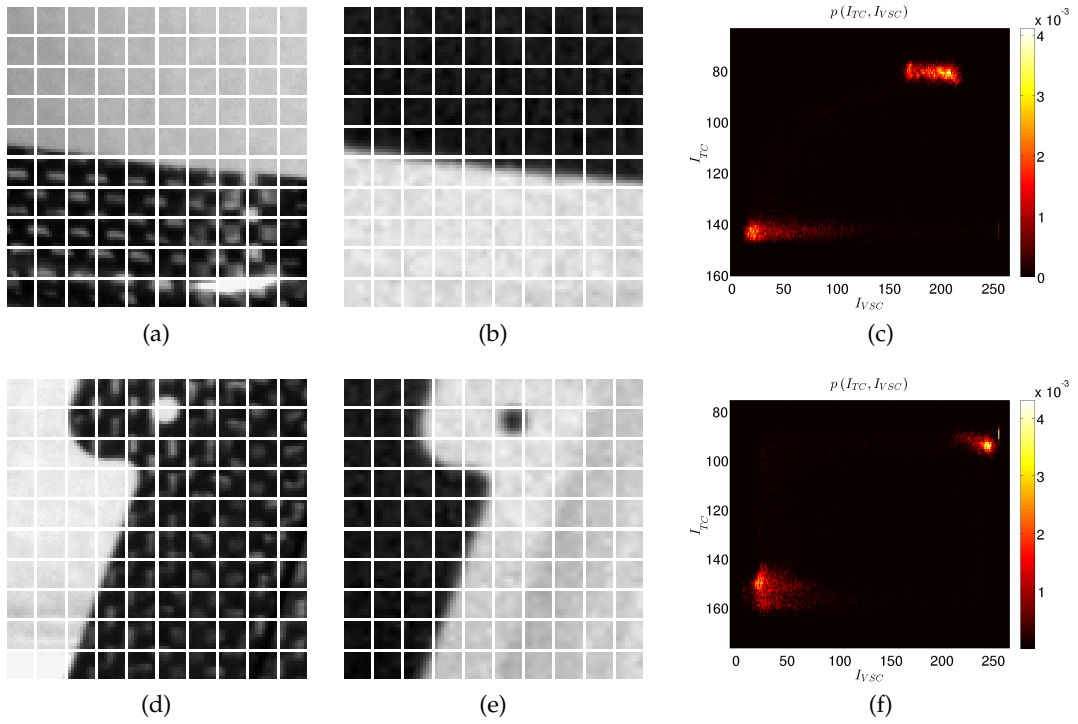


Figure 2.5: Figures 2.5a and 2.5d show 2 local regions of the VSC image and figures 2.5b and 2.5e their counter parts in the TC image. The histograms in figures 2.5c and 2.5f exhibit two maxima. The maxima can be thought to be the support points of a line in the histogram. Thus *locally* the TC and the VSC image have linear relationships

exists multi-modal setups where the cameras are *not* sensitive to the same spectra. Figure 2.4 shows two images recorded from a visual spectrum camera (VSC, figure 2.4a, I_{vsc}) and a thermographic camera (TC, figure 2.4b, y_{tc}). The recorded object, here a carbon-fiber reinforced polymer (CFRP) has physically different absorption and emission properties in the visual spectrum domain recorded by the VSC then in the infra-red domain recorded by the TC. Thus the intensities in I_{vsc} (figure 2.4a) follow a completely different distribution then those in y_{tc} (figure 2.4b). This is shown in the joint histogram between I_{vsc} and y_{tc} in figure 2.4c which fails to admit the correspondence $I_{vsc} \stackrel{d}{\approx} y_{tc}$ where the superscript d means *equally distributed*.

We will now discuss three statistical similarity measures for arbitrary images y and I which avoid the assumption of brightness constancy. For this we will take the two images y and I to be random variables with the marginal distributions $p(y)$ and $p(I)$. Then the mean and the variance are defined as

$$\mathbb{E}(X) = \int X \cdot p(X) \quad (2.147)$$

$$\text{Var}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) \quad (2.148)$$

The three similarity measures all avoid the brightness constancy assumption implied by the data term in eq. (2.143) by only taking into account the *statistical* features of the images y and I such as their joint entropy and joint covariance.

Mutual Information

Mutual Information (MI) [51, 52, 45, 53] is a popular similarity measure used mainly in medical imaging where images from different modalities including MR, CT and PET are registered against each other. For images y and I from two different modalities capturing the same scene, MI is defined with the joint distribution $p(y, I)$ by

$$MI(y, I) = \int p(\hat{y}, \hat{I}) \ln \frac{p(\hat{y}, \hat{I})}{p(\hat{y}) \cdot p(\hat{I})} d\hat{y}d\hat{I} \quad (2.149)$$

MI measures how strong the images y and I *statistically* depend on each other. In the case that y and I are statistically independent, $p(y, I) = p(y) \cdot p(I)$, then by eq. (2.149) MI is zero. On the other side, MI is maximal when y completely determined by I or vice versa. In the context of optical flow MI is used to measure

the similarity between y and I_d

$$E_{y,I}^{data}(\mathbf{d}) = -MI(y, I_d) \quad (2.150)$$

However, as [54] puts it, MI does not explain the kind of dependency between images y and I , its maxima are statistically but not visually meaningful, since it disregards any spacial information, which is essential for optical flow. Thus optical flow likelihoods based on MI usually tend to have many local minima rendering MI too unconstrained for optical flow.

Correlation Ratio

To alleviate the problems with MI, [54] argument that a better similarity measure would be one that measures the *functional* relation between the images y and I . The base key ingredient for their proposal is that the pixel values $I(\mathbf{x})$ and $y(\mathbf{x})$ are assumed to be the realizations of random variables, which by abuse of notation we denote by \hat{I} and \hat{y} . Then the normalized joint histogram of the images I and y can be interpreted as the joint probability distribution $p(\hat{y}, \hat{I})$, and the conditional distribution

$$p(\hat{y} | \hat{I} = I) = \frac{p(\hat{y}, \hat{I} = I)}{p(\hat{I} = I)} \quad (2.151)$$

encodes the spacial functional relationship between y and I . They introduced the Correlation Ratio (CR) [55, 52, 45, 46]

$$\eta(I|y) = \frac{\text{Var}(\phi^*(y))}{\text{Var}(I)} \quad E_{y,I}^{data}(\mathbf{d}) = -\eta(I_d|y) \quad (2.152)$$

The optimal function ϕ^* was shown to be the expectation value of \hat{I} , conditioned on a realization of \hat{y}

$$\phi^*(y) = \mathbb{E}(\hat{I} | \hat{y} = y) = \int I p(I|y) dI \quad (2.153)$$

The function $\phi(\hat{y})$ maps any realization of \hat{y} to an expectation value of \hat{I} . As \hat{y} is a random variable, $\phi(\hat{y})$ is also at random. Its variance measures how well I is *functionally explained* by a realization of \hat{y} . The measure in eq. (2.152) is bounded between 0 and 1, 0 indicating that y and I are independent, 1 indicating a functional relationship $I = \phi^*(y)$. The function ϕ^* although not necessarily continuous, is measurable in the L_2 -sense. Thus CR is a much stronger constraint than MI and has fewer, but more meaningful minima [54].

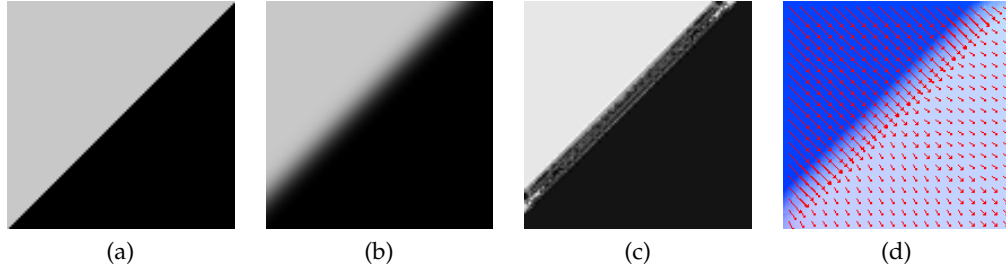


Figure 2.6: Figure 2.6a shows a synthetic high resolution image I^{syn} . In figure 2.6b we show a low resolution image H_y^{syn} . H_y^{syn} is computed by convolution of I^{syn} with Gaussian G_σ with standard deviation $\sigma = 5$ and translated by 10 pixels relative to I^{syn} . Figure 2.6d shows the flow d computed with the model in eq. (2.142), which does not incorporate knowledge of the scale difference between H_y^{syn} and I^{syn} and figure 2.6c show the warped image I_d^{syn}

Cross Correlation

Cross Correlation [56, 57, 58, 45] is the strongest constrained similarity measure. It is basically an additional constraint to CR, namely that the functional relationship in eq. 2.152 must be linear. Then η reduces to

$$\eta(I|y) = \frac{Cov(y, I)}{\text{Var}(I) \cdot \text{Var}(y)} \quad I = \lambda \cdot y \quad (2.154)$$

As we will see in section 3.4 a measure similar to eq. (2.154) will be computed based on the assumption that both y and I are Gaussian. The Gaussian assumption is valid when both cameras y and I produce Gaussian noise and the joint histogram is predominantly linear. Linearity in the joint histogram occurs when the recorded scene contains materials with uniform luminosity in the frequency bands of the cameras y and I .

All three similarity measures have in common that the images y and I must have the same spacial resolution in order to compute the measure. For instance in eq. (2.150) [51] computed the joint probability $p(y, I)$ as a normalized histogram $h(\hat{y}, \hat{I})$ created from the samples $\hat{y} = y(\mathbf{x})$ and $\hat{I} = I(\mathbf{x})$ drawn from all locations $\mathbf{x} \in \Omega$. However the image domain of I , Ω_I can have a larger size then the domain of y , Ω_y . In an ad-hoc fashion we could first filter and down-sample the image I with a convolution filter G to obtain an image h_I with the same spacial

resolution of Ω_y

$$h_I(\mathbf{x}) = (G \star I)(\mathbf{x}), \quad \mathbf{x} \in \Omega_y \quad (2.155)$$

and evaluate the similarity measures on the image pair y and h_I . The negative impact is that we could only estimate an optical flow d with the same low resolution as the image y . Conversely we could up-sample the image y with some interpolation scheme to produce a high resolution image H_y and evaluate the similarity measures on the pair H_y and I . This situation is shown in figure 2.6. I^{syn} in figure 2.6a shows a sharp linear boundary and H_y^{syn} (figure 2.6b) is a convolution of I^{syn} with a Gaussian G_σ of standard deviation $\sigma = 5$ which is translated by 10 pixels. The images H_y^{syn} and I^{syn} thus differ in optical scale, and the scale difference is the parameter σ . We used the model of Horn et. al

$$E(d^{syn}) = \frac{1}{2} \int_{\Omega} \left(H_y^{syn}(\mathbf{x}) - I_{d^{syn}}^{syn}(\mathbf{x}) \right)^2 dx + \frac{\lambda}{2} \sum_i \int_{\Omega} \|\nabla d_i^{syn}(\mathbf{x})\|^2 dx \quad (2.156)$$

(see eq. (2.142)) to compute the optical flow d^{syn} mapping I^{syn} to H_y^{syn} (see figure 2.6d). Figure 2.6c shows the image $I_{d^{syn}}^{syn}(\mathbf{x}) = I^{syn}(\mathbf{x} + d^{syn}(\mathbf{x}))$. We can see that the optical flow d corrupts the sharp boundary of I^{syn} in order to match it to the varying gray levels of the blurred boundary in H_y^{syn} (figure 2.6b). The problem that the model in eq. (2.156) can account for the *difference in size* of the images y^{syn} and I^{syn} but it does not take the *difference in optical scale* σ into account. Thus we need a model that can account for the optical scale σ .

2.7 Image Fusion

In this section we will introduce the image fusion method of Hardie et. al. [60]. In that paper the authors solved the problem of refining the low optical resolution of an image y_{tc} obtained by a thermographic camera (TC) using the image I_{vsc} obtained by a visual spectrum camera (VSC). The result of their method is a thermographic image Y_{tc} with improved optical resolution (see figure 2.7a). They used this method for the subject of remote sensing where the TC and the VSC are built in a co-aligned fashion (see figure 2.7a) within the body of a satellite. In section 3.3 the goal is to extend this method to the case where the TC and the VSC are not co-aligned. In this case low resolution image y_{tc} and thus the high resolution Y_{tc} have a natural separation from the VSC image I_{vsc} similar to figure 2.3. We will show that it is possible to jointly estimate the image Y_{tc} and the optical flow $d(\mathbf{x})$ between Y_{tc} and I_{vsc} .

The method of [60] goes as follows: In figure 2.8a a model of the CCD of the low

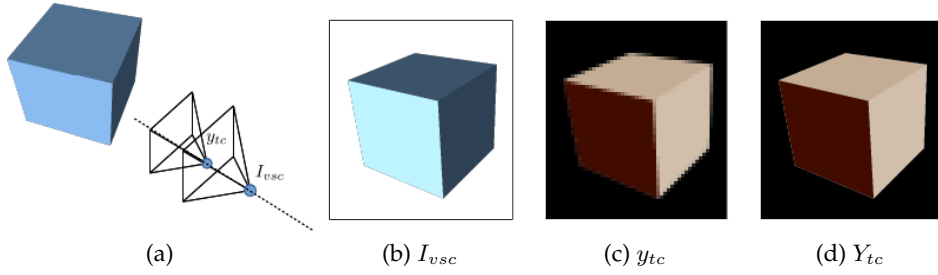


Figure 2.7: Figure 2.7a shows a schematic setup of the camera configuration considered by Hardie et. al. [60]. The dashed line indicates the orthogonal direction to both the TC and the VSC image planes. The camera centers (indicated by the blue circles) are aligned along the dashed line. Figure 2.7b shows the image I_{vsc} captured by the VSC and figure 2.7c the image y_{tc} captured by the TC. The image y_{tc} has a smaller optical resolution than the image I_{vsc} . The method in [60] takes the data y_{tc} and I_{vsc} to produce a higher resolution thermo-graphic image Y_{tc} , shown in figure 2.7d

resolution TC is shown overlaid with a higher resolution grid representing the VSC. The gray region in figure 2.8a symbolizes one pixel of the TC and it can be seen that each pixel of the TC covers a group of pixels of the VSC. Since the TC pixel has a finite surface, we need to specify how this pixel absorbs photons landing at different points in its area in order to relate the covered pixels of the VSC to it. The response of each individual pixel in the TC is called the point spread function (PSF), $W_\sigma(x, y)$, the vector (x, y) being the location on the surface of the TC pixel with respect to the VSC coordinate frame. Figure 2.8b is the result of a theoretical model of a FLIR TC [?]. The model, obtained by Hardie et al. [59], combines absorption properties of the CCD pixel with physical properties of the camera lens. We can see that each TC pixel has a non uniform response to incoming photons. Using this information we can model a super-resolved version Y_{tc} of the TC image y_{tc} with the help of the PSF W_σ , by stating that y_{tc} is the result of the convolution of Y_{tc} with W_σ

$$y_{tc} = W_\sigma Y_{tc} + n \quad n \sim \mathcal{N}(0|C_n) \quad (2.157)$$

The problem of estimating Y_{tc} is that there is an infinite amount of high resolution TC images Y_{tc}^* which relate to y_{tc} via eq. (2.157) since the high spacial frequency components of Y_{tc} are filtered out by W_σ . In [60] Hardie suggested use of a high resolution imager I_{vsc} whose camera center is co-aligned (see figure 2.7a) with the TC image y_{tc} and correlated with Y_{tc} . The rationale behind their approach is to combine the desired features such as sharp edges and corners of I_{vsc} with

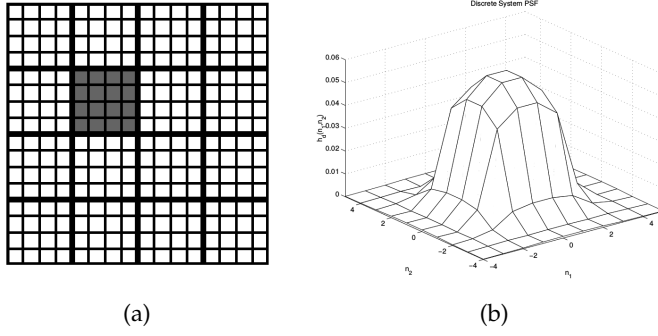


Figure 2.8: Figure 2.8a The thick grid depicts the CCD of the low resolution thermographic camera. The finer grid a virtual super-resolved version of the pixels in the TC. Figure 2.8b shows the point spread function $W_\sigma(x, y)$ of the gray pixel in figure 2.8a, taken from Hardie et al. [59]. It shows that each pixel in the TC image has a non uniform response over its surface to incoming photons.

the intensity spectrum of y_{tc} into the super-resolved image Y_{tc} , while avoiding limitations such as the noise model of y_{tc} .

The key ingredient in the model of [60] is that the intensities of Y_{tc} and I_{vsc} are assumed to be samples drawn from the joint Gaussian $p(Y_{tc}, I_{vsc})$. As I_{vsc} is already fixed as input data we can derive a conditional distribution for Y_{tc} via the Bayesian rule

$$p(Y_{tc}|I_{vsc}) = \frac{p(Y_{tc}, I_{vsc})}{p(I_{vsc})} \sim \mathcal{N}(\mu_{Y_{tc}|I_{vsc}} | C_{Y_{tc}|I_{vsc}}) \quad (2.158)$$

$$C_{Y_{tc}|I_{vsc}} = C_{Y_{tc}, Y_{tc}} - C_{Y_{tc}, I_{vsc}}^2 \cdot C_{I_{vsc}, I_{vsc}}^{-1} \quad (2.159)$$

$$\mu_{Y_{tc}|I_{vsc}}(\mathbf{x}) = \mu_{Y_{tc}} + C_{Y_{tc}, I_{vsc}} \cdot C_{I_{vsc}, I_{vsc}}^{-1} (I_{vsc}(\mathbf{x}) - \mu_{I_{vsc}}) \quad (2.160)$$

where the variances are computed globally

$$C_{u,v} = \int_{\Omega} (u(\mathbf{x}) - \mu_u) \cdot (v(\mathbf{x}) - \mu_v) dx \quad (2.161)$$

We see that the mean of Y_{tc} conditioned on I_{vsc} , $\mu_{Y_{tc}|I_{vsc}}$ is linear in the values of I_{vsc} , thus in this model the intensities of Y_{tc} are assumed to be *globally* linearly related to the intensities of I_{vsc} . We combine eq. (2.158) with the Gaussian likelihood in eq. (2.157) to the posterior

$$p(Y_{tc}|y_{tc}, I_{vsc}) \sim p(y_{tc}|Y_{tc}) \cdot p(Y_{tc}|I_{vsc}) = \exp(-E(Y_{tc})) \quad (2.162)$$

with the associated energy

$$\begin{aligned}
E(Y_{tc}) &= \frac{1}{2} \int_{\Omega} \left(y_{tc}(\mathbf{x}) - W_{\sigma} Y_{tc}(\mathbf{x}) \right)^2 \cdot C_n^{-1} d^x \\
&\quad + \frac{1}{2} \int_{\Omega} \left(Y_{tc}(\mathbf{x}) - \mu_{Y_{tc}|I_{vsc}}(\mathbf{x}) \right)^2 \cdot C_{Y_{tc}|I_{vsc}}^{-1} d^x
\end{aligned} \tag{2.163}$$

The minimization of eq. (2.163) and thus maximization of (2.162) with respect to Y_{tc} gives the analytical solution [60]

$$Y_{tc}^* = \mu_{Y_{tc}|I_{vsc}} + C_{Y_{tc}|I_{vsc}} \cdot W_{\sigma}^T \left(W_{\sigma} \cdot C_{Y_{tc}|I_{vsc}} \cdot W_{\sigma}^T + C_n \right)^{-1} \left(y_{tc} - W_{\sigma} \mu_{Y_{tc}|I_{vsc}} \right) \tag{2.164}$$

Eq. (2.164) is intractable to compute due to the dense operator W_{σ} and the matrix-inverse operation. In [61] a computationally tractable approximation was introduced

$$\hat{Y}_{tc} = \mu_{Y_{tc}|I_{vsc}} + C_{\tilde{Y}_{tc}|\tilde{I}_{vsc}} \cdot \left(C_{\tilde{Y}_{tc}|\tilde{I}_{vsc}} + C_n \right)^{-1} \left(y_{tc} - \tilde{\mu}_{Y_{tc}|I_{vsc}} \right) \tag{2.165}$$

$$\tilde{I}_{vsc} = W_{\sigma} I_{vsc} \quad \tilde{Y}_{tc} = W_{\sigma} Y_{tc} \approx y_{tc} \tag{2.166}$$

The key issue is that this method requires both modalities, I_{vsc} and y_{tc} , to be co-aligned. Since we are dealing with an optical flow problem y_{tc} and thus Y_{tc} is shifted by a disparity $\mathbf{d}(\mathbf{x})$ from I_{vsc} . This disparity has to be taken in to account by our model in chapter 3.3. The second issue is that the assumption that Y_{tc} and I_{vsc} are *globally* joint Gaussian is not supported by our data. However by computing $C_{Y_{tc}|I_{vsc}}$ in *local* sub-domains of the space Ω we can show that Y_{tc} and I_{vsc} are *locally* joint Gaussian. This will also be shown in chapter 3.3.

3 Linearized Priors

3.1 The Linear Structure Tensor

We shall now proceed to introduce a prior based on the considerations made in chapter 2.2. We will concentrate on the translation group \mathbb{T} for which the Lie algebra \mathfrak{t} is characterized by the set of vectors \mathbf{v} which are constant within a sub domain $A \subset \Omega$. The basis operators X_e^i are the Cartesian differential operators $\{\partial_x, \partial_y\}$, and the spacial component V_e^Ω of a vector $V_e \in T_e\mathbb{T}$ has the representation

$$V_e^\Omega = v_x(\mathbf{x}) \partial_x + v_y(\mathbf{x}) \partial_y \in \mathfrak{t} \quad \mathbf{v}(\mathbf{x})|_A = \text{const} \quad (3.1)$$

The translation group \mathfrak{t} is a commutative algebra so the basis $\{\partial_x, \partial_y\}$ is commutative and any vector V_e^Ω commutes with the $\{\partial_x, \partial_y\}$

$$\left[V_e^\Omega, \partial_{x,y} \right] = 0 \quad (3.2)$$

Thus V_e^Ω is translation invariant. Consider an image $\phi(\mathbf{x})$. The level-sets S_X corresponding to the vector V_e^Ω are defined by

$$S_X = \left\{ \mathbf{x} \mid \mathbf{v}^T \cdot \nabla \phi(\mathbf{x}) = 0 \right\} \quad (3.3)$$

We would like to characterize the dominant strength and the orientation of $\nabla \phi$ within the sub domain $A \subset \Omega$. In [62] it was suggested that the tangential vector \mathbf{v} of the level sets S_X can be computed by minimizing the energy

$$J(\mathbf{v}) = \frac{1}{2} \int_A w(\|\mathbf{x}\|) \mathbf{v}^T \cdot \left(\nabla \phi(\mathbf{x}) \nabla^T \phi(\mathbf{x}) \right) \mathbf{v} = \frac{1}{2} \mathbf{v}^T S \mathbf{v} \quad (3.4)$$

$$S = \int_A w(\|\mathbf{x}\|) \left(\nabla \phi(\mathbf{x}) \nabla^T \phi(\mathbf{x}) \right) d^2x \quad (3.5)$$

The weight function $w(\|\mathbf{x}\|)$ is normalized and weights the contributions of the gradient $\nabla \phi(\mathbf{x})$ at various points $\mathbf{x} \in A$. Typically a Gaussian is deployed for the weight function, $w(\|\mathbf{x}\|) = G_\kappa(\mathbf{x} - \mathbf{x}_0)$ where \mathbf{x}_0 is the center pixel of A . This way the gradient $\nabla \phi(\mathbf{x}_0)$ is the dominant contribution to the integral in eq. (3.5).

The matrix S is called the structure tensor. Since S is a symmetric matrix there exists an orthogonal decomposition

$$S = V^T D V \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad V = (\mathbf{V}_1, \mathbf{V}_2) \quad (3.6)$$

The eigenvalues give of the squared strength of the gradient in the basis defined by the columns of V . They characterize the structure in A in the following way

- $\lambda_1 > \lambda_2$: Strong linear level set with tangential vector $\mathbf{v} = \mathbf{V}_2$
- $\lambda_1 \approx \lambda_2 \approx 0$: No strong gradient, image is approximately constant
- $\lambda_1 \approx \lambda_2 \gg 0$: No *linear* level sets, level sets have strong curvature

We want to study the variation of the structure tensor S under the $SO(2)$ at the unit element e . Let S_θ be the structure tensor where the local coordinate frame A is rotated by the $SO(2)$ (see eq. (2.74))

$$S_\theta = \int_A w(\|\mathbf{x}(\theta)\|) \left(\nabla_{\mathbf{x}(\theta)} \phi(\mathbf{x}(\theta)) \nabla_{\mathbf{x}(\theta)}^T \phi(\mathbf{x}(\theta)) \right) d^2 x_\theta \quad (3.7)$$

The $SO(2)$ only rotates the domain A and does not deform it otherwise, thus the integral measure $d^2 x_\theta$ is independent of θ , $d^2 x_\theta = d^2 x$. Since the weighting function w only depends on the norm $\|\mathbf{x}(\theta)\|$ which is preserved by the $SO(2)$, it is also invariant. The only component which changes is the gradient $\nabla_{\mathbf{x}(\theta)}$. Using eq. (2.75) and the product rule we can compute the derivative of S_θ at $\theta = 0$

$$\left. \frac{d}{d\theta} S_\theta \right|_{\theta=0} = \int_A w(\|\mathbf{x}\|) \left(\mathbf{M}_\theta \nabla \phi \nabla^T \phi + \nabla \phi \nabla^T \phi \mathbf{M}_\theta^T \right) d^2 x \quad (3.8)$$

$$= \mathbf{M}_\theta \cdot S - S \cdot \mathbf{M}_\theta = [\mathbf{M}_\theta, S] \quad (3.9)$$

In eq. (3.9) we used $\mathbf{M}_\theta^T = -\mathbf{M}_\theta$. We can get some information on the magnitude of the rate of change $\left. \frac{d}{d\theta} S_\theta \right|_{\theta=0}$ by multiplying the commutator in eq. (3.9) with the eigenvectors $\mathbf{v}_{1,2}$. It is easy to show that both projection have the same norm

$$\|\tilde{\mathbf{v}}_{1,2}\| = |\lambda_1 - \lambda_2|, \quad \tilde{\mathbf{v}}_{1,2} = [\mathbf{M}_\theta, S] \mathbf{v}_{1,2} \quad (3.10)$$

With the help of eq. (3.10) we can reformulate our characterization of the eigenvalues $\lambda_{1,2}$ in the following way

- $\lambda_1 > \lambda_2$: Structure tensor S has strong change under $SO(2)$

- $\lambda_1 \approx \lambda_2$: Structure tensor S is largely invariant under the $SO(2)$ and approximately diagonal

$$S \approx \lambda \mathbb{1}, \quad \lambda \approx \lambda_1 \approx \lambda_2 \quad (3.11)$$

3.2 Structure Tensor Based Prior

Since the vector field V_e^Ω in eq. (3.1) is translation invariant the structure tensor S is also translation invariant. Under the rotation group $SO(2)$ the structure tensor is *not* invariant. Nonetheless it has an important transformation property: the transformed structure tensor S' may be written in terms of the old matrix S and the rotation matrix $R_\theta \in SO(2)$

$$S' = R_\theta^T S R_\theta \quad (3.12)$$

We would like to construct a prior $p(\nabla\phi)$ which is conditionally invariant conditionally invariant to the combined group $\mathbb{G} = \mathbb{T} \times SO(2)$. Since the eigenvalues λ_i of the structure tensor S are positive definite we propose as an energy prior for ϕ the integral over the determinant of S

$$E_{ST}^{prior} = \int_{\Omega} \mathcal{E}_{ST}(S) d^2x \quad (3.13)$$

$$\mathcal{E}_{ST}(S) = \frac{\lambda}{2} \det(S) \quad (3.14)$$

We want to show that E_{ST}^{prior} is invariant under the $SO(2)$. We insert S_θ from eq. (3.7) into the determinant in eq. (3.14) and evaluate the derivative of E_{ST}^{prior} with respect to θ

$$\left. \frac{d}{d\theta} E_{ST}^{prior}(S_\theta) \right|_{\theta=0} = \int_{\Omega} \text{Tr} \left(\mathbf{P}^{ST} \cdot [\mathbf{M}_\theta, S] \right) d^2x, \quad P_{ij}^{ST} = \frac{\delta \mathcal{E}_{ST}}{\delta S_{ij}} \quad (3.15)$$

The matrix \mathbf{P}^{ST} is the canonical momentum with respect to the structure tensor S , thus \mathbf{P}^{ST} it has the same transformation properties under the $SO(2)$ as S . The trace in eq. (3.15) can be further transformed

$$\text{Tr} \left(\mathbf{P}^{ST} \cdot [\mathbf{M}_\theta, S] \right) = \text{Tr} \left(\mathbf{P}^{ST} \cdot \mathbf{M}_\theta \cdot S \right) \quad (3.16)$$

The matrix under the trace on the right hand side of eq. (3.16) is a product of a symmetric and an anti-symmetric matrix, and thus itself anti-symmetric. Since traces over anti-symmetric matrices vanish, it follows that the prior E_{ST}^{prior} is

invariant under the $SO(2)$

$$\left. \frac{d}{d\theta} E_{ST}^{prior}(S_\theta) \right|_{\theta=0} = \int_{\Omega} \text{Tr} \left(\mathbf{P}^{ST} \cdot [\mathbf{M}_\theta, S] \right) d^2x = 0 \quad (3.17)$$

We note that the symmetry expressed by eq. (3.17) is a non-trivial symmetry, since only the trace as a whole vanishes.

3.3 Geometrical Optical Flow Model

describe opt flow, registration

In this chapter we will introduce our new model optical flow based on the image fusion algorithm from Hardie et al. [60]. We will address the two issues outlined in section 2.7, namely that the images y and I (Figures 2.4a and 2.4b) are *not* co-aligned and *not* joint Gaussian.

3.4 Disparity

The main objective of this chapter is to introduce a model which is capable of estimating the optical flow $\mathbf{d}(x)$ mapping the low resolution TC image y (figure 2.4b) to the high resolution VSC image I (figure 2.4a). There basically three problems with the data y and I :

- a: The images y and I have different intensity distributions, since the TC and the VSC are sensitive to different spectra.
- b: The images y and I have different resolutions.
- c: The image I contains textural information which is not contained in y

As is explained in the background (see section 2) the optical flow \mathbf{d} can only be estimated with a likelihood $p(y, I | \mathbf{d})$ which measures how similar the images y and I are given \mathbf{d} . However a likelihood that measures the similarity of the intensities of y and I like the one in eq. (2.142) would fail since the intensities cannot be compared due to problem a.

The difference in resolution in problem b causes an ambiguity of the optical field \mathbf{d} since the features in the lower resolved image y are blurred and it is not clear which pixel in I relates to which pixel in y . To demonstrate the issue we have created test data y^{syn} and I^{syn} in figure 2.6. I^{syn} in figure 2.6a shows a sharp linear boundary and y^{syn} (figure 2.6b) is a convolution of I^{syn} with a Gaussian

G_σ of standard deviation $\sigma = 5$ which is translated by 10 pixels. We used the model of Horn et. al

$$E(\mathbf{d}^{syn}) = \frac{1}{2} \int_{\Omega} \left(y^{syn}(\mathbf{x}) - I_{\mathbf{d}^{syn}}^{syn}(\mathbf{x}) \right)^2 dx + \frac{\lambda}{2} \sum_i \int_{\Omega} \|\nabla d_i^{syn}(\mathbf{x})\|^2 dx \quad (3.18)$$

(see eq. (2.142)) to compute the optical flow \mathbf{d}^{syn} mapping I^{syn} to y^{syn} (see figure 2.6d). Figure 2.6c shows the image $I_{\mathbf{d}^{syn}}^{syn}(\mathbf{x}) = I^{syn}(\mathbf{x} + \mathbf{d}^{syn}(\mathbf{x}))$. We can see that the optical flow \mathbf{d} corrupts the sharp boundary of I^{syn} in order to match it to the varying gray levels of the blurred boundary in y^{syn} (figure 2.6b).

In order to solve problem a and b we need a method to transform I to an image S which has the same intensity distribution as y but the same resolution as I . A putative likelihood $p(y, S|\mathbf{d})$ can measure how similar the images y and S are given \mathbf{d} .

If I contains a feature not existent in y or vice versa, the optical flow \mathbf{d} is ambiguous and the ambiguity may only be resolved upon removal of the contradicting feature.

In section 2.7 a method was introduced which produces a super-resolved image S given co-aligned data y and I_c .

The model is defined by the posterior distribution for S (see eq. (2.158))

$$p(S|y, I_c) = p(y|S) \cdot p(S|I_c) \quad (3.19)$$

$$-\ln(p(y|S)) = \frac{1}{2} \int_{\Omega} \left(y(\mathbf{x}) - W_\sigma S(\mathbf{x}) \right)^2 \cdot C_n^{-1} dx \quad (3.20)$$

$$-\ln(p(S|I_c)) = \frac{1}{2} \int_{\Omega} \left(S(\mathbf{x}) - \mu_{S|I_c}(\mathbf{x}) \right)^2 \cdot C_{S|I_c}^{-1} dx \quad (3.21)$$

with the conditional variance and mean

$$C_{S|I_c} = C_{S,S} - C_{S,I_c}^2 \cdot C_{I_c,I_c}^{-1} \quad (3.22)$$

$$\mu_{S|I_c} = \mu_S + C_{S,I_c} \cdot C_{I_c,I_c}^{-1} (I - \mu_I) \quad (3.23)$$

In the conditional prior $p(S|I_c)$ in eq. (3.21) pixels in S and in I_c have a one-on-one relationship, so that it is natural to map pixels in I to S rather than to y directly. We model the disparity between the images S and I by setting the co-aligned VSC image I_c to be the result of the original VSC I , warped by an unknown optical flow field $\mathbf{d}(\mathbf{x})$,

$$I_c(\mathbf{x}) = I(\mathbf{x} + \mathbf{d}(\mathbf{x})) = I_{\mathbf{d}}(\mathbf{x}) \quad (3.24)$$

Substituting eq. (3.24) into eq. (3.19) and following, we obtain the posterior

$$p(S|y, I, \mathbf{d}) = p(S|y, I_d) \quad (3.25)$$

with the energy

$$E_{post}(S, \mathbf{d}) = \frac{1}{2} \int_{\Omega} \left(y(\mathbf{x}) - W_{\sigma} S(\mathbf{x}) \right)^2 \cdot C_n^{-1} dx + \frac{1}{2} \int_{\Omega} \left(S(\mathbf{x}) - \mu_{s|I_d}(\mathbf{x}) \right)^2 \cdot C_{s|I_d}^{-1} dx \quad (3.26)$$

While keeping \mathbf{d} fixed we minimize $E_{post}(S, \mathbf{d})$ with respect to S and obtain similar to eq. (2.165) a closed form solution for S

$$\hat{S} = \mu_{s|I_d} + C_{s|\tilde{I}_d} \cdot \left(C_{s|\tilde{I}_d} + C_n \right)^{-1} \left(y - \tilde{\mu}_{s|I_d} \right) \quad (3.27)$$

We insert the simplified closed form expression for \hat{S} from eq. (3.27) into E_{post} and obtain an energy measuring the similarity between y and $\tilde{I}_d = W_{\sigma} I_d$

$$E_{data}(\mathbf{d}) = E_{post}(\hat{S}, \mathbf{d}) \quad (3.28)$$

$$= \frac{1}{2} \int_{\Omega} \left(y(\mathbf{x}) - f \cdot \tilde{I}_d(\mathbf{x}) \right)^2 \cdot C_{s|\tilde{I}_d} \left(C_{s|\tilde{I}_d} + \lambda C_n \right)^{-2} \quad (3.29)$$

$$f = C_{y, \tilde{I}_d} C_{\tilde{I}_d, \tilde{I}_d}^{-1} \quad (3.30)$$

The data term E_{data} defines a likelihood for \mathbf{d}

$$p(y, I|\mathbf{d}) = \exp(-E_{data}(\mathbf{d})) \quad (3.31)$$

We remember that the problems with the data y and I are that they (a) have different intensity distributions and (b) different resolutions. The likelihood in eq. (3.31) solves the problems a and b elegantly in one approach by introducing the latent variable S . The low resolution component of S , $W_{\sigma} S$ is coupled through the likelihood $p(y|S)$ in eq. (3.20) to the TC image y . The prior $p(S|I)$ in eq. (3.21) couples S to the high resolution image I . As a result I_d in E_{data} in eq. (3.35) is filtered by the PSF W_{σ} to match the scale of y . Furthermore the factor f transforms the intensity range of the filtered image \tilde{I}_d to a range similar to that of y so that E_{data} is a measure for the similarity between y and $f \cdot \tilde{I}_d$.

To demonstrate that our likelihood E_{data} in eq. (3.31) respects the difference in scale between y and I we have estimated the flow with E_{data} as the similarity measure for the data y^{syn} and I^{syn} in figure 2.6. The standard deviation σ in E_{data} was set to $\sigma = 5$ and the factor f is automatically computed as $f \approx 1$ since the intensity distributions of y^{syn} and I^{syn} are approximately the same. The

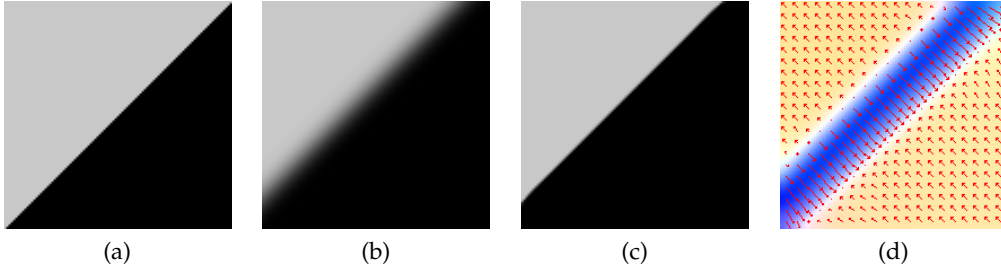


Figure 3.1: Figure 3.1a shows a synthetic high resolution image I^{syn} . In figure 3.1b we show a low resolution image y^{syn} . y^{syn} is computed by convolution of I^{syn} with Gaussian G_σ with standard deviation $\sigma = 5$ and translated by 10 pixels relative to I^{syn} . Figure 3.1d shows the flow \mathbf{d} computed with the model in eq. (3.31), which incorporates knowledge of the scale difference between y^{syn} and I^{syn} and figure 3.1c show the warped image I_d^{syn}

image I_d^{syn} is convolved with W_σ . The resulting image \tilde{I}^{syn} has the same scale as y^{syn} . The resulting optical flow \mathbf{d}^{syn} is shown in figure 3.1d. Notice the blurred boundary \mathbf{d}^{syn} around the linear feature in I^{syn} (figure 3.1a). This is the result of E_{data} in eq. (3.31) measuring the difference between y^{syn} and the blurred image $\tilde{I}_d^{syn} = W_\sigma I_d^{syn}$. In eq. (3.1c) we see I_d^{syn} . The linear boundary has been warped by \mathbf{d}^{syn} without being corrupted like in figure 2.6c.

3.5 Localization

The assumption that the intensities of the images y and I are *globally* linear related is a very strong constraint that can hold in most cases only unimodal data. In the case of the VSC and TC data in figure 2.4 the assumption of linearity fails. In figure 3.2 the global joint histogram of the VSC and the TC image is shown. We can see that the distribution in the joint histogram lacks a linear relationship between the TC and the VSC. However in figure 3.3 we have evaluated the histogram within local region of interests. The histograms in figures 3.3c and 3.3f show that within the ROI's the assumption of linearity between the intensities of the TC and the VSC is well supported. Therefore we propose a local version of the variance in eq. (2.161)

$$C_{u,v}(\mathbf{x}_0) = \int_{\Omega} \omega(\mathbf{x} - \mathbf{x}_0) (u(\mathbf{x}) - \mathbb{E}(u, \mathbf{x}_0)) \cdot (v(\mathbf{x}) - \mathbb{E}(v, \mathbf{x}_0)) \quad (3.32)$$

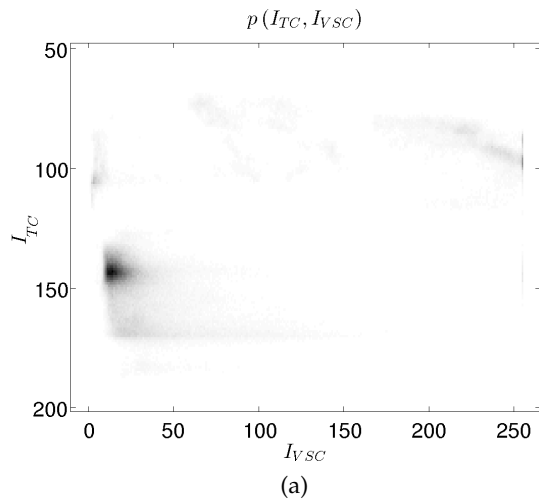


Figure 3.2: Joint Histogram of the TC image figure 2.4b and the VSC image figure 2.4a . We observe that there is no linear relationship between the TC and the VSC

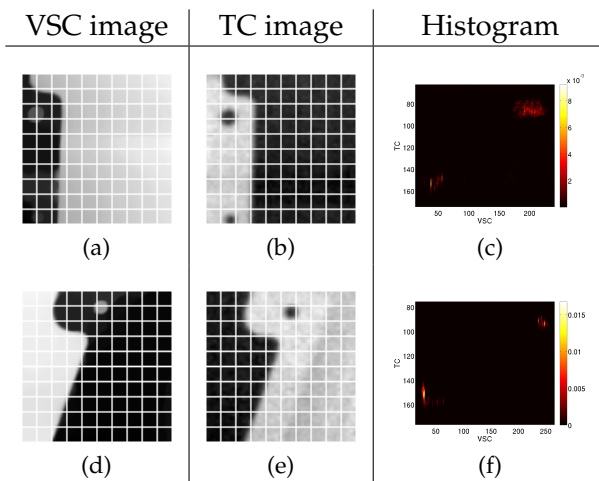


Figure 3.3: Different ROI's and their joint histograms. A grid is shown in the VSC and the TC image to emphasize the disparity between them. The gridsize is 10 pixels. In each histogram there are two maxima since the majority of pixels with high intensity in the VSC ROI correspond to pixels of low intensity in the TC ROI and vice versa. Since two maxima are sufficient for a linear relationship the assumption of local joint Gaussianity of the VSC and the TC image is valid.

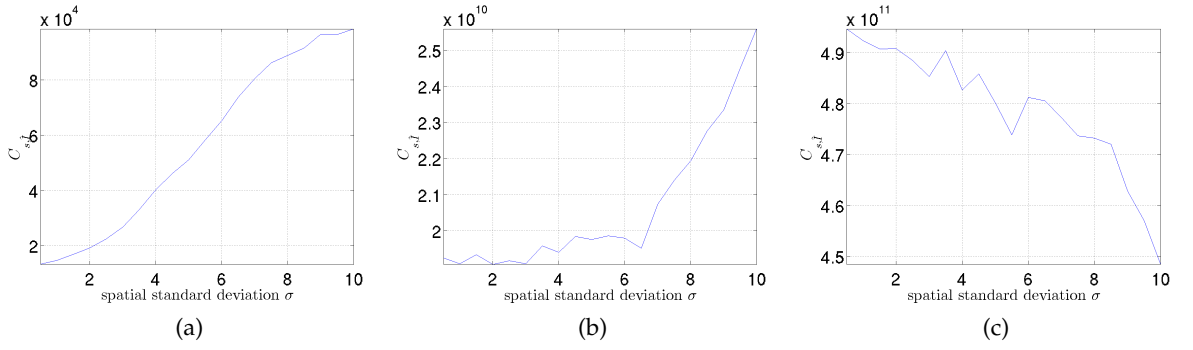


Figure 3.4: Median conditional variance $\hat{C}_{s|I}^{\sigma,a}$ for the likelihood window size (see eq. (3.33)) $a = 5$ (figure 3.4a), $a = 23$ (figure 3.4b) and $a = 33$ (figure 3.4c). We can see that for small values of a $\hat{C}_{s|I}^{\sigma,a}$ has a minimum at $\sigma < 2$, and for larger values of a the profile changes so that the minimum of $\hat{C}_{s|I}^{\sigma,a}$ is at $\sigma \geq 10$

where ω is a window function which we take to be constant within a subset $W \subset \Omega$

$$\omega(\mathbf{x}) = \begin{cases} \frac{1}{|W|-1} & 0 \leq x, y \leq a \\ 0 & \text{else} \end{cases} \quad (3.33)$$

Then $C_{s|I}^{\sigma,a}(\mathbf{x})$ becomes a local measure that measures how linear the intensities of y and I are within the sub domain W . The problem that arises is how large to set the window size a . If it is set too small the signal to noise ratio will be too small so that not enough information of the features in the TC and the VSC image are captured to robustly register them. On the other hand if a is set too large we eventually lose the local linearity between the TC and the VSC image. In figure 3.4 we have plotted the median conditional variance

$$\hat{C}_{s|I}^{\sigma,a} = \text{median} \left(C_{s|I}^{\sigma,a}(\mathbf{x}) \right) \quad (3.34)$$

as a function of σ for three fixed values of the window size a . In figure 3.4a ($a = 5$) $\hat{C}_{s|I}^{\sigma,a}$ has a minimum for $\sigma < 2$, and in figure 3.4c ($a = 33$) it is minimal for $\sigma \geq 10$. The profile of $\hat{C}_{s|I}^{\sigma,a}$ changes from monotonic increasing to monotonic decreasing for small to large values of a . Since we know the value for the scale parameter σ , $\sigma^* = 2$ from the ccd resolutions of the cameras, the idea find the optimal value a^* such that $\hat{C}_{s|I}^{\sigma,a^*}$ is minimal at $\sigma = \sigma^*$. For $a = 23$ this is the case as we see in figure 3.4b. Thus for the data in figure 2.4 $a^* = 23$ is the optimal

value so that $\hat{C}_{s|I}^{\sigma, a^*}$ has physically meaningful minimum $\sigma^* = 2$. The local data term E_{data} now has the form

$$E_{data}(\mathbf{d}) = \frac{1}{2} \int_{\Omega} \left(y(\mathbf{x}) - f(\mathbf{x}) \cdot \tilde{I}_d(\mathbf{x}) \right)^2 \cdot C_{s|\tilde{I}_d}^{\sigma^*, a^*}(\mathbf{x}) \left(C_{s|\tilde{I}_d}^{\sigma^*, a^*}(\mathbf{x}) + \lambda C_n \right)^{-2} \quad (3.35)$$

$$f(\mathbf{x}) = C_{y, \tilde{I}_d}(\mathbf{x}) C_{\tilde{I}_d, \tilde{I}_d}^{-1}(\mathbf{x}) \quad (3.36)$$

and together with our prior from chapter 3 the energy for the complete optical flow model is

$$E(\mathbf{d}) = E_{data}(\mathbf{d}) + \frac{\lambda}{2} \left(\sum_i \text{Det}(S(d_i)) \right) \quad (3.37)$$

The matrix $S(d_i)$ is the structure tensor (see eq. (3.4)) acting on each component of the optical flow \mathbf{d} . In this model we are making the assumption that the motion boundaries are locally linear. This assumption is valid for object boundaries with small curvature but as we will see in chapter ?? this assumption fails at junction points in the optical flow field, since those are where objects are partially occluding each other and moving in opposite directions.

3.6 The solution algorithm

To minimize 3.37 and obtain the optimum flow field $\hat{\mathbf{d}}$ we deploy a simple newton scheme with a nested linearization of 3.37. The linearized model is solved by a conjugate gradients algorithm with block Jacobi preconditioning. The problem with this approach is that the regularizer is quartic in the flow field components and thus the linearization becomes numeric instable for the initial steps of the algorithm.

The problem arises in step 9 of the iterative algorithm. The second functional derivative Q_k of the energy model 3.37 consists of one part coming from the likelihood and one part coming from the prior, $Q_k = Q_k^{data} + \lambda Q_k^{reg}$. The matrix Q_k^{reg} is the second derivative of the prior in 3.37 with respect to \mathbf{d} . At small k its eigenvalues are small due to the initial guess $\mathbf{d}_0 = 0$. The matrix Q_k^{data} is the second derivative of the likelihood in eq. (3.37). In regions where there is no motion the eigen values of Q_k^{data} are also small. This makes the linearized solution in step 9 numerically instable. Our solution to this problem is to extend

Algorithm 1 Optical Flow with Structure Tensor prior

Initialize $\mathbf{d}_0 = 0$
Set $\mathbf{r}_0 = \frac{\delta E(\mathbf{d})}{\delta \mathbf{d}}(\mathbf{d}_0)$
scale $s = s_{\text{Max}}$
while $s > 1$ **do**
 downsample $y_s = G_s \star y_0, I_s = G_s \star I_0$
 while $\|\mathbf{r}\| > \epsilon$ or $k < N$ **do**
 set $\mathbf{d}_{k+1} = \mathbf{d}_k + \alpha \delta$
 expand $E(\mathbf{d}_{k+1}) = E(\mathbf{d}_k) + \alpha \mathbf{b}_k^T \delta + \frac{\alpha^2}{2} \delta^T Q_k \delta$
 solve $Q_k \delta = \mathbf{b}_k$ for δ with conjugate gradients and suitable preconditioning
 compute $\mathbf{d}_{k+1} = \mathbf{d}_k + \alpha \delta, k \rightarrow k + 1$
 end while
 upsample d_N , set $d_0 = d_N, k = 0$
 $s = s - 1$
end while

3.37 to include an L_2 prior on the flow field \mathbf{d} but with a small lagrange multiplier λ_2

$$E(\mathbf{d}) = \int (y - \hat{s}_{I,\mathbf{d}})^2 \cdot C_{s|I_d} + \frac{\lambda}{2} \left(\sum_i (\text{Det}(S(d_i)) + \lambda_2 \|\nabla d_i\|) \right) \quad (3.38)$$

With the L_2 prior in 3.38 the linearized solution in step 9 becomes numerically stable.

3.7 Results

3.8 Uni-Modal Data

We will now discuss the results of our optical flow method on the middleburry data set for which there exists ground truth (GT). As the GT is the true flow field for the data we use it to asses the quality of the computed optical flow. To do this we define the Endpoint error (EPE) and the angular error (AE) as

$$e_{EPE} = \|\mathbf{v} - \mathbf{v}_{gt}\| \quad (3.39)$$

$$e_{AE} = \cos(\angle(\mathbf{v}, \mathbf{v}_{gt})) \in \{-1, 1\} \quad (3.40)$$

The EPE e_{EPE} measures how well the computed optical flow v fits the true optical flow v_{gt} . In cases where v does not match v_{gt} well, we would still like to check how both vectors are aligned. This alignment is depicted by the AE values ranging between -1 , for minimal alignment (worst case), and 1 for maximal alignment (best case).

#Need to work on this chapter

3.8.1 Structure Tensor Prior

3.8.2 Total Variation Prior

Figure	Filtersize	Median, Min, Max EPE	Median, Min, Max AE
figure 3.7a	7	2.36, 0.01, 7.24	0.42, -1.00 , 1.00
	9	1.32, 0.00, 6.02	0.87, -1.00 , 1.00
	11	1.15, 0.00, 6.45	0.91, -1.00 , 1.00
figure 3.7f	7	0.84, 0.01, 13.35	0.87, -1.00 , 1.00
	9	0.46, 0.01, 8.23	0.97, -1.00 , 1.00
	11	0.40, 0.00, 8.25	0.98, -1.00 , 1.00
figure 3.8a	7	0.47, 0.01, 5.22	0.97, -0.96 , 1.00
	9	0.28, 0.00, 3.71	0.99, -1.00 , 1.00
	11	0.25, 0.00, 2.50	0.99, -1.00 , 1.00
figure 3.8f	7	0.44, 0.00, 2.73	0.98, -1.00 , 1.00
	9	0.34, 0.00, 2.65	0.99, -1.00 , 1.00
	11	0.30, 0.00, 3.12	0.99, -1.00 , 1.00

Table 3.1: EPE and AE analysis

EPE and AE values for different region of interests and filter sizes (Figures 3.7a to 3.8f). The second column shows the median, minimum and maximum EPE per roi. The third column shows the median, minimum and maximum AE per roi. The table shows that the EPE gets better with increasing filtersize. Despite this the values for roi's with non-linear geometry (figure 3.7) generally have higher EPE values than the roi's with linear or constant geometry (figure 3.8)

In figure 3.5 the rubber whale sequence of the middleburry data set is shown, and in figure 3.5b the corresponding ground truth v_{gt} . In figure 3.5d the computed flow-field v is shown for a filter size of 11, while in figure 3.5c the resulting flow for the TV model is shown. Figures 3.7 and 3.8 show different region of interrests (roi) for which the EPE and AE are shown on a pixel basis for the structure tensor model and Figures 3.9 and 3.10 show the same for the TV model. We can observe from the comparison between figures 3.5d and 3.5c that the TV model produces

Figure	Median, Min, Max EPE	Median, Min, Max AE
figure 3.9a	1.38, 0.00, 5.83	0.92, -1.00, 1.00
figure 3.9f	0.20, 0.00, 3.34	1.00, -1.00, 1.00
figure 3.10a	0.04, 0.00, 2.08	1.00, -1.00, 1.00
figure 3.10f	0.09, 0.00, 2.06	1.00, -1.00, 1.00

Table 3.2: EPE and AE analysis

EPE and AE values for different region of interests for the TV model (Figures 3.9a to 3.10f). The first column shows the median, minimum and maximum EPE per roi. The second column shows the median, minimum and maximum AE per roi. Compared to table 3.1 the median EPE is lower for nearly all roi's, while the median AE do not differ that much

Figure	Filtersize	Median, Min, Max EPE	Median, Min, Max AE
figure 3.11a	7	0.73, 0.00, 6.80	0.99, -1.00, 1.00
	9	0.60, 0.00, 7.29	0.99, -1.00, 1.00
	11	0.96, 0.01, 15.60	0.98, -1.00, 1.00
figure 3.11f	7	0.36, 0.00, 7.00	1.00, 0.00, 1.00
	9	0.27, 0.00, 6.79	1.00, 0.00, 1.00
	11	0.41, 0.01, 6.55	1.00, 0.00, 1.00

Table 3.3: EPE and AE analysis

EPE and AE values for different region of interests and filter sizes (Figures ?? to ??). Since the motion boundaries in figure 3.6a are all curvilinear there is no correlation between the filtersize and the EPE.

smoother results which are closer to the ground truth (figure 3.5b). In tables 3.1 and 3.4 the median values for the EPE and AE in various region of interrests are listed. Indeed we can observe that the EPE for the TV is approximately half the value of that of the structure tensor model. We chose the median as opposed to the mean EPE as it is robust outlier values of the EPE at single pixel locations and thus gives a better assessment of the quality of the flow within the roi.

Table 3.1 shows also how the EPE and the AE vary with increasing filtersize: The EPE decreases while the AE increases. In figure 3.8 the roi's have mostly a constant motion field or a motion field with linear boundary, thus according to table 3.1 they have lower EPE values then the roi's in figure 3.7. The roi with the largest discrepancy from the group of linear motions is figure 3.7a which depicts a rotating wheel. Since the wheel is largely free of texture, the motion field (figure 3.7d) is penalized by the structure tensor prior in such a way that it acquires spurious linear motion boundaries. This is the reason for its high EPE value. The roi in figure 3.7f shows another case of a motion field violating the assumption of linear motion boundaries. In the ground truth roi in figure 3.7j

Figure	Median, Min, Max EPE	Median, Min, Max AE
figure 3.12a	0.44, 0.00, 6.12	1.00, -1.00, 1.00
figure 3.12f	0.12, 0.01, 7.38	1.00, 0.00, 1.00

Table 3.4: EPE and AE analysis

EPE and AE values for different region of interests for the TV model (Figures ?? to ??). The first column shows the median, minimum and maximum EPE per roi. The second column shows the median, minimum and maximum AE per roi. Compared to table 3.1 the median EPE is lower for nearly all roi's, while the median AE do not differ that much

there are two junction points where three objects are occluding and moving against each other. This type of motion is penalized by the structure tensor prior so that the flow at these points is oversmoothed. The TV model (ref!) like the structure tensor model penalizes non linear motion boundaries. figure 3.9d shows the result of the TV model for the wheel roi. Just like in the structure tensor model, the flow on the circumference of the wheel is heavily penalized resulting in high EPE values and wrong AE values (see table 3.4). figure 3.9i shows the resulting flow of the TV model at the two junctions in figure 3.9f. Similar to our proposed prior the flow is oversmoothed at the junctions resulting in high EPE values (see table 3.4).

On the otherside both models are faithful to roi's with constant motion or linear motion boundaries (see figures 3.8 and 3.10). In figure 3.8d we see that the structure tensor model inflicts more of the texture from the underlying data (figure 3.8a) on the estimated flow than the TV model (see figure 3.10d for the result of the TV model) thus leading to a slightly higher EPE value (table 3.1). Figure 3.8i shows an example of an roi with a linear motion boundary for the structure tensor model. Comparing it to the corresponding result for the TV model figure 3.8i, we see that the structure tensor model produces sharper lineat motion boundaries.

In figure 3.6 another sequence of the middleburry data set is shown. In this sequence the camera is rotating around a hydrangea. As the ground truth shows there are no linear motion boundaries, thus only the constant motion of the background is correctly detected (upto some artifacts in the upper left corner in figure 3.6d), see the EPE and AE values in figure 3.11 and table 3.3.

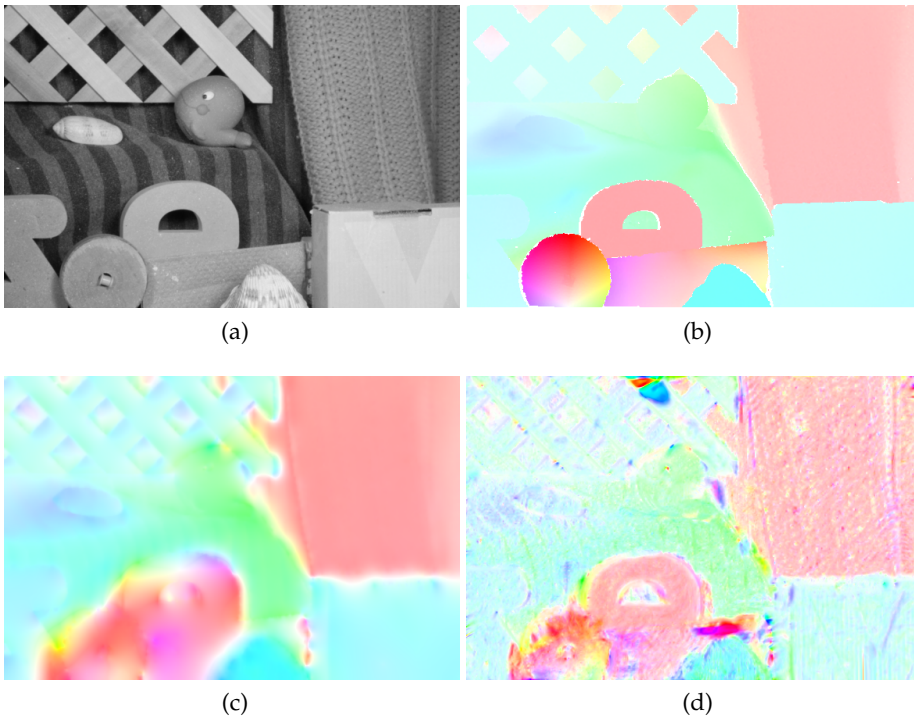


Figure 3.5: Rubberwhale Sequence

Figure 3.5a shows one frame of the sequence. figure 3.5d shows the estimated optical flow, figure 3.5c the result of the TV model and figure 3.5b shows the provided ground truth

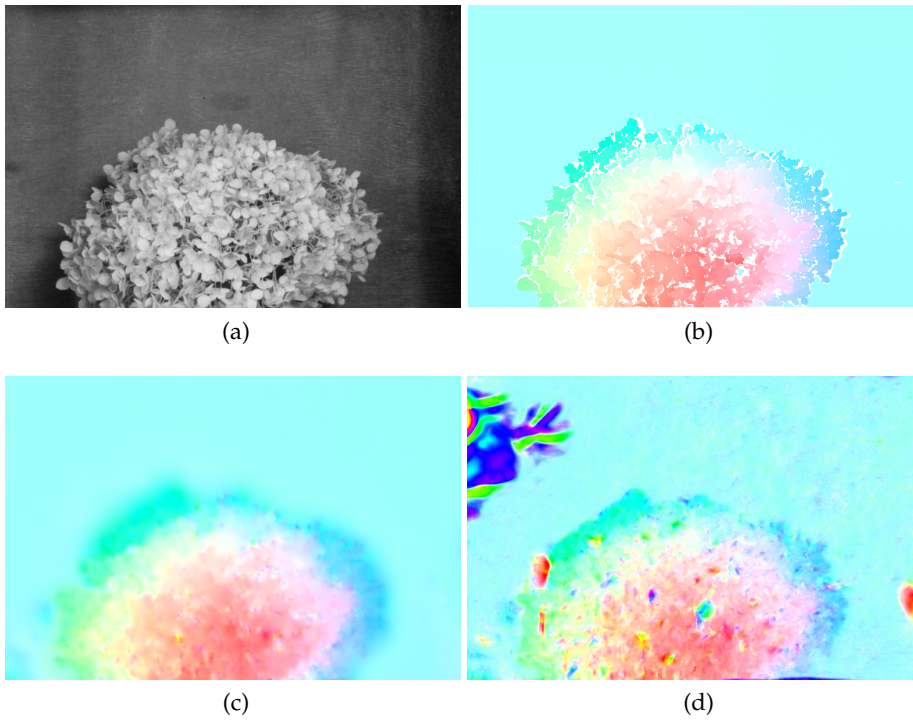


Figure 3.6: Hydrangea Sequence

Figure 3.6a shows one frame of the sequence. figure 3.6d shows the estimated optical flow, figure 3.6c the result of the TV model and figure 3.6b shows the provided ground truth


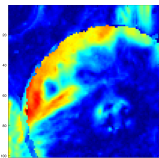
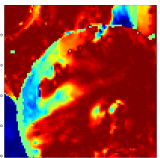
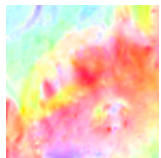


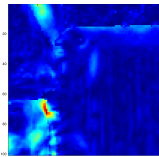
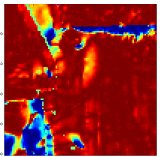
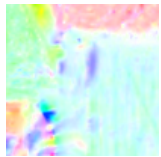

Region of interest	EPE	AE	Resulting Flow	Ground Truth v_{gt}
				
				

Figure 3.7: Error Analysis ST model: This figure shows two examples of motion field with nonlinear boundaries. In figure 3.7c we see that along the circumference of the wheel the EPE has the largest values and in figure 3.7h the is largest the junction point where three objects are moving against each other.

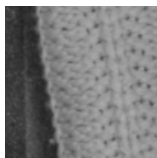
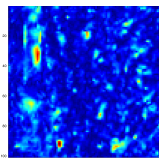
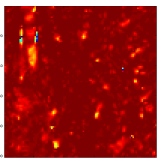
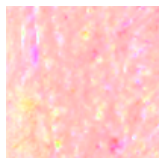


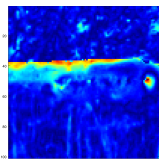
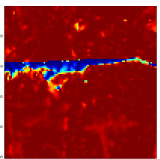
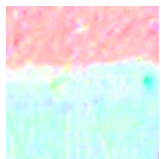

Region of interest	EPE	AE	Resulting Flow	Ground Truth v_{gt}
				
				

Figure 3.8: Error Analysis ST: This figure shows two examples of motion fields with linear boundaries. In figures 3.8d and 3.8i we can see that the resulting flow with texture inflicted from the data. Nevertheless the EPE values are nearly homogenous and small (see figures 3.8c and 3.8h)

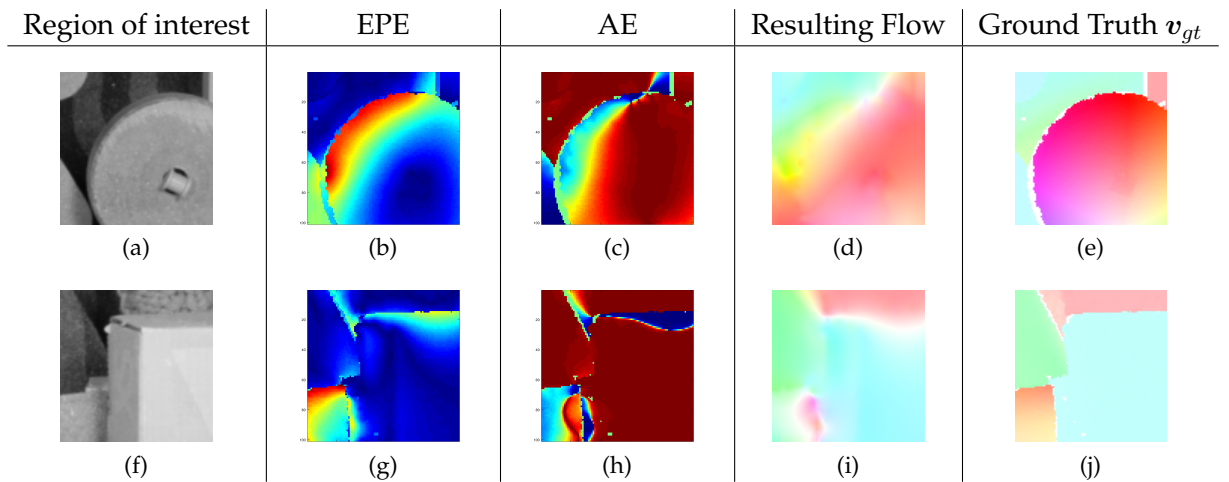


Figure 3.9: Error Analysis TV model: This figure shows two examples of motion field with nonlinear boundaries. In figure 3.9c we see that along the circumference of the wheel the EPE has the largest values and in figure 3.9h the is largest the junction point where three objects ar moving against each other.

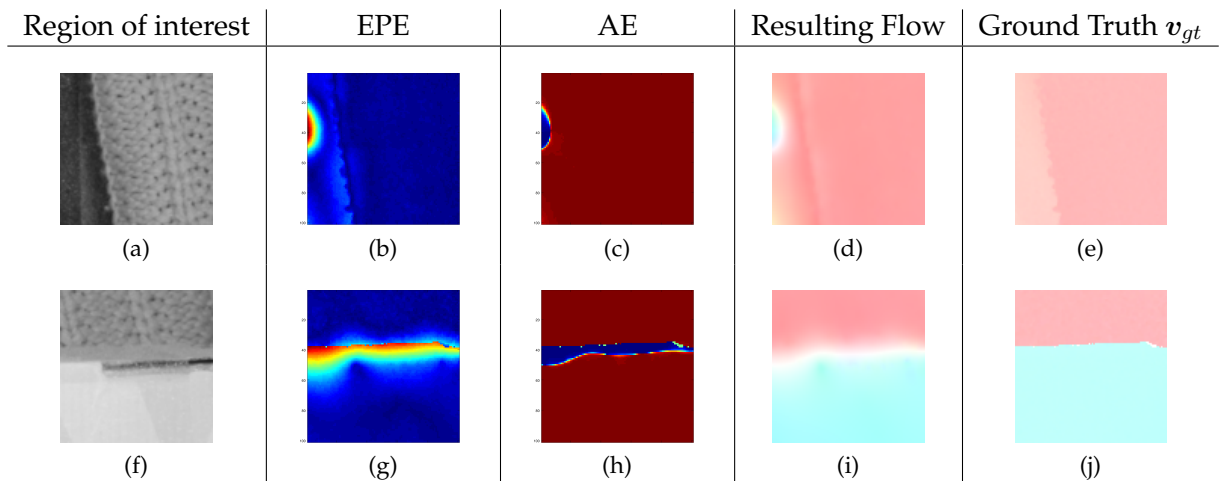


Figure 3.10: Error Analysis TV: This figure shows two examples of motion fields with linear boundaries. In figures 3.10d and 3.10i we can see that the resulting flow with texture inflicted from the data. Nevertheless the EPE values are nearly homogenous and small (see figures 3.10c and 3.10h)

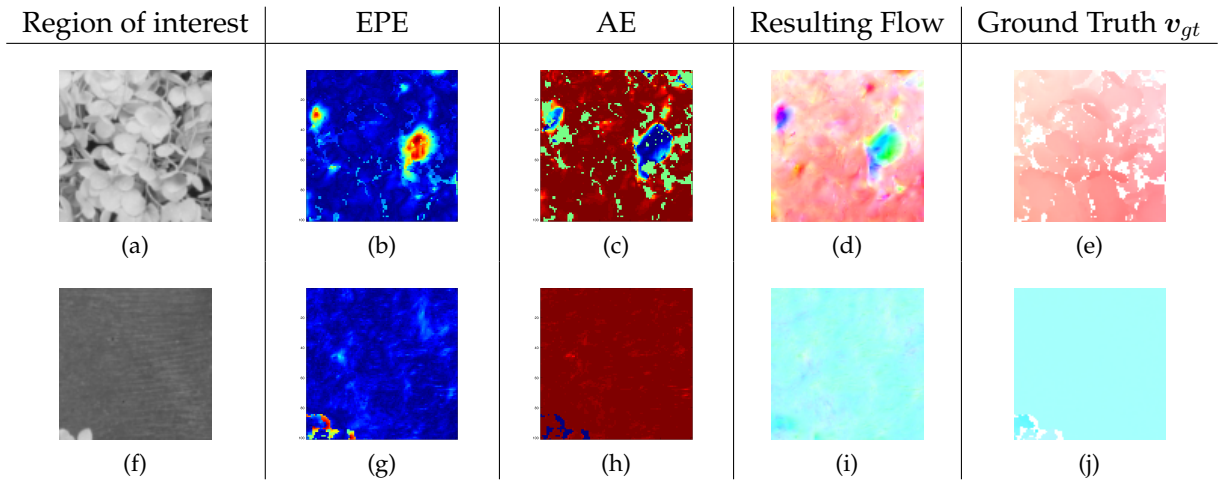


Figure 3.11: Error Analysis:
Second Column: Endpoint Error, Third Column: Angular Error.

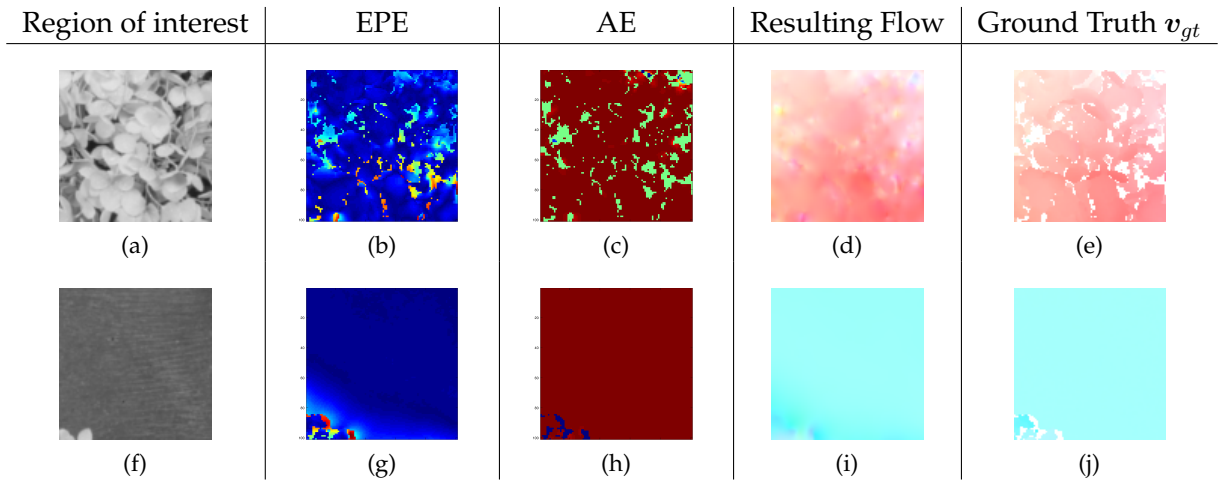


Figure 3.12: Error Analysis:
Second Column: Endpoint Error, Third Column: Angular Error.

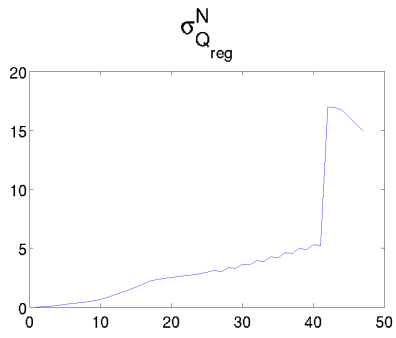


Figure 3.13: $\lambda_2 = 10^{-3}$

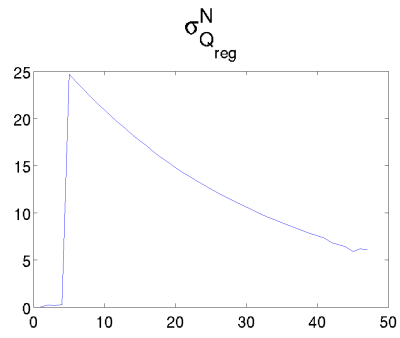


Figure 3.14: $\lambda_2 = 10^{-6}$

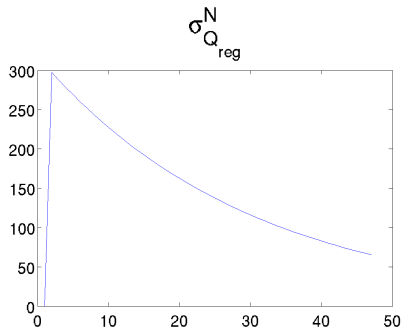


Figure 3.15: $\lambda_2 = 10^{-9}$

Figure 3.16: Analysis of the largest eigenvalue σ_Q^i of Q^{reg}

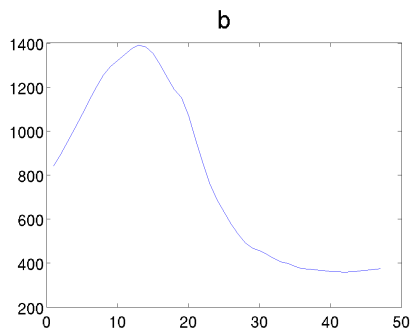


Figure 3.17: $\lambda_2 = 10^{-3}$

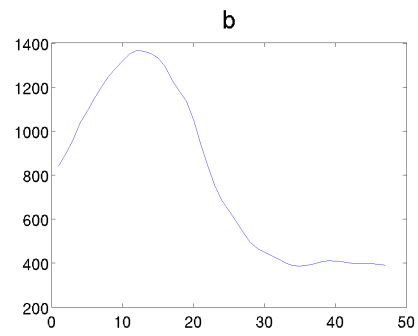


Figure 3.18: $\lambda_2 = 10^{-6}$

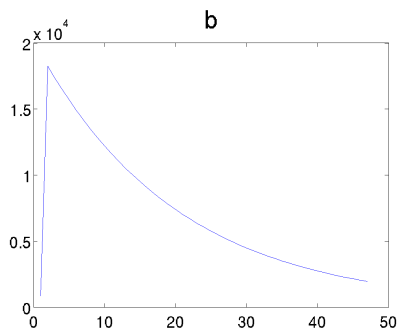


Figure 3.19: $\lambda_2 = 10^{-9}$

Figure 3.20: Analysis of the Euler-Lagrange vector b

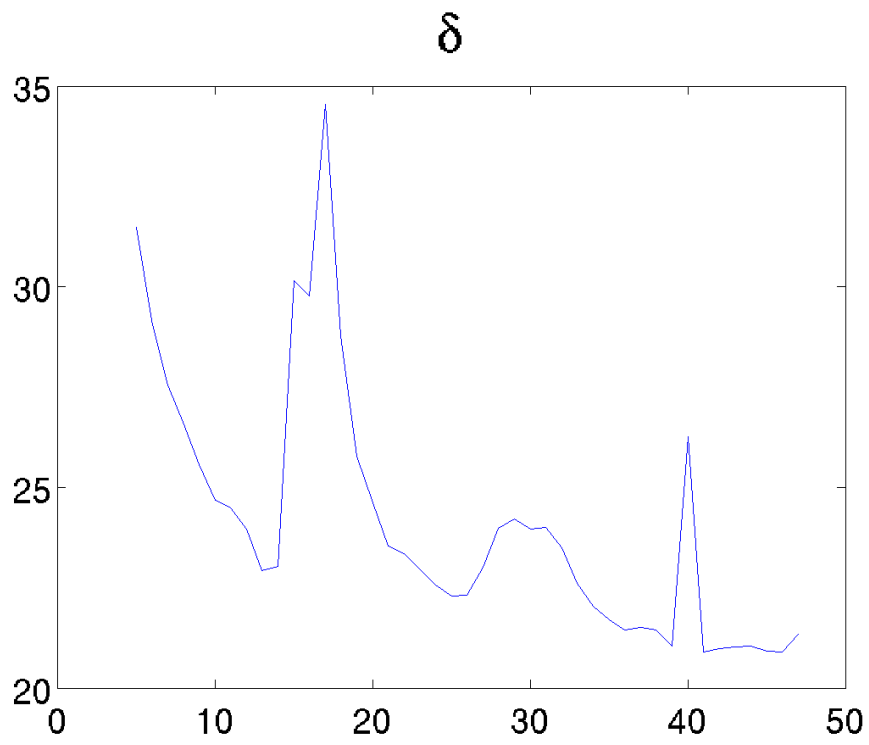


Figure 3.21: Analysis of the Euler-Lagrange vector $\delta_{eq. (??)}$

3.9 Eigenvalue analysis and the stabilization parameter

λ_2

In chapter 3.3 we stated that the L_2 term in eq. (3.38) is needed to support the numerical stability of the model. We will back this statement now. Figures 3.13, 3.14 and 3.15 show the largest eigenvalue of Q_{reg}^i , σ_Q^i at each iteration on the coarsest scale of the pyramid for different values of λ_2 . They all show that σ_Q^N rises to a maximum after which it decreases and converges. The initial value of σ_Q^i is of the order of λ_2 indicating that in the initial steps the L_2 term in eq. (3.38) governs the regularization. As the number of iterations increases the structure tensor determinant gets more weight, until the point where its influence overweighs that of the L_2 term. As can be seen this point comes after fewer iterations the smaller λ_2 is set. On the other side Figures 3.17, 3.18 and 3.19 show the vector \mathbf{b} , that is the Euler-Lagrange equation vector for different values of λ_2 . Comparing the magnitude of \mathbf{b} in Figures 3.17, 3.18 and 3.19 we see that for $\lambda_2 = 10^{-9}$ \mathbf{b} is several orders of magnitude larger than the other cases, which leads to longer convergence rates or numerically instable solution. This means we have a tradeoff between

- $\lambda_2 \sim 10^{-3}$: Faster convergence but less influence of structure tensor (need $i > 40$ iterations for ST to act)
- $\lambda_2 \sim 10^{-9}$: slower convergence but more influence of structure tensor (need only $i > 1$ iterations for ST to act)

We choose $\lambda_2 = 10^{-6}$ since in this case \mathbf{b} is of the same order of magnitude as for $\lambda_2 = 10^{-3}$ but as we see in figure 3.14 the structure tensor only needs 4 iterations until its eigenvalues overweigh the eigenvalues of the L_2 term. We also choose $N = 10$ for the number of iterations per pyramid scale, since according to ?? the update vector δ gets unstable after 15 iterations.

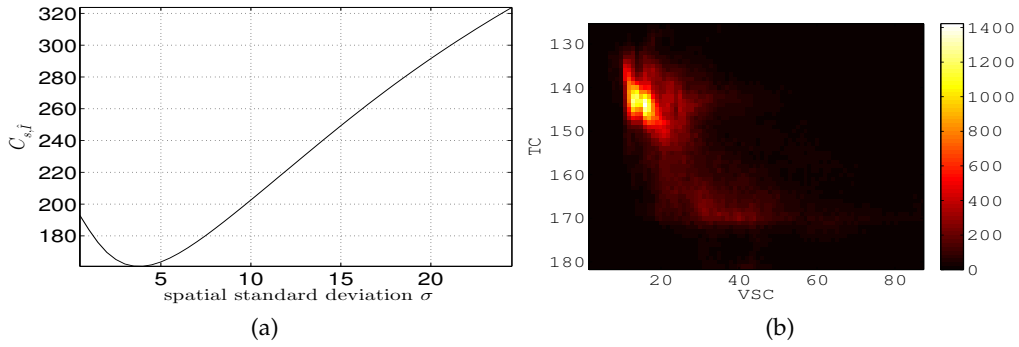


Figure 3.22: 3.22a: Dependence of $C_{s|I}$ on the scaling parameter σ . 3.22b: Joint Histogram $p(y, I)$ of the TC and smoothed VSC image pair y and \tilde{I} at the optimum $\sigma^* = 4$, the scale at which y and \tilde{I} are maximally linear.

3.10 Multimodal Optical Flow

3.10.1 Estimation of the resolution parameter σ

3.10.2 Structure Tensor Prior

3.10.3 Total Variation Prior

Our optical flow model eq. (??) is based on the assumption that the modalities to be registered have a linear relationship in their intensity spectrum. This is not the case for TC images and VSC images of arbitrary objects. However in the case of bare CFRPs the linearity assumption holds. CFRPs are black bodies when in thermal equilibrium at 30 deg C since the emmissivity of carbon is approximately 0.98 (see [?]). It is in this case that in the amplitude image in ??b the CFRP has a uniform amplitude. In the visual spectrum domain (??d) the CFRP is not a perfect black body due to the reflective nature of the epoxy coating, however the epoxy coating is uniformly distributed so that the reflections do not cause image gradients, which are not correlated to geometric features. Since the TC and the VSC have different resolutions we must take the difference in resolution into account. This difference in resolution is encoded in the scale parameter σ of our local likelihood model in eq. (2.157). The local conditional variance $C_{s|I}(\mathbf{x})$ in eq. (??) is a measure for the similarity of the TC image y , and thus s and the VSC image I with a local subdomain $W \subset \Omega$. The local conditional variance $C_{s|I}(\mathbf{x})$ has two parameters we need to estimate: the scale parameter σ from

the likelihood in eq. (2.157) and the window size a of the window function ω . Since $C_{\tilde{s}|\tilde{I}}(\mathbf{x})$ varies spatially we compute its median value $\hat{C}_{\tilde{s}|\tilde{I}}$. In ?? we have plotted for various window sizes a the median conditional variance $\hat{C}_{\tilde{s}|\tilde{I}}$ over the filter size σ . We can see that for window sizes $a \leq 23$ $\hat{C}_{\tilde{s}|\tilde{I}}$ has minima at $\sigma \approx 0$ while for larger window sizes $a \geq 31$ it tends to be minimal at filtersizes $\sigma > 6$. ?? show their optimum σ^* plotted over the window size a . We see that window sizes $a < 21$ and $A > 31$ lead to unrealistic scale differences $\sigma^* \approx 0$ and $\sigma^* \geq 6$, since the actual difference in scale must be $\sigma \approx 2$ judged by the resolutions of the VSC and the TC. This value is produced only at $a = 23$ and $a = 27$ and we choose $a = 23$ since $C_{\tilde{s}|\tilde{I}}(\mathbf{x})$ is smaller compared to the case $a = 27$.

In ?? we show the resulting optical flow for different region of interests (roi). Figures 3.23a and 3.23f show the resulting optical flow \mathbf{d} which match the corresponding VSC image I and TC image y in the table. For each roi we computed the joint histogram $p(y, I)$ (Figures 3.23b and 3.23g). In figure 3.23b $p(y, I)$ has two isolated maxima which is sufficient for a linear relationship between y and I . In figure 3.23g the linearity is obstructed to a minor degree since the TC image in figure 3.23j has a slight structural difference in the lower left corner compared to figure 3.23h.

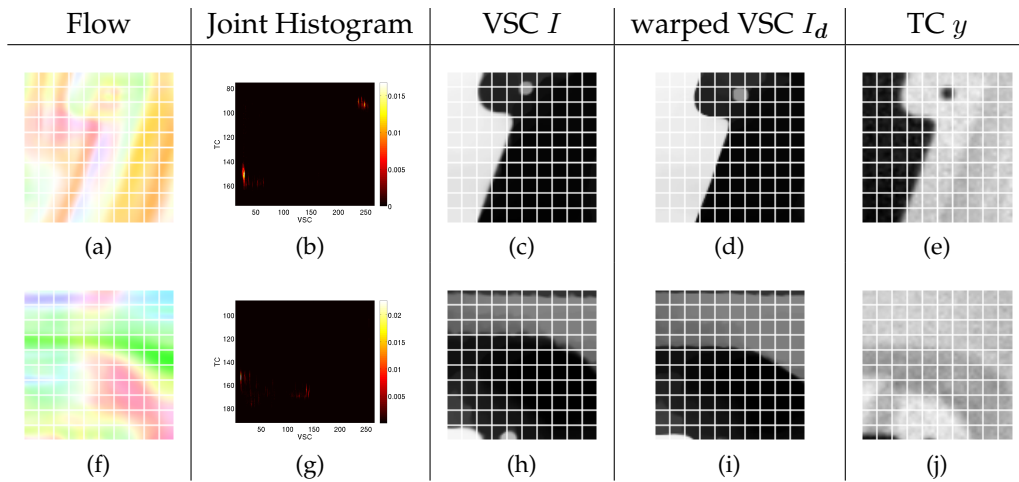


Figure 3.23: Multimodal Optical Flow: The resulting flow d , VSC image I , the warped VSC I_d , the TC image y as well as the joint histogram $p(y, I)$ are shown for different region of interests. We can observe that the boundaries of the flow are blurred. This comes from the window function ω in the local likelihood. The joint likelihood $p(y, I)$ was evaluated only for the roi's. It has at most two maxima, which suffices to constitute a linear relationship between y and I . A grid is overlaid on the roi's for I , I_d and y with 10 pixels per element to visually assess the quality of the flow. We can see the larger features are correctly matched, while smaller features are matched in a suboptimal fashion

??

4 The Generalized Newton Algorithm

4.1 Motivation

In this section we want to motivate a new type of algorithm for the minimization of the energy

$$E(\phi, \nabla\phi) = E^{data}(\phi) + \lambda E^{prior}(\nabla\phi) = \int_{\Omega} \mathcal{E}_{tot}(\phi(\mathbf{x}), \nabla\phi(\mathbf{x})) d^2x \quad (4.1)$$

based on the considerations in section 2.2. Traditional algorithms for the minimization of the energy functional in eq. (4.1) are based around the concept that a variation of ϕ around the minimum of E lead to a set of vanishing differentials $[\mathcal{E}_{tot}]$ called the Euler-Lagrange differentials

$$\left. \frac{dE(\phi')}{dt} \right|_{t=0} = \int_{\Omega} [\mathcal{E}_{tot}](\phi^*(\mathbf{x}), \nabla\phi^*(\mathbf{x})) v^{\phi}(\mathbf{x}) d^2x = 0 \quad (4.2)$$

$$\phi' = \phi^* + \tau v^{\phi} \quad (4.3)$$

We can derive the same result if we take ϕ' to be the result of the action of a one parameter Lie group $g_{\gamma t}$ acting only on ϕ^* and not on Ω

$$\left. \frac{dE(\phi')}{dt} \right|_{t=0} = \int_{\Omega} V_e^{\phi}(\mathcal{E}_{tot}) d^2x = 0, \quad \phi' = g_{\gamma t} \circ \phi^* \quad (4.4)$$

$$V_e^{\phi} = v^{\phi} \frac{d}{d\phi} + \frac{dv^{\phi}}{dx_{\nu}} \frac{d}{d\partial_{\nu}\phi} \quad (4.5)$$

The vector V_e^{ϕ} is obtained from eqs. 2.55 and 2.56 simply by setting the spacial variations ω_i^{μ} in the basis operators X_e^i to zero, $\omega_i^{\mu} = 0$ and setting $v^{\phi} = \sum_i \alpha_i \omega_i^{\phi}$. Using integration by parts we show that eq. (4.2) and eq. (4.4) are equal.

One of the basic algorithms for solving the minimization problem in eq. (4.2) is the method of steepest descent (citation!!!). Beginning with an initial guess ϕ^0 , the basic idea of steepest descent is to compute a new estimate of the field ϕ by advancing a previous estimate ϕ^n along the negative direction of the gradient of

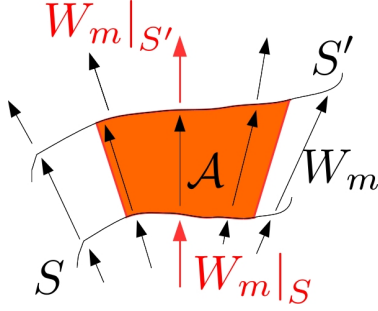


Figure 4.1: This figure shows a transformation of the level-set S to S' along the vector $\mathbf{W}_m(\mathbf{x})$. The region $\mathcal{A} \subset \Omega$ is the region a section of S traverses as it is shifted along \mathbf{W}_m to the end position S' . If the divergence of \mathbf{W}_m vanishes, this means that the incoming flux of \mathbf{W}_m equals the outgoing flux (both indicated by the red arrows), $\mathbf{W}_m|_S = \mathbf{W}_m|_{S'}$

$E(\phi, \nabla\phi)$ which is provided by the Euler-Lagrange differentials $[\mathcal{E}_{tot}]$

$$\phi^{n+1} = \phi^n - \tau^\phi [\mathcal{E}_{tot}] (\phi^n, \nabla\phi^n) \quad (4.6)$$

The scheme is repeated (see algorithm 2) until either the Euler-Lagrange differentials vanish or a fixed number N of iterations is reached.

In the eqs. 4.2 and 4.4 we take only the variation of the field ϕ into account. However the discussion in chapter 2.3 which led to Noether's theorem is based upon a more general Lie group \mathbb{G} which also contains possible variations of the coordinate frame (see eq. (??))

$$\mathbf{x}' = \mathbf{x} + \tau^\Omega \mathbf{v}(\mathbf{x}) \quad (4.7)$$

where τ is a parameter that controls the extent of the deformation of Ω , similar to how τ^ϕ controls the deformation of ϕ . In the following we would like to introduce a methodology that enables the use of the **entire group \mathbb{G} for the minimization of the total energy E in eq. (4.1)**. Our new methodology is centered around the concept of steepest descent for the **spacial coordinate frame Ω of the form**

$$\mathbf{x}^{n+1} = \mathbf{x}^n - \tau^\Omega \mathbf{b}(\mathbf{x}^n) \quad (4.8)$$

The exact form of the vector $\mathbf{b}(\mathbf{x})$ will soon be deduced, now we wish to give an intuitive idea of $\mathbf{b}(\mathbf{x})$ should look like. In section 2.2.1 (??) we explained that the role of the prior E^{prior} is to enforce certain geometric constraints onto the level-sets S_X of the minimizers ϕ^* . The geometric constraints are encoded in the Lie algebra \mathcal{G} of the Lie group \mathbb{G} under which E^{prior} and thus E is invariant. If \mathbb{G}

is a pure spacial Lie group, $v^\phi = 0$, then eq. (2.110) simplifies to

$$\frac{d}{dt} g_{\gamma_t} \circ E \Big|_{t=0} = \int_{\Omega} \left(\sum_m \alpha_m \operatorname{div}(\mathbf{W}_m) - V_e^\Omega(\phi) [\mathcal{E}_{tot}] \right) d^2x = 0 \quad (4.9)$$

which is independent of the integration region. Since eq. (4.9) must hold for any coefficient vector α , by virtue of the expansion $V_e^\Omega = \sum_m \alpha_m X_e^{m,\Omega}$ the individual divergences must satisfy

$$\int_{\Omega} (\operatorname{div}(\mathbf{W}_m)) d^2x = \int_{\Omega} \left(X_e^{m,\Omega}(\phi) [\mathcal{E}_{tot}] \right) d^2x \quad (4.10)$$

Eq. (4.10) must hold for any integration domain Ω which means that the integrands themselves must be equal

$$\operatorname{div}(\mathbf{W}_m) = X_e^{m,\Omega}(\phi) [\mathcal{E}_{tot}] \quad (4.11)$$

By Gauss' law the integrated divergence of \mathbf{W}_m within any subset $\mathcal{A} \subset \Omega$ equals the integral of the flux of \mathbf{W}_m over the surface $\partial\mathcal{A}$

$$\int_{\mathcal{A}} \operatorname{div}(\mathbf{W}_m) d^2x = \int_{\partial\mathcal{A}} \mathbf{W}_m d\mathbf{S} \quad (4.12)$$

thus from eq. (4.11) we have

$$\int_{\partial\mathcal{A}} \mathbf{W}_m d\mathbf{S} = \int_{\mathcal{A}} \left(X_e^{m,\Omega}(\phi) [\mathcal{E}_{tot}] \right) d^2x \quad (4.13)$$

In figure 4.1 we have depicted the situation where a level-set S is shifted along the vector \mathbf{W}_m with S' is the result of the shift and \mathcal{A} is the region traversed by the shift of a section of S . We denote this transformation by $g_{\mathbf{W}}^m(t) \in \mathbb{G}^{\mathbf{W}}$ where $\mathbb{G}^{\mathbf{W}}$ is another group. Whether or not it is a Lie group possibly intersecting with \mathbb{G} , $\mathbb{G}^{\mathbf{W}} \cap \mathbb{G} \neq \{\emptyset\}$ is not known by the author. However this irrelevant for the following argument. The boundary $\partial\mathcal{A}$ consists of two lines tangential to \mathbf{W}_m besides the sections of S and S' . Since the flux over the tangential lines vanishes we have

$$\int_S \mathbf{W}_m d\mathbf{S} - \int_{S'} \mathbf{W}_m d\mathbf{S}' = \int_{\mathcal{A}} X_e^{m,\Omega}(\phi) [\mathcal{E}_{tot}] dx^2 \quad (4.14)$$

From Eq. (4.14) we see that the Euler-Lagrange differentials $[\mathcal{E}_{tot}]$ and the basis element $X_e^{m,\Omega}$ act as a source that drives the transformation $g_{\mathbf{W}}^m(t)$ in the sense that the S' propagates until it traverses a region in which the integrand of the right hand side in eq. (4.14) vanishes. More precisely eq. (4.14) can interpreted as an equation of motion for the normal vector on S' , $\mathbf{n}_{S'}$.

The main question we want to raise in this section is: Does the eq. (4.14), understood as an equation of motion of the normal vector \mathbf{n}_S on S under the influence of the Euler-Lagrange differentials $[\mathcal{E}_{tot}]$, suffice to explain the evolution of the level-set S ? In other words can we formulate an equation of motion for a tangential vector \mathbf{b}_S on S which satisfies $\mathbf{b}_S \perp \mathbf{n}_S$? The tangent vector \mathbf{b}_S is taken to be normalized.

Technically speaking, an operator B that would potentially manipulate S along the tangent \mathbf{b}_S can be formulated in terms of the Cartesian basis $\{\partial_\mu\}$

$$B = b_S^\mu(\mathbf{x}) \partial_\mu \quad (4.15)$$

A naive answer to our second question is: No! To illustrate this we construct a one dimensional potential Lie group g_t^B for which B is the basis of its algebra

$$\left. \frac{d}{dt} g_t^B \right|_{t=0} = B \quad (4.16)$$

The very definition of the level-set S is that along the tangent \mathbf{b}_S the field variable $\phi(\mathbf{x})$ is constant

$$\left. \frac{d}{dt} g_t^B \circ \phi \right|_{t=0} = B\phi = b_S^\mu \partial_\mu \phi = 0 \quad (4.17)$$

As a result we are led to believe that the energy E in eq. (4.1) is invariant to the transformation g_t^B since it depends on ϕ which in turn obeys eq. (4.17). Thus by Noether's theorem there must exist a vector valued function $\mathbf{W}_B(\mathbf{x})$ for which similar to eq. (4.10) the identity

$$\int_\Omega (\text{div}(\mathbf{W}_B)) d^2x = \int_\Omega (B(\phi) [\mathcal{E}_{tot}]) d^2x \quad (4.18)$$

must hold. However due to eq. (4.17) the integrand of the right hand side of eq. (4.18) vanishes uniformly on Ω , leading to the conclusion that the divergence of $\mathbf{W}_B(\mathbf{x})$ must vanish uniformly too

$$\text{div}(\mathbf{W}_B)(\mathbf{x}) = 0, \quad \forall \mathbf{x} \in \Omega \quad (4.19)$$

The fallacy in the argumentation of our naive answer lies in the last statement in eq. (4.19). We recall the generic form of the divergence vector \mathbf{W}_m from eq. (2.108)

$$W_m^\mu = \omega_m^\mu \mathcal{E} + \sum_i \omega_i^\mu \left(\omega_m^\phi - X_e^{m,\Omega}(\phi) \right) \frac{\delta \mathcal{E}}{\delta (X_e^i \phi)} \quad (4.20)$$

and replace $X_e^{m,\Omega}$ with B as the generating operator, thereby setting $\omega_m^\phi = 0$ (B is a purely spacial operator)

$$W_B^\mu = b_S^\mu \mathcal{E} - \sum_i \omega_i^\mu \cdot B(\phi) \frac{\delta \mathcal{E}}{\delta (X_e^i \phi)} \quad (4.21)$$

Again by virtue of eq. (4.17) W_B simplifies even more

$$W_B^\mu = b_S^\mu \mathcal{E} \quad (4.22)$$

To show why $\text{div}(\mathbf{W}_B) = 0$ cannot hold, we construct an integral over a region A_B which is enclosed by two level-sets S_1 and S_2 and two curves T_1 and T_2 connecting the level-sets

$$\int_{A_B} \text{div}(\mathbf{W}_B) d^2x = \int_{S_1} \mathcal{E} \cdot (\mathbf{b}_S \cdot \mathbf{n}_{S_1}) dS_1 + \int_{S_2} \mathcal{E} \cdot (\mathbf{b}_S \cdot \mathbf{n}_{S_2}) dS_2 \quad (4.23)$$

$$+ \int_{T_1} \mathcal{E} \cdot (\mathbf{b}_S \cdot \mathbf{n}_{T_1}) dT_1 + \int_{T_2} \mathcal{E} \cdot (\mathbf{b}_S \cdot \mathbf{n}_{T_2}) dT_2 \quad (4.24)$$

Since the vector valued function $\mathbf{b}_S(\mathbf{x})$ is the tangent vector on the level-sets S_1 and S_2 the line integrals in eq. (4.23) vanish. Furthermore T_1 and T_2 have opposite orientation so we can choose the gauge

$$(\mathbf{b}_S|_{T_1} \cdot \mathbf{n}_{T_1}) = +1, \quad (\mathbf{b}_S|_{T_2} \cdot \mathbf{n}_{T_2}) = -1 \quad (4.25)$$

Thus we have for the divergence

$$\int_{A_B} \text{div}(\mathbf{W}_B) d^2x = \int_{T_1} \mathcal{E}(\phi(\mathbf{x}), \nabla \phi(\mathbf{x})) dT_1 - \int_{T_2} \mathcal{E}(\phi(\mathbf{x}), \nabla \phi(\mathbf{x})) dT_2 \quad (4.26)$$

If $\text{div}(\mathbf{W}_B) = 0$ was to be true then the two line integrals in eq. (4.26) would have to cancel. However we did not make any assumptions on the length of the curves T_i or on the distance between them so eq. (4.26) must hold for *any* configuration of the T_i . Hence we conclude that the energy density $\mathcal{E}(\phi(\mathbf{x}), \nabla \phi(\mathbf{x}))$ must be a constant function on Ω *regardless* of the field ϕ . We have derived a clear contradiction. It follows that the energy E *cannot* be an invariant of the transformation g_t^B !

The question that remains is that if g_t^B does not change the values of $\phi(\mathbf{x})$ what does it change? Our answer is: g_t^B is an operator on the coordinates \mathbf{x} themselves

$$\mathbf{x}(t) = g_t^B \circ \mathbf{x}_0 \quad (4.27)$$

The constraint in eq. (4.17) guaranties that the level-set of $\phi(\mathbf{x}(t))$ *move along*

the flow in eq. (4.27). The image ϕ just *appears* to be transformed if we view it from an absolute reference frame Ω_0

$$\tilde{\phi}(\mathbf{x}_0) = \phi \left(g_t^B \circ \mathbf{x}_0 \right), \quad \mathbf{x}_0 \in \Omega_0 \quad (4.28)$$

However the particular reference frame Ω_0 is irrelevant as two frames Ω and Ω' may always be connected by the **flow in eq. (4.27)**.

4.2 The Generalized Newton Algorithm

Our motivated procedure in eq. (4.8) depends on the assumption that from the r equations $\text{div}(\mathbf{W}_m) = 0$ the vector \mathbf{b} can be constructed from a basis of r vectors \mathbf{b}_m which are perpendicular to the \mathbf{W}_m . To demonstrate this we remember how the energy E transforms under an arbitrary sub group $g_{\gamma(t)} \subset \mathbb{G}$ (eq. (2.110))

$$\left. \frac{d}{dt} g_{\gamma t} \circ E \right|_{t=0} = \int_{\Omega} \left(\sum_i [V_e^{\Omega}, X_e^{\Omega, i}] (\phi) \cdot P_i + v^{\phi}[\mathcal{E}_{tot}] \right) d^2x = 0 \quad (4.29)$$

$$= \int_{\Omega} \left(\sum_m \alpha_m \sum_i (\mathbf{M})_{mi} \cdot P_i + v^{\phi}[\mathcal{E}_{tot}] \right) d^2x = 0 \quad (4.30)$$

$$(\mathbf{M})_{mi} = [X_e^{\Omega, m}, X_e^{\Omega, i}] (\phi), \quad P_i \left(\mathbf{X}_e^{\Omega} \phi(\mathbf{x}) \right) = \frac{\delta \mathcal{E}_{prior} \left(\mathbf{X}_e^{\Omega} \phi(\mathbf{x}) \right)}{\delta \left(X_e^{\Omega, i} \phi \right)} \quad (4.31)$$

As we explained in section 2.4 the variation to the field ϕ , v^{ϕ} not only contains the instantaneous per-pixel variation \tilde{v}^{ϕ} but also the component $V_e^{\Omega} \phi$ which comes from the variation of the spacial coordinates \mathbf{x}

$$\phi' = \phi + v^{\phi} \quad v^{\phi}(\mathbf{x}) = \tilde{v}^{\phi}(\mathbf{x}) + V_e^{\Omega} \phi(\mathbf{x}) \quad (4.32)$$

Many methods for minimizing the energy E are of the basic Newton type in algorithm 2 which is based around the concept of steepest decent where the variation v^{ϕ} is chosen such that the field ϕ is advanced in the direction of the negative gradient of E

$$\phi' = \phi - \tau^{\phi} \cdot [\mathcal{E}_{tot}] \quad (4.33)$$

The step-size parameter τ^{ϕ} is either chosen to be constant or, in more advanced algorithms adjusted dynamically, for example on the conjugate gradients algorithm ([63]).

What is new **our** treatment of the minimization problem in eq. (4.30) is that we consider the change of the differential operators X_e^i themselves under the action of the spacial operator V_e^Ω which is encoded in the matrix commutator M . We explained in section 4.30 that if we set $v^\phi = 0$, that is we consider only the variation to the operators X_e^i in eq. (4.30) then the relation

$$M \cdot P = 0 \quad (4.34)$$

must hold independently from the coefficient vector α , since eq. (4.30) must hold for any $g_{\gamma_t} \in \mathbb{G}$. It is this relation that defines the perpendicular vector $\mathbf{b}(\mathbf{x})$. We will now show how \mathbf{b} is constructed. Since the commutator matrix M in eq. (2.107) is an element of the Lie algebra \mathcal{G} it can be represented in terms of the basis X_e^i

$$(M)_{mi} = \sum_l C_{mi}^l X_e^{\Omega,l}(\phi) \quad (4.35)$$

As the basis elements $X_e^{\Omega,l}$ are represented by the Cartesian gradient operator ∇

$$X_e^{\Omega,l} = \omega_l^\mu(\mathbf{x}) \partial_\mu \quad (4.36)$$

the product $M \cdot P$ in eq. (2.107) takes the form

$$(M \cdot P)_m = \|\mathbf{b}_m\| B_m \phi(\mathbf{x}) \quad (4.37)$$

$$B_m = \frac{1}{\|\mathbf{b}_m\|} b_m^\mu(\mathbf{x}) \partial_\mu, \quad b_m^\mu(\mathbf{x}) = \sum_i P_i(\phi(\mathbf{x})) C_{mi}^l \omega_l^\mu(\mathbf{x}) \quad (4.38)$$

We would like to discuss the operator B_m . By eq. (4.37) and eq. (4.34) the operator B_m must satisfy the level-set equation

$$B_m \phi = b_m^\mu \partial_\mu \phi = 0 \quad (4.39)$$

From eq. (4.39) we see that the vector \mathbf{b}_m is tangential to the level-sets of $\phi(\mathbf{x})$. The r operators B_m are left invariant vector fields. They are the basis of a Lie algebra \mathcal{G}^B and an associated Lie group \mathbb{G}^B . For reasons soon to be clear we call the group \mathbb{G}^B the *bending group* and the algebra \mathcal{G}^B the *bending algebra*. The dimension r_B of \mathcal{G}^B is not necessarily equal to the dimension r of the original algebra \mathcal{G} , since $B_m \neq 0$ holds only for the non trivial elements of \mathbb{G} . For instance in the case $\mathbb{G} = \mathbb{T} \times SO(2)$ only the $SO(2)$ group allows for the construction of a bending group $SO^B(2)$ since E is trivially invariant under \mathbb{T} ($M|_{\mathbb{T}} = [\partial_x, \partial_y] = 0$).

The bending algebra \mathcal{G}^B is an algebra of spacial differential operators. Thus we

can use it to define a diffusion equation for the coordinate frame Ω

$$\mathbf{x}(t) = g_t^B \circ \mathbf{x}, \quad \left. \frac{d\mathbf{x}(t)}{dt} \right|_{t=0} = \sum_{m=1}^r \beta_m B_m \mathbf{x} = \mathbf{B} \cdot \mathbf{x} \quad (4.40)$$

The operator \mathbf{B} is a linear combination of the r operators B_m from eq. (4.38) and the coefficient vector β is an arbitrary constant vector. The diffusion process in eq. (4.40) is a non-linear process since the coefficient vector $\mathbf{b}_m(\mathbf{x})$ itself (eq. (4.38)) is a function of the coordinates $\mathbf{x}(t)$. It is guided along those operators B_m which do not vanish due to trivial symmetries. To understand the dynamical properties of the diffusion equation in eq. (4.40) we calculate the rate of change of the energy E under the transformation g_t^B . First we notice that due to the level-set equation in eq. (4.39) the first order derivative of the data term \mathcal{E}^{data} under the action of g_t^B vanishes

$$\left. \frac{d}{dt} \left(g_t^B \circ \mathcal{E}^{data}(\mathbf{x}) \right) \right|_{t=0} = \frac{\delta \mathcal{E}^{data}}{\delta \phi} \cdot \mathbf{B} \phi = 0 \quad (4.41)$$

This means the diffusion process in eq. (4.40) is independent from the data term \mathcal{E}^{data} . The action of g_t^B on the prior \mathcal{E}^{prior} allows for a geometrical explanation of how the diffusion in eq. (4.40) proceeds. We let g_t^B act on \mathcal{E}^{prior} and compute the derivative with respect to t

$$\left. \frac{d}{dt} \left(g_t^B \circ \mathcal{E}^{prior} \right) \right|_{t=0} = \sum_i P_i \cdot [\mathbf{B}, X_e^i] \phi \quad (4.42)$$

The commutator $[\mathbf{B}, X_e^i] \phi$ is the change the gradient $X_e^i \phi$ undergoes under the diffusion process in eq. (4.40) at time $t = 0$. Its projection onto the canonical momentum \mathbf{P} is a measure for the curvature of the level-set defined by eq. (4.39).

The bending operator \mathbf{B} depends on the vector \mathbf{P} in contrast to the Euler-Lagrange differentials $[\mathcal{E}_{tot}]$, which only depend on the divergence $\text{div}(\mathbf{P})$. Hence although the divergence may vanish after some iterations of the classical Newton algorithm, the actual vectors \mathbf{P} and thus \mathbf{B} may be non-vanishing upon the diffusion process in eq. (4.40)

$$\mathbf{b}(\mathbf{x}(t)) \neq 0 \quad t \rightarrow \infty \quad (4.43)$$

However the prior \mathcal{E}^{prior} and thus the energy E converge to zero under the diffusion process. Thus projection of $[\mathbf{B}, X_e^i] \phi$ onto \mathbf{P} vanishes at some point $t = t^*$

$$\sum_i P_i \cdot [\mathbf{B}, X_e^i] \phi = 0, \quad t = t^* \quad (4.44)$$

To show why eq. (4.44) holds we need to argue that $g_t^B \circ \mathcal{E}^{prior}$ is convex with respect to t . For brevity we will assume $\mathcal{E}^{prior}(X_e^i \phi(\mathbf{x}))$ only depends on the derivative $X_e^i \phi(\mathbf{x})$, the other derivatives $X_e^{j \neq i} \phi$ are taken to be fixed. We define $f(\mathbf{x}, t) = g_t^B \circ X_e^i \phi(\mathbf{x})$. The function $f(\mathbf{x}, t)$ is neither globally convex in Ω with respect to t nor concave. However we can split the domain Ω into two sub-domains, one domain Ω_{\smile} in which f is globally convex and Ω_{\frown} in which f is globally concave. We may construct a globally convex function $h(\mathbf{x}, t)$

$$h(\mathbf{x}, t) = \begin{cases} f(\mathbf{x}, t) & \mathbf{x} \in \Omega_{\smile} \\ -f(\mathbf{x}, t) & \mathbf{x} \in \Omega_{\frown} \end{cases} \quad (4.45)$$

and insert $h(\mathbf{x}, t)$ into the prior \mathcal{E}^{prior} and since \mathcal{E}^{prior} is convex in its arguments the combined function $\mathcal{E}^{prior}(h(\mathbf{x}, t))$ is convex with respect to t . Now \mathcal{E}^{prior} is also a positive semi-definite functional meaning it is invariant to the reversal of the signs of its arguments, $\mathcal{E}^{prior}(h(\mathbf{x}, t)) = \mathcal{E}^{prior}(-h(\mathbf{x}, t))$. Thus we conclude that $\mathcal{E}^{prior}(f(\mathbf{x}, t))$ is also convex with respect to t . The same arguments can be applied to all derivatives X_e^i for $1 \leq i \leq q$ so the result is that $g_t^B \circ \mathcal{E}^{prior}(X_e^1 \phi, \dots, X_e^q \phi)$ is convex with respect to t .

The positive definiteness and convexity of $g_t^B \circ \mathcal{E}^{prior}$ with respect to t has the important consequence that $g_t^B \circ \mathcal{E}^{prior}$ has a global minimizer g_t^{B*} . However the minimizer g_t^{B*} is not unique! Since by construction \mathcal{E}^{prior} is invariant to the group \mathbb{G} , $\tilde{g}_t^{B*} = g_t^{B*} \circ g_\Omega$ is also a minimizer of \mathcal{E}^{prior} for all $g_\Omega \in \mathbb{G}$.

Eq. (4.42) gives us a geometrical explanation of the minimizer g_t^{B*} since for $g_t^B = g_t^{B*}$ it vanishes

$$\sum_i P_i \cdot [\mathbf{B}, X_e^i] \phi = 0 \quad (4.46)$$

As we had explained the commutator in eq. (4.46) has the geometrical meaning of being the change $X_e^i \phi$ under goes during parallel transport along \mathbf{B} . The vectors \mathbf{B} and \mathbf{P} span the tangential vectorspace in Ω since $\mathbf{B} \neq \mathbf{P}$ anywhere in Ω . The projection of $[\mathbf{B}, X_e^i] \phi$ on to \mathbf{B} vanish due to the level-set equation in eq. (4.39). If eq. (4.46) holds then the commutator must vanish in Ω

$$[\mathbf{B}, X_e^i] \phi = 0 \quad (4.47)$$

Eq. (4.47) means that the operator X_e^i does not change its orientation when parallel transported along \mathbf{B} .

However they are not elements of the original Lie algebra B_m is a left invariant vector field and thus an element of the Lie algebra \mathcal{G} . This is due to the commutator matrix \mathbf{M} being left invariant and the canonical momentum \mathbf{P} , which is the

Algorithm 2 Basic Newton Method

Set $n = 0$
Set Initial guess ϕ^0
Compute residual $r^n = -[\mathcal{E}_{tot}] (\phi^n)$
while $\|r\| > \delta$ and $n < N$ **do**
 Compute $\phi^{n+1}(\mathbf{x}) = \phi^n(\mathbf{x}) - \tau^\phi [\mathcal{E}_{tot}] (\phi^n(\mathbf{x}))$
 Recompute $r^{n+1} = -[\mathcal{E}_{tot}] (\phi^{n+1})$
 Set $n \rightarrow n + 1$
end while

Algorithm 3 Diffusion Method

Set $n = 0$
Set Initial guess ϕ^0, \mathbf{x}^0
while $n < N$ **do**
 Compute $\mathbf{x}^{n+1} = \mathbf{x}^n - \tau^\Omega \mathbf{b}(\mathbf{x}^n)$
 Compute $\phi^{n+1}(\mathbf{x}^{n+1}) = \phi^n(\mathbf{x}^{n+1})$
 Set $n \rightarrow n + 1$
end while

functional derivative of \mathcal{E}^{prior} and as such also left invariant. The consequence of the left invariance of B_m is that eq. (4.39) holds for any coordinate transformation of an arbitrarily chosen reference frame Ω_0 .

In algorithm 3 we have sketched an algorithm for the minimization of the energy E in eq. (4.1) which only deploys the diffusion process in eq. (4.40). Since the coefficient vector \mathbf{b} is a unit vector, a stopping condition in the sense of algorithm 2 ($\|r\| < \delta$) does not make sense and instead we chose terminating the algorithm after a fixed amount of steps N . However a sensible stopping condition could be derived from eq. (4.42), like $\frac{d}{dt}(g_t^B \circ E^{prior}) < \delta$. We combined both algorithms 2 and 3 into the generalized Newton algorithm 4. As we discussed above eq. (4.44) must characterize the minimum of E in eq. (4.1) just like the Euler-Lagrange equations $[\mathcal{E}] = 0$. Thus it is sufficient to implement the same stopping conditions in the algorithm 4 as in the basic Newton method in algorithm 2.

$$B_m \phi(g \circ \mathbf{x}_0) = 0 \quad \forall \mathbf{x}_0, g \quad g \in \mathbb{G}, \quad \mathbf{x}_0 \in \Omega_0 \quad (4.48)$$

Algorithm 4 Generalized Newton Method

```

Set  $n = 0$ 
Set Initial guess  $\phi^0, \mathbf{x}^0$ 
Compute residual  $r^n = -[\mathcal{E}_{tot}] (\phi^n)$ 
while  $\|r\| > \delta$  and  $n < N$  do
  Compute  $\mathbf{x}^{n+1} = \mathbf{x}^n - \tau^\Omega \mathbf{b}(\mathbf{x}^n)$ 
  Compute  $\phi^{n+1}(\mathbf{x}^{n+1}) = \phi^n(\mathbf{x}^{n+1}) - \tau^\phi [\mathcal{E}_{tot}] (\phi^n(\mathbf{x}^n))$ 
  Recompute  $r^{n+1} = -[\mathcal{E}_{tot}] (\phi^{n+1})$ 
  Set  $n \rightarrow n + 1$ 
end while

```



Figure 4.2: Figure 4.2a shows a picture ϕ^c of a person. ϕ^c is taken to be free of noise. Figure 4.2b is a noise corrupted version of ϕ^c in figure 4.2a, $\phi^d = \phi^c + n$ where n is iid Gaussian noise with a standard deviation $\sigma = 100$.

4.2.1 Image De-noising

Image de-noising is the problem of estimating a *clean* image ϕ^* given a noisy image ϕ^d . The image ϕ^0 is connected to ϕ^* via

$$\phi^d = \phi^* + n \quad n \sim \mathcal{D} \quad (4.49)$$

where \mathcal{D} is some distribution and n is a noise term drawn from \mathcal{D} . ϕ^* is estimated from the family of functionals

$$E(\phi) = \frac{1}{q} \int_{\Omega} |\phi - \phi^d|^q d^2x + \int_{\Omega} \mathcal{E}^{prior}(\phi(\mathbf{x}), \mathbf{X}_e \phi(\mathbf{x})) d^2x \quad (4.50)$$

The degree q of the data term depends on the type of noise distribution \mathcal{D} . We have simulated a noise corrupted image ϕ^d in figure 4.2b in which the distribution \mathcal{D} is the Gaussian $\mathcal{N}(0, \sigma)$ ($\sigma = 100$). We have run both the Newton and our generalized Newton algorithm to minimize the energy in eq. (4.50) for the total variation prior and our new structure tensor based prior.

TV-Prior

The TV based image de-noising model is defined by the energy

$$E(\phi, \nabla\phi) = \frac{1}{q} \int_{\Omega} |\phi - \phi^d|^q d^2x + \int_{\Omega} \mathcal{E}^{prior}(\nabla\phi(\mathbf{x})) d^2x \quad (4.51)$$

$$\mathcal{E}^{prior}(\nabla\phi(\mathbf{x})) = \lambda \sqrt{\nabla^T \phi \cdot \nabla \phi} \quad (4.52)$$

The prior \mathcal{E}^{prior} in eq. (4.52) is an invariant of the Lie group $\mathbb{G} = \mathbb{T} \times SO(2)$, the group of translation and rotations. However the associated bending operator $B_{\mathbb{T}}$ vanishes for the translation group \mathbb{T} vanishes since

$$B_{\mathbb{T}}^x = P^\nu [\partial_\nu, \partial_x] = 0, \quad B_{\mathbb{T}}^y = P^\nu [\partial_\nu, \partial_y] = 0 \quad (4.53)$$

that is \mathbb{T} is a trivial symmetry of \mathcal{E}^{prior} and $E(\phi, \nabla\phi)$. The bending operator B_θ associated with the rotation group $SO(2)$ does not vanish, but computes to

$$B_\theta = b_\theta^\mu \partial_\mu, \quad \mathbf{b}_\theta = \frac{\nabla^\perp \phi}{\sqrt{\nabla^T \phi \cdot \nabla \phi}} = \mathbf{P}^\perp \quad (4.54)$$

We ran algorithm 3 and algorithm 4 with the bending operator in eq. (4.54) and compared it to the original Newton method in algorithm 2. In figure 4.3a we have plotted the energy $E(\phi, \nabla\phi)$ over the current iteration n for both algorithm 4 and 2 and in figure 4.3b the prior energy E^{prior} . The pure diffusion algorithm (PDA) 3 initially minimizes the prior E^{prior} and the total energy E faster than the basic Newton algorithm (BNA). At around $n = 40$ iterations the PDA starts to converge to a higher total energy E than the BNA. We relate this finding to the fact that the bending operator B_θ does not depend on the data term \mathcal{E}^{data} and that the PDA does not involve the steepest descent step in eq. (4.6). The prior energy E^{prior} however converges for both the PDA and the BNA to the same value albeit the PDA converging slower than the BNA. The GNA as a combination of both the PDA and the BNA converges approximately twice as fast as the BNA alone according to figures 4.3a and 4.3b.

In figure 4.4a the variance $\sigma^2 = \mathbb{V}(\phi)$ and in figure 4.4b the curvature $\|\mathbf{K}\|$ are plotted against the iterations n for the BNA, PDA and GNA. The curvature $\|\mathbf{K}\|$

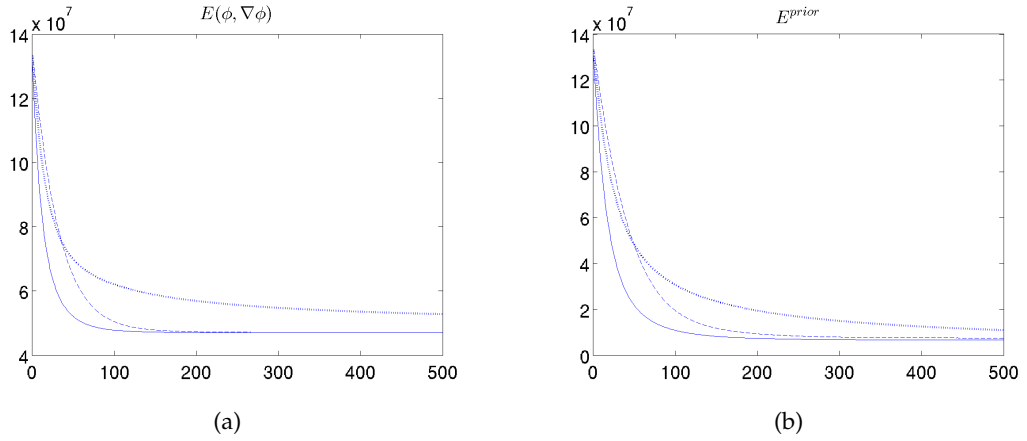


Figure 4.3: Figure 4.3a is a plot of the energy in eq. (4.52) and figure 4.3b a plot of the prior energy E^{prior} over the current iteration n for algorithm 4 (solid line), algorithm 2 (dashed line) and algorithm 3 (dotted line). We can see that the diffusion process in eq. (4.40) alone (algorithm 3) minimizes the prior energy E^{prior} and the total energy E . We see that the bending operator in the diffusion process in eq. (4.40) causes an acceleration in the decrease of the prior energy E^{prior} and thus of the total energy E .

is the mean curvature over the domain Ω . Similar as for the energies E and E^{prior} the mean curvature is initially minimized by the PDA faster than the BNA and after $n = 40$ iterations slower than the BNA. It converges fastest for the GNA. The variance σ^2 has a similar behavior like the total energy E in that since the PDA is not bounded by the data term \mathcal{E}^{data} the variance σ^2 converges to a higher level than the BNA and the GNA. Again since the GNA creates regions of equal intensity, level-set regions in ϕ via the diffusion in eq. (4.40) and simultaneously matching the level-set regions to each other via the data term in the Euler-Lagrange differential in eq. (4.33) the variance converges the fastest for the GNA.

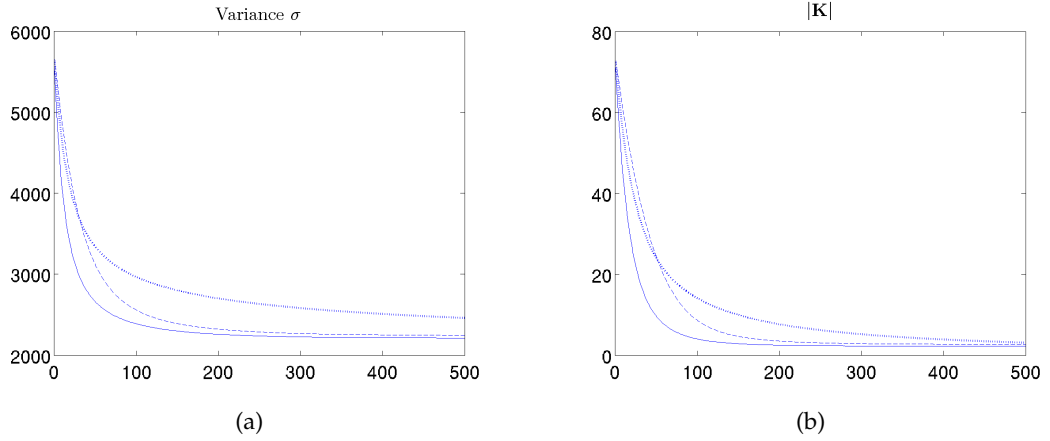


Figure 4.4: Figure 4.4a shows the variance σ^2 of the image ϕ and figure 4.4b the curvature vector \mathbf{K} for algorithm 3 (dotted line), algorithm 2 (dashed line) and algorithm 4 (solid). The curvature $\|\mathbf{K}\|$ converges for all three algorithms 3, 2 and 4 to the same value. However the variance σ^2 for algorithms 4 and 2 converge to a lower value than algorithm 3. The explanation is that while algorithm 3 only creates regions of equal intensity, level-set regions, the algorithms 4 and 2 not only create level-set regions but actually match the intensities of the level-set regions to each other.

Structure Tensor Prior

In this section we applied our structure tensor prior from section ?? to the image de-noising problem

$$E(\phi, \nabla\phi) = \frac{1}{q} \int_{\Omega} |\phi - \phi^d|^q d^2x + \int_{\Omega} \mathcal{E}^{prior}(\nabla\phi(\mathbf{x})) d^2x \quad (4.55)$$

$$\mathcal{E}^{prior}(\nabla\phi(\mathbf{x})) = \det(S) \quad (4.56)$$

In order to apply the GNA in algorithm 4 to the model in eq. (4.56) we need to compute the coefficient vector \mathbf{b} of the bending operator \mathbf{B} . From section ?? we know that \mathcal{E}_{ST}^{prior} is invariant to the group $\mathbb{G} = \mathbb{T} \times SO(2)$. Like the TV prior the translation group \mathbb{T} is a trivial symmetry so that it suffices to compute the bending operator \mathbf{B}_{θ} corresponding to the group $SO(2)$. We remember that the structure tensor prior in eq. (4.56) transforms under the $SO(2)$ in the following

way

$$\left. \frac{d}{dt} E^{prior} (S_{gt}) \right|_{t=0} = \int_{\Omega} \text{Tr} (S \cdot M_{\theta} \cdot S) d^2x = 0 \quad (4.57)$$

where M_{θ} is the Pauli matrix (the generator of the algebra $\mathfrak{so}(2)$). We can expand the structure tensor S in terms of its eigenvectors $\mathbf{v}_{1,2}$ and the corresponding eigenvalues $\lambda_{1,2}$. The eigenvector \mathbf{v}_1 is parallel to the *mean gradient* (over the window covered by S) of ϕ

$$\mathbf{v}_1 \parallel \langle \nabla \phi \rangle \quad (4.58)$$

and the eigenvalue λ_1 is the squared magnitude of $\langle \nabla \phi \rangle$. The vector \mathbf{v}_2 is orthogonal to \mathbf{v}_1 and λ_2 is the squared rate of change of ϕ along the direction of \mathbf{v}_2 . The structure tensor S can be written as

$$S = \lambda_1 \mathbf{v}_1 \mathbf{v}_1^T + \lambda_2 \mathbf{v}_2 \mathbf{v}_2^T \quad (4.59)$$

The trace in eq. (4.57) then reduces to the equation

$$\text{Tr} (S \cdot M_{\theta} \cdot S) = \langle \nabla \phi \rangle^T \tilde{\mathbf{b}}_{\theta}, \quad \tilde{\mathbf{b}}_{\theta} = S \cdot \left(\langle \nabla^{\perp} \phi \rangle - \frac{\lambda_2}{\sqrt{\lambda_1}} \mathbf{v}_2 \right) \quad (4.60)$$

The vector $\tilde{\mathbf{b}}_{\theta}$ shows in the direction of \mathbf{v}_2 and is thus perpendicular to the mean gradient $\nabla \phi$ and the trace in eq. (4.57) vanishes. We insert $\tilde{\mathbf{b}}_{\theta}$ into the definition of the bending operator B in eq. (4.38) and get

$$B_{\theta} = v_2^{\mu} \partial_{\mu} \quad (4.61)$$

In figures 4.5 and 4.6 the energies E^{prior} , E the variance σ^2 and the $\|\mathbf{K}\|$ are plotted over the iterations n for both algorithms 2 and 4. Like for the TV prior we see that the energies E and E^{prior} converge for the GNA faster than for the BNA. The variance σ^2 (figure 4.6) however converges for both algorithms at the same rate albeit to a lower value for the GNA.

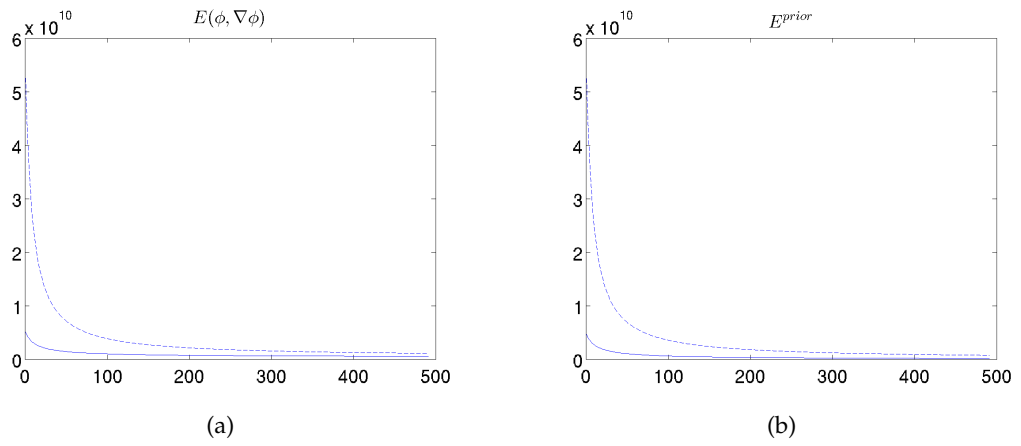


Figure 4.5: Figure 4.5a is a plot of the energy in eq. (4.52) and figure 4.5b a plot of the prior energy E^{prior} over the current iteration n for algorithm 4 (solid line), algorithm 2 (dashed line) and algorithm 3 (dotted line). We can see that the diffusion process in eq. (4.40) alone (algorithm 3) minimizes the prior energy E^{prior} and the total energy E . We see that the bending operator in the diffusion process in eq. (4.40) causes an acceleration in the decrease of the prior energy E^{prior} and thus of the total energy E .

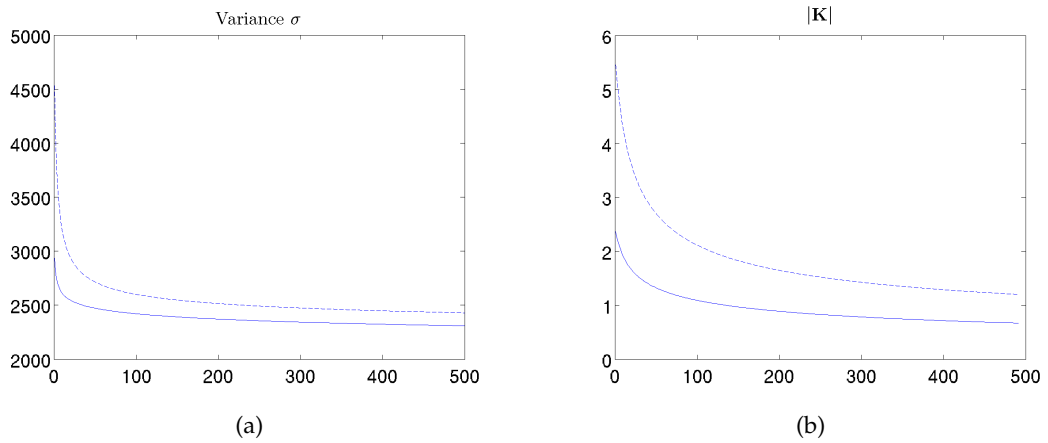


Figure 4.6: Figure 4.6a shows the variance σ^2 of the image ϕ and figure 4.6b the curvature vector \mathbf{K} for algorithm 3 (dotted line), algorithm 2 (dashed line) and algorithm 4 (solid). The curvature $\|\mathbf{K}\|$ converges for all three algorithms 3, 2 and 4 to the same value. However the variance σ^2 for algorithms 4 and 2 converge to a lower value than algorithm 3. The explanation is that while algorithm 3 only creates regions of equal intensity, level-set regions, the algorithms 4 and 2 not only create level-set regions but actually match the intensities of the level-set regions to each other.

5 Conclusions

Bibliography

- [1] D. Sun, S. Roth, and M. J. Black, "Secrets of optical flow estimation and their principles," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 2432–2439. [7](#), [33](#)
- [2] Paul Fieguth, *Statistical image processing and multidimensional modeling*, Springer Science & Business Media, 2010. [8](#)
- [3] Leonid I Rudin, Stanley Osher, and Emad Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992. [12](#), [29](#)
- [4] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical imaging and vision*, vol. 20, no. 1-2, pp. 89–97, 2004. [12](#), [29](#)
- [5] Mark Nitzberg and Takahiro Shiota, "Nonlinear image filtering with edge and corner enhancement," *IEEE transactions on pattern analysis and machine intelligence*, vol. 14, no. 8, pp. 826–833, 1992. [12](#)
- [6] J. Weickert, "Theoretical foundations of anisotropic diffusion in image processing," *Computing, Suppl.*, vol. 11, pp. 221–236, 1996. [12](#)
- [7] Joachim Weickert, *Anisotropic diffusion in image processing*, vol. 1, Teubner Stuttgart, 1998. [12](#)
- [8] D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated variational problems," *Comm. Pure and Applied Mathematics*, vol. 42, pp. 577–685, 1989. [12](#), [31](#)
- [9] L. Ambrosio and V. M. Tortorelli, "Approximation of functional depending on jumps by elliptic functional via t-convergence," *Communications on Pure and Applied Mathematics*, vol. XLIII, pp. 999–1036, 1990. [12](#)
- [10] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the mumford and shah model," *International Journal of Computer Vision*, vol. 50 no 3, pp. 271–293, 2002. [12](#), [31](#)

- [11] Hans-Hellmut Nagel and Wilfried Enkelmann, "An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, , no. 5, pp. 565–593, 1986. [13](#)
- [12] Alexander A Kirillov, *An introduction to Lie groups and Lie algebras*, vol. 113, Cambridge University Press Cambridge, 2008. [16](#), [19](#)
- [13] Elizabeth Louise Mansfield, *A practical guide to the invariant calculus*, vol. 26, Cambridge University Press, 2010. [16](#), [17](#), [27](#)
- [14] Mario Ferraro and Terry M. Caelli, "Relationship between integral transform invariances and lie group theory," *Journal of the Optical Society of America A*, vol. 5 no 5, pp. 738–742, 1988. [16](#)
- [15] Taichiro Kugo, *Eichtheorie*, 1997. [22](#)
- [16] Michael E. Peskin and Daniel V. Schroeder, *An Introduction to Quantum Field Theory (Frontiers in Physics)*, Perseus Books, 1995. [22](#)
- [17] Richard Phillips Feynman, Robert B. Leighton, and Matthew Sands, *The Feynman Lectures on Physics*, 1963. [22](#)
- [18] E. Noether, "Invariante variationsprobleme," *Nachrichten Der Koeniglichen Gesellschaft Der Wissenschaften Zu Goettingen*, vol. Mathematische-Physikalische Klasse, pp. 235–257, 1918. [23](#)
- [19] E. Noether, "Invariant variation problems," *Transport Theory and Statistical Problem*, vol. 1 (3), pp. 183–207, 1971. [23](#), [27](#)
- [20] A. Wedel, T. Pock, C. Zach, H. Bischof, and D. Cremers, "An improved algorithm for tv-l 1 optical flow," in *Statistical and Geometrical Approaches to Visual Motion Analysis*, pp. 23–45. Springer, 2009. [29](#), [33](#)
- [21] J. M. Fadili and G. Peyre, "Total variation projection with first order schemes," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 657–669, Mar. 2011. [29](#)
- [22] Kristian Bredies and Dirk Lorenz, "Mathematische bildverarbeitung," *Vieweg+ Teubner*, vol. 4, no. 6, pp. 12, 2011. [29](#), [30](#), [32](#)
- [23] Stanley Osher and Leonid I Rudin, "Feature-oriented image enhancement using shock filters," *SIAM Journal on Numerical Analysis*, vol. 27, no. 4, pp. 919–940, 1990. [29](#)
- [24] Leonid I Rudin, "Images, numerical analysis of singularities and shock filters," 1987. [29](#)

- [25] Otmar Scherzer, Markus Grasmair, Harald Grossauer, Markus Haltmeier, and Frank Lenzen, "Variational methods in imaging, volume 167 of applied mathematical sciences," 2009. [29](#), [32](#)
- [26] LC Evans and J Spruck, "Motion of level sets by mean curvature. i," in *Fundamental contributions to the continuum theory of evolving phase interfaces in solids*, pp. 328–374. Springer, 1999. [32](#)
- [27] Vincent Caselles, Ron Kimmel, and Guillermo Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22 no 1, pp. 61–79. [32](#)
- [28] Satyanad Kichenassamy, Arun Kumar, Peter Olver, Allen Tannenbaum, and Anthony Jr. Yezzi, "Conformal curvature flows: From phase transitions to active vision," *Archive for rational mechanics and analysis*, vol. 134 issue 3, pp. 275–301, 1996. [32](#)
- [29] B. K.P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, issues 1-3, pp. 185–203, 1981. [33](#), [34](#), [36](#)
- [30] B. D. Lucas, T. Kanade, et al., "An iterative image registration technique with an application to stereo vision.," in *IJCAI*, 1981, vol. 81, pp. 674–679. [33](#), [34](#), [36](#)
- [31] Michael J Black and Paul Anandan, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields," *Computer vision and image understanding*, vol. 63, no. 1, pp. 75–104, 1996. [33](#)
- [32] T. Bruhn, J. Weickert, and C. Schnörr, "Lucas/kanade meets horn/schunck: Combining local and global optic flow methods," *International Journal of Computer Vision*, vol. 61,no 3, pp. 211–231, 2004. [33](#)
- [33] Claudia Kondermann, Daniel Kondermann, and Christoph S. Garbe, "Post-processing of optical flows via surface measures and motion inpainting," *Lecture Notes in Computer Science*, vol. 5096, pp. 355–364, 2008. [33](#), [34](#)
- [34] Soo-Chul Han and Christine I Podilchuk, "Video compression with dense motion fields," *IEEE Transactions on Image Processing*, vol. 10, no. 11, pp. 1605–1612, 2001. [33](#)
- [35] Soo-Chul Han and Christine I Podilchuk, "Efficient encoding of dense motion fields for motion-compensated video compression," in *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*. IEEE, 1999, vol. 1, pp. 84–88. [33](#)
- [36] M. Pollefeys and L. Koch, R.and Gool, "Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters," *International Journal of Computer Vision*, vol. 32, no. 1, pp. 7–25, 1999. [33](#)

- [37] J. Repko and M. Pollefeys, "3d models from extended uncalibrated video sequences: Addressing key-frame selection and projective drift," in *3-D Digital Imaging and Modeling, 2005. 3DIM 2005. Fifth International Conference on*. IEEE, 2005, pp. 150–157. [33](#)
- [38] P. Tanskanen, K. Kolev, L. Meier, F. Camposeco, O. Saurer, and M. Pollefeys, "Live metric 3d reconstruction on mobile phones," in *Computer Vision (ICCV), 2013 IEEE International Conference on*. IEEE, 2013, pp. 65–72. [33](#)
- [39] N. Snavely, S. M. Seitz, and R. Szeliski, "Modeling the world from internet photo collections," *International Journal of Computer Vision*, vol. 80, no. 2, pp. 189–210, 2008. [33](#)
- [40] N. Snavely, S. M. Seitz, and R. Szeliski, "Photo tourism: exploring photo collections in 3d," in *ACM transactions on graphics (TOG)*. ACM, 2006, vol. 25, pp. 835–846. [33](#)
- [41] Young Min Kim, Christian Theobalt, James Diebel, Jana Kosecka, Branislav Miscusik, and Sebastian Thrun, "Multi-view image and tof sensor fusion for dense 3d reconstruction," in *Proc. IEEE 12th Int Computer Vision Workshops (ICCV Workshops) Conf*, 2009, pp. 1542–1549. [33](#)
- [42] Daniel Scharstein and Richard Szeliski, "High-accuracy stereo depth maps using structured light," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, 2003, vol. 1. [33](#)
- [43] Mahua Bhattacharya and Arpita Das, "Registration of multimodality medical imaging of brain using particle swarm optimization," in *Proceedings of the first international conference on intelligent human computer interaction*. Springer, 2009, pp. 131–139. [33](#)
- [44] Anthoni Yezzi, Lilla Zöllei, and Tina Kapur, "A variational framework for joint segmentation and registration," *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, p. 44, 2001. [33](#)
- [45] Gerardo Hermosillo Valadez, *Variational Methods for Multimodal Image Matching*, Ph.D. thesis, École Doctorale Sciences et Technologies de l'Information et de la Communication, 2002. [34](#), [38](#), [39](#), [40](#)
- [46] Christophe Chef d'Hotel, Gerardo Hermosillo, and Olivier Faugeras, "A variational approach to multi-modal image matching," in *Variational and Level Set Methods in Computer Vision, 2001. Proceedings. IEEE Workshop on*. IEEE, 2001, pp. 21–28. [34](#), [39](#)
- [47] Christopher M Bishop, *Neural networks for pattern recognition*, Oxford university press, 1995. [34](#)

- [48] Christoph Zetsche and Erhardt Barth, "Fundamental limits of linear filters in the visual processing of two-dimensional signals," *Vision research*, vol. 30, no. 7, pp. 1111–1117, 1990. [34](#)
- [49] Nils Papenberg, Andrés Bruhn, Thomas Brox, Stephan Didas, and Joachim Weickert, "Highly accurate optic flow computation with theoretically justified warping," *International Journal of Computer Vision*, vol. 67, no. 2, pp. 141–158, 2006. [36](#)
- [50] Christopher Zach, Thomas Pock, and Horst Bischof, "A duality based approach for realtime tv-l 1 optical flow," in *Joint Pattern Recognition Symposium*. Springer, 2007, pp. 214–223. [36](#)
- [51] Frederik Maes, Andre Collignon, Dirk Vandermeulen, Guy Marchal, and Paul Suetens, "Multimodality image registration by maximization of mutual information," *IEEE transactions on Medical Imaging*, vol. 16, no. 2, pp. 187–198, 1997. [38](#), [40](#)
- [52] Alexis Roche, Gregoire Malandain, and Nicholas Ayache, *Unifying maximum likelihood approaches in medical image registration*, Ph.D. thesis, Inria, 1999. [38](#), [39](#)
- [53] Heiko Hirschmüller, "Stereo processing by semiglobal matching and mutual information," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 2, pp. 328–341, 2008. [38](#)
- [54] Alexis Roche, Grégoire Malandain, Nicholas Ayache, and Xavier Pennec, "Multimodal image registration by maximization of the correlation ratio," 1998. [39](#)
- [55] A. Roche, G. Malandain, X. Pennec, and N. Ayache, "The correlation ratio as a new similarity measure for multimodal image registration," in *Medical Image Computing and Computer-Assisted Intervention*, pp. 1115–1124. Springer, 1998. [39](#)
- [56] Olivier Faugeras and Renaud Keriven, *Variational principles, surface evolution, pde's, level set methods and the stereo problem*, IEEE, 2002. [40](#)
- [57] Thomas Netsch, Peter Rosch, Arianne van Muiswinkel, and Jürgen Weese, "Towards real-time multi-modality 3-d medical image registration," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*. IEEE, 2001, vol. 1, pp. 718–725. [40](#)
- [58] Pascal Cachier and Xavier Pennec, "3d non-rigid registration by gradient descent on a gaussian-windowed similarity measure using convolutions," in *Mathematical Methods in Biomedical Image Analysis, 2000. Proceedings. IEEE Workshop on*. IEEE, 2000, pp. 182–189. [40](#)

- [59] Russell C Hardie, Kenneth J Barnard, John G Bognar, Ernest E Armstrong, and Edward A Watson, "High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system," *Optical Engineering*, vol. 37, no. 1, pp. 247–260, 1998. [42](#), [43](#)
- [60] R. C. Hardie, M. T. Eismann, and G. L. Wilson, "Map estimation for hyperspectral image resolution enhancement using an auxiliary sensor," *IEEE Transactions on Image Processing*, vol. 13, no. 4, 2004. [41](#), [42](#), [43](#), [44](#), [48](#)
- [61] Y. Zhang, "Spatial resolution enhancement for hyperspectral image based on wavelet bayesian fusion," *4th International Congress on Image and Signal Processing*, 2011. [44](#)
- [62] J. Bigun and G. H. Granlund, "Optimal orientation detection of linear symmetry," *Proceedings of the IEEE First International Conference on Computer Vision : London, Great Britain*, pp. 433–438, 1987. [45](#)
- [63] Jonathan Richard Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain, 1994," URL <http://www-2.cs.cmu.edu/jrs/jrspapers.html#cg>, 1994. [76](#)