Contents

1 Introduction

Etwas umständlich formuliert

Many objects in nature posses 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. A ball of uniform color for instance does not change its appearance to an observer upon rotation around an arbitrary axis 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 $\phi_{\mathcal{O}}(\boldsymbol{x}) = const$ (e.g. $\phi_{\mathcal{O}}(\boldsymbol{x}) = x^2 + y^2 + z^2 = 1$ for a ball of unit radius) then our intuition of global symmetry is equal to the statement that $\phi_{\mathcal{O}}(\boldsymbol{x}) = const$ is invariant under the global transformation $\boldsymbol{x}' = g \circ \boldsymbol{x}$

through the center

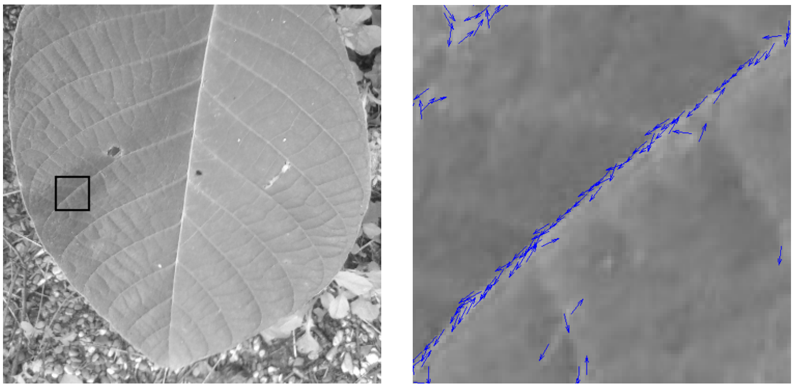
$$\phi_{\mathcal{O}}(g \circ \boldsymbol{x}) = \phi_{\mathcal{O}}(\boldsymbol{x}) \tag{1.1}$$

Sind doch die wenigsten, oder?

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 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 posses 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 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 $\boldsymbol{\omega}^{\Omega}(\boldsymbol{x})$ such that the local transformation $g^{\Omega}(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{\omega}^{\Omega}(\boldsymbol{x})$ leaves the image ϕ invariant

$$\phi(\boldsymbol{x} + \boldsymbol{\omega}^{\Omega}(\boldsymbol{x})) = \phi(\boldsymbol{x}) \tag{1.2}$$

In general we cannot assume that $g^{\Omega}(\boldsymbol{x})$ in eq. (1.2) is unique since there can always exist a vector field $\boldsymbol{\omega}^{\Omega\prime}(\boldsymbol{x}) \neq \boldsymbol{\omega}^{\Omega}(\boldsymbol{x})$ which satisfies eq. (1.2). On the other side any transformation g^{Ω} satisfying eq. (1.2) uniquely determines the



(a) (b)

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.

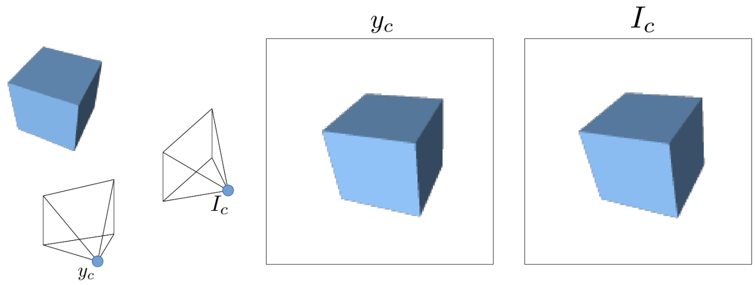
geometry of ϕ for if we were to draw lines along the tangential vectors $\boldsymbol{\omega}^{\Omega}(\boldsymbol{x})$ by connecting $\boldsymbol{x}$ with $\boldsymbol{x} + \boldsymbol{\omega}^{\Omega}(\boldsymbol{x})$ we would reconstruct the object $\mathcal{O}$ from $g^{\Omega}(\boldsymbol{x})$.

Auch etwas umständlich, nicht klar, worauf Du hinaus willst.

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.

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



(a) (b) (c)

Figure 1.2: Figure 1.2a: Two cameras are shown recording a scene from different positions. The scene could ~~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.

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

$$Y, \phi : \Omega \to \mathbb{R}^n \tag{1.3}$$

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

$$Y(\boldsymbol{x}) \xrightarrow{T_Y} \phi(\boldsymbol{x}) \tag{1.4}$$

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'

$$\phi'(\boldsymbol{x}) = \phi(g \circ \boldsymbol{x}) \tag{1.5}$$

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

$$Y \xrightarrow{T_Y} \phi = Y \xrightarrow{T_Y^1} \phi^1 \xrightarrow{T_Y^2} \phi^2 \cdots \xrightarrow{T_Y^{n-1}} \phi^{n-1} \xrightarrow{T_Y^n} \phi \tag{1.6}$$

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

$$g' \circ T_Y^i = T_Y^{i+k} \tag{1.7}$$

we may minimize the number steps in eq. (1.6) und 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

$$\phi^\star = \text{argmin}_\phi E_Y(\phi) \quad \leftrightarrow \quad Y \xrightarrow{T_Y} \phi^\star \tag{1.8}$$

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

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

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 $\boldsymbol{W}_m$

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

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 if 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 [17]. 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 $\boldsymbol{x}$ in Ω (see eq. (1.2)) into account to facilitate the search for the shortest path in the inference problem in eq. (1.6).

Generell sehr schön und verständlich geschrieben. Einziger Schwachpunkt: Es ist sehr lange unklar, worauf Du hinaus willst. Was ist eigentlich das Thema?. Vielleicht noch ein paar einleitende Sätze, um klarer den Zusammenhang zu geben.

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 $\phi = (\phi_1, \ldots, \phi_n)$. The physical system C relates the vector ϕ to a set of observables $Y = \{Y_1, \ldots, Y_k\}$

$$Y = C\left(\phi^{\star}\right) \tag{2.1}$$

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

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

called a Gibbs-Random-Field (GRF) [1]. 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 $\phi(\boldsymbol{x})$ 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

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

which is related to the probability distribution

$$p\left(\phi|Y\right) = p\left(Y|\phi\right) \cdot p(\phi) \sim \exp\left(-E_Y\left(\phi\right)\right) \tag{2.4}$$

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

$$p\left(\phi\right) = \exp\left(-E^{prior}\left(\phi\right)\right) \tag{2.6}$$

The value of the probability distribution $p(\phi|Y)$ at $\phi(\boldsymbol{x}) = \hat{\phi(\boldsymbol{x})}$ describes the probability that the GRF $\phi(\boldsymbol{x})$ assumes the values $\hat{\phi}(\boldsymbol{x})$ at each point $\boldsymbol{x} \in \Omega$. The set of values $\hat{\phi}(\boldsymbol{x})$ is what is called a *configuration* of the GRF ϕ

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

$$\phi^\star = \operatorname{argmin}_\phi \left(E_Y(\phi)\right) \tag{2.7}$$

The particular value $\phi^\star(\boldsymbol{x})$ 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

$$\phi^\star = C^{-1}(Y) \tag{2.8}$$

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 $I_O(\boldsymbol{x})$ is the light intensity the object O reflects to the point $\boldsymbol{x}$ 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 $\boldsymbol{x}_{i,j}$ called pixels

$$I^c_{ij} \in \mathbb{R}, \quad \boldsymbol{x}_{i,j} \in \Omega \quad 1 < i < n, 1 < j < m \tag{2.9}$$

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

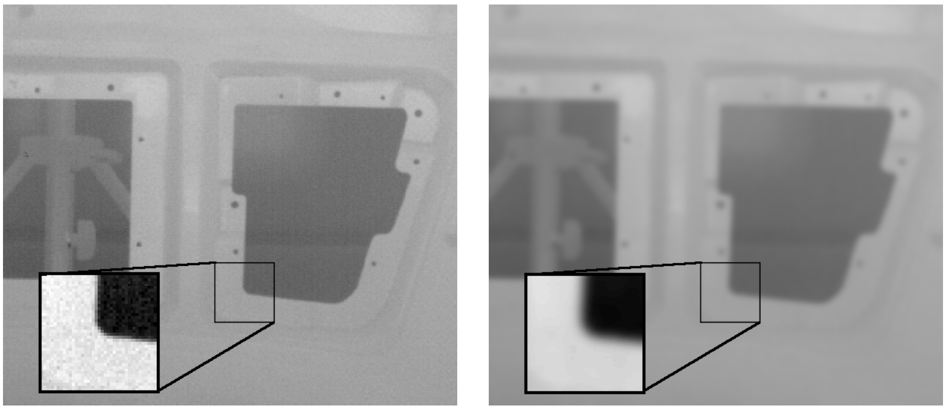
$$I^c_{ij} = C_{ij}(I_O) \tag{2.10}$$

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

$$\hat{I}^c_{ij} = I_O(\boldsymbol{x}_{i,j}) + n \quad n \sim p\left(I^c_{ij}|I_O(\boldsymbol{x}_{i,j})\right) \tag{2.11}$$

$p(I^c_{ij}|I_O(\boldsymbol{x}_{i,j}))$ is the likelihood that I^c_{ij} assumes the value $\hat{I}^c_{ij}$ given the incoming intensity $I_O(\boldsymbol{x}_{i,j})$ at the pixel $\boldsymbol{x}_{i,j}$. 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 $I_O(\boldsymbol{x}_{i,j})$ at the pixels $\boldsymbol{x}_{i,j}$ from the noisy data I^c_{ij} we need to pose some form of regularity on the values $I_O(\boldsymbol{x})$ to counter the pixel-wise noise imposed by the CCD in eq. (2.11). This can be achieved by



(a) (b)

Figure 2.1: Figure 2.1a shows an image 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 $I_O^\star$ 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 $I_O(\boldsymbol{x})$ at all pixels with each other in the prior

$$p\left(I_O\right) = \exp\left(-E^{prior}\left(I_O\right)\right) \tag{2.12}$$

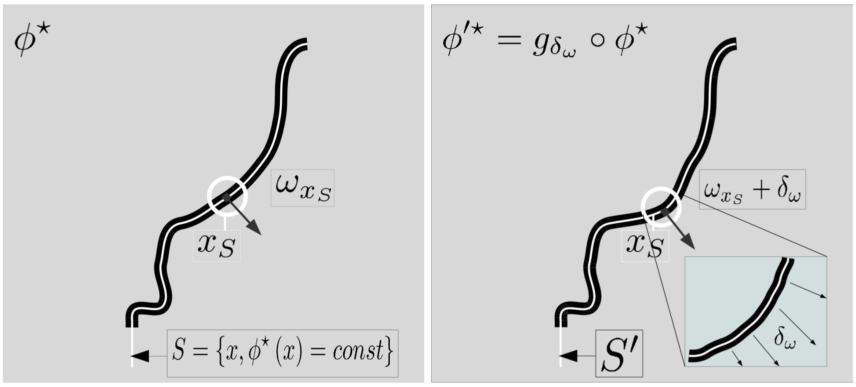
$$E^{prior}\left(I_O\right) = \int_\Omega \mathcal{E}\left(I_O\left(\boldsymbol{x}\right), I_O\left(\Omega / \left\{\boldsymbol{x}\right\}\right)\right) dx \tag{2.13}$$

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

$$I_O^\star = \text{argmin}_{I_O}\left(E_{I^c}\left(I_O\right)\right), \quad E_{I^c}\left(I_O\right) = E_{I^c}^{data}(I_O) + E^{prior}\left(\nabla I_O\right) \tag{2.14}$$

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 $n^2 - 1$ pixels. Even for medium sized images with $n = 500$ the computations involved in eq. (2.14) are practically infeasible.

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



(a) (b)

Figure 2.2: Local transformation of an image ϕ with a level-set S. Figure 2.2a shows an image $\phi(\boldsymbol{x})$ with a line S along which the intensity values are constant. At each point $\boldsymbol{x}_S$ the vector $\boldsymbol{\omega}_S$ 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 $\boldsymbol{\omega}_S$ a spacial dependent vector $\boldsymbol{\delta}_\omega(\boldsymbol{x})$

gradient $\nabla I_O(\boldsymbol{x})$

$$E^{prior}_{L_2}(\nabla I_O) = \int_\Omega \|\nabla\ I_O(\boldsymbol{x})\|^2\, dx \tag{2.15}$$

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). $E^{prior}_{L_2}$ 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

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

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

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

$$A_c = \left\{\phi_c^\star \middle| \phi_c^\star = \text{argmin}_\phi(E_{L_2}(\nabla\phi)) = c, \quad c \in \mathbb{R}\right\} \tag{2.17}$$

A slightly different description of the set of minimizers A_c goes as follows: given $\phi_0^\star(\boldsymbol{x}) = c_0$ we can generate all other possible minimizers $\phi_c^\star$ by adding any $c \in \mathbb{R}$ to $\phi_0^\star(\boldsymbol{x})$. We can label the action of adding a real number $c \in \mathbb{R}$ on to any function $\phi(\boldsymbol{x})$ by g_c

$$A_c = \{\phi_c^\star(\boldsymbol{x}) | \phi_c^\star = g_c \circ \phi_0^\star = \phi_0^\star + c, \quad c \in \mathbb{R}\} \tag{2.18}$$

Under the group of such actions $\mathbb{G}_{const} = \{g_c\}$ our set of minimizers A_c is invariant

$$g_d \circ A_c = A_{c+d} = \{\phi_c^\star(\boldsymbol{x}) | \phi_{c+d}^\star = \phi_0^\star + c + d, \quad c \in \mathbb{R}\} = A_c \tag{2.19}$$

as well as the L_2 energy

$$g_d \circ E_{L_2}(\phi) = \int_\Omega \|\nabla(\phi + d)\|^2 = \int_\Omega \|\nabla\phi\|^2 \tag{2.20}$$

In this sense we can state that the L_2 prior $p_{L_2}(\nabla\phi)$ is actually conditioned on the group of constant transformations

$$p_{L_2}(\nabla\phi) = p_{L_2}(\nabla\phi|\mathbb{G}_{const}) \tag{2.21}$$

since it is invariant under the entire set $\mathbb{G}_{const}$ but under no other set. This is why we call p_{L_2} *conditionally invariant* with respect to $\mathbb{G}_{const}$. We observe that $\mathbb{G}_{const}$ is not a discrete set but a continuous set since the parameters c and d in eq. (2.18) and eq. (2.19) are real valued numbers. In eq. (2.18) the set $\mathbb{G}_{const}$ acts on the functions $\phi^\star(\boldsymbol{x})$ by shifting their *function values* by constants.

Now consider the set of transformations $\mathbb{G}_\Omega$ whose elements $g_{\omega\Omega} \in \mathbb{G}_\Omega$ operate on the coordinate space Ω by warping it with the vector-field $\boldsymbol{\omega}^{\Omega}(\boldsymbol{x})$

$$g_{\omega^\Omega} \circ \boldsymbol{x} = \boldsymbol{x} + \boldsymbol{\omega}^{\Omega}(\boldsymbol{x}) \tag{2.22}$$

$\mathbb{G}_\Omega$ is the set of all possible deformations of the space Ω. Obviously any element $\phi^\star(\boldsymbol{x}) \in A_c$ is invariant under the action of $\mathbb{G}_\Omega$ since A_c is the set of constant functions. Thus the prior $p_{L_2}(\nabla\phi)$ is conditionally invariant under the combined set $\mathbb{G}_{\Omega c} = \mathbb{G}_\Omega \times \mathbb{G}_{const}$

$$p_{L_2}(\nabla\phi) = p_{L_2}(\nabla\phi|\mathbb{G}_{\Omega c}) \tag{2.23}$$

In the following we will argument that it is possible to introduce priors $p(\nabla\phi)$ which allow for conditional invariance with respect to a larger set of transformations $\mathbb{G} = \mathbb{G}_\Omega \times \mathbb{G}_i$

$$p(\nabla\phi) = p(\nabla\phi|\mathbb{G}) \tag{2.24}$$

where the elements $g_{\omega^\phi} \in \mathbb{G}_i$ operate in a similar fashion like the $g_{\omega\Omega}$ in eq. (2.22) but on non constant functions $\phi(\boldsymbol{x})$

$$\phi\tilde{(\boldsymbol{x})} = g_{\omega^\phi}\phi(\boldsymbol{x}) = \phi(\boldsymbol{x}) + \omega^{\phi}(\boldsymbol{x}) \tag{2.25}$$

Similar to the definition of A_c in eq. (2.18) we can describe the maximizers of $p(\nabla\phi)$ as being related to each other by the elements of $\mathbb{G}$

$$A = \{\phi^\star \,|\phi^\star = g \circ \phi_0^\star \quad g \in \mathbb{G}\} \tag{2.26}$$

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

$$A_\Omega(\phi_c^\star) = \{\phi^\star \,|\phi^\star(\boldsymbol{x}) = \phi_c^\star(g_{\omega^\Omega} \circ \boldsymbol{x}), \quad g_{\omega^\Omega} \in \mathbb{G}_\Omega\} \tag{2.27}$$

$$A = \{A_\Omega(\phi_c^\star)\,|\phi_c^\star = g_{\omega^\phi} \circ \phi_0^\star, \quad g_{\omega^\phi} \in \mathbb{G}_i\} \tag{2.28}$$

This is significant for the following reason: knowledge of the geometric set of transformations $\mathbb{G}_\Omega$ 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 $\phi_c^\star \in A_{red}$ are not related to each other by $\mathbb{G}_\Omega$

$$A_{red} = \{\phi_c^\star \,|\phi_c^\star = g_{\omega^\phi} \circ \phi_0^\star, \quad g_{\omega^\phi} \in \mathbb{G}_i\} \tag{2.29}$$

$$\phi_d^\star(\boldsymbol{x}) \neq \phi_c^\star(g_{\omega^\Omega}\boldsymbol{x}) \quad \forall g_{\omega^\Omega} \in \mathbb{G}_\Omega, \phi_{c,d}^\star \in A_{red} \tag{2.30}$$

We may also turn the argument around: we could specify the geometric set of transformations $\mathbb{G}_\Omega$ and design a prior $p(\nabla\phi)$ which is conditionally invariant under $\mathbb{G}_\Omega$, 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 $\phi_0^\star$ is a sub set $S_c \subset \Omega$ defined by

$$S_c = \{\boldsymbol{x}\,|\phi_0^\star(\boldsymbol{x}) = c\} \tag{2.31}$$

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

$$g \circ \phi(\boldsymbol{x}) = g_{\omega^\phi}\phi(g_{\omega^\Omega} \circ \boldsymbol{x}) \tag{2.32}$$

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_{ω^Ω} in eq. (2.22) we see that g_{ω^Ω} is a geometrical transformation that deforms the level-sets S_c (see figure 2.2). We are free to define g_{ω^ϕ} so that it is orthogonal to g_{ω^Ω} in the sense that the level-sets S_c are invariant under g_{ω^ϕ}

$$S_{c'} = g_{\omega^\phi} \circ S_c = S_c \tag{2.33}$$

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 $\boldsymbol{\omega}_\delta(\boldsymbol{x})$ which (see figure 2.2) is the set of vectors *tangent* to S_c

$$S_c = \{\boldsymbol{x}\,|\boldsymbol{\omega}_\delta(\boldsymbol{x}) \cdot \nabla\phi_0^\star(\boldsymbol{x}) = 0\} \tag{2.34}$$

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 $\boldsymbol{\omega}'_\delta(\boldsymbol{x}) = \boldsymbol{\omega}_\delta(\boldsymbol{x}) + \boldsymbol{\delta}(\boldsymbol{x})$ as tangent vectors.

$$S'_c = \{\boldsymbol{x}\,|(\boldsymbol{\omega}_\delta(\boldsymbol{x}) + \boldsymbol{\delta}(\boldsymbol{x})) \cdot \nabla\phi_0^\star(\boldsymbol{x}) = 0\} \tag{2.35}$$

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

$$S'_c = \{\boldsymbol{x}'\,|\boldsymbol{\omega}_\delta(\boldsymbol{x}') \cdot \nabla_\delta\phi_0^\star(\boldsymbol{x}') = 0\} \tag{2.36}$$

The operator ∇_δ loosely speaking encodes a reversal of the action of g_{ω^Ω} on $\boldsymbol{x}$ so that S'_c can be represented with the same tangential vector-field as S_c but in the new frame $\boldsymbol{x}' = g_{\omega_\delta} \circ \boldsymbol{x}$. The operator ∇_δ is called a *pull-back* of the gradient ∇. With the help of the pull-backs ∇_δ it is possible to translate the notion of conditional invariance with respect to $\mathbb{G}_\Omega$ to the requirement that $p(\nabla_\delta\phi)$ must be constant with respect to variations of the vector-field $\boldsymbol{\delta}(\boldsymbol{x})$

$$\frac{\delta}{\delta \boldsymbol{\delta}(\boldsymbol{x})} p(\nabla_\delta \phi) = 0 \tag{2.37}$$

Given a specific form of the operators in $\mathbb{G}_\Omega$, eq. (2.37) poses constraints on the form of the differential operators in the prior $p(\nabla_\delta \phi)$. Eq. (2.37) also ensures that $p(\nabla_\delta \phi)$ is indifferent to a large class of level-sets $\{S\}$, which are generated by $\mathbb{G}_\Omega$ acting on S (see eq. (2.36)).

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* [2, 3, 4] if the group operation

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

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

$$l_g : \mathbb{G} \to \mathbb{G} \quad l_g(h) = g \cdot h \quad g, h \in \mathbb{G} \tag{2.38}$$

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 $\mathcal{A}$

$$g : \mathcal{A} \to \mathbb{G}, \quad (a_1, \ldots, a_r) \to g_{a_1,\ldots,a_r} \tag{2.39}$$

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 $\cdot$ is the matrix multiplication. In section 2.2.1 we argument that the set $\mathbb{G}$ acts in a two-fold manner on the functions $\phi(\boldsymbol{x}) \in \mathcal{C}^\infty(\Omega)$, namely by acting on the spacial coordinates $\boldsymbol{x} \in \Omega$ in eq. (2.22) and on the function values $\phi(\boldsymbol{x})$ them selves in eq. (2.25). 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 $\phi(\boldsymbol{x})$ 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* $J^k\left(\mathcal{C}^\infty\left(\Omega\right)\right)$, $\mathcal{M} = \Omega \times J^k\left(\mathcal{C}^\infty\left(\Omega\right)\right)$. $J^k\left(\mathcal{C}^\infty\left(\Omega\right)\right)$ is the set of smooth differentiable functions with compact support in Ω and their derivatives up to order k. The points $\boldsymbol{z} \in \mathcal{M}$ are vectors of the *independent* variables $\boldsymbol{x}$, the *dependent* variable $\phi\left(\boldsymbol{x}\right)$ and its derivatives $\phi_{,K}$

$$\boldsymbol{z} = \left(\boldsymbol{x}, \phi\left(\boldsymbol{x}\right), \phi_{,K}\left(\boldsymbol{x}\right)\right) \tag{2.40}$$

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

$$\boldsymbol{z} = \left(\boldsymbol{x}, \phi\left(\boldsymbol{x}\right), \nabla\phi\left(\boldsymbol{x}\right)\right) \tag{2.41}$$

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

$$\tilde{\boldsymbol{z}} = \left(\tilde{\boldsymbol{x}}, \tilde{\phi}\left(\tilde{\boldsymbol{x}}\right), \tilde{\nabla}\tilde{\phi}\left(\tilde{\boldsymbol{x}}\right)\right) \tag{2.42}$$

$$\tilde{\boldsymbol{x}} = g_{a_1 \ldots a_r} \circ \boldsymbol{x} \tag{2.43}$$

$$\tilde{\phi} = g_{a_1 \ldots a_r} \circ \phi \tag{2.44}$$

$$\tilde{\nabla} = J^{-1}\nabla, \quad J_{\mu\nu} = \frac{d\tilde{x_\mu}}{dx_\nu} \tag{2.45}$$

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

$$\gamma : t \to \left(a_1\left(t\right) \ldots a_r\left(t\right)\right) \tag{2.46}$$

$$g_{\gamma(0)} = e \tag{2.47}$$

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 $e \in \mathbb{G}$, $T_e\mathbb{G}$

$$\left.\frac{d}{dt} g_{\gamma(t)}\right|_{t=0} = X_e \in T_e\mathbb{G} \tag{2.48}$$

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 $\mathcal{M}$, $\mathcal{F}\left(\mathcal{M}\right)$. The action of X on $\mathcal{F}\left(\mathcal{M}\right)$ can be computed by evaluating $F \in \mathcal{F}\left(\mathcal{M}\right)$ on the tranformed vector $\tilde{\boldsymbol{z}} = g_{\gamma(t)} \circ \boldsymbol{z}$ and the taking the derivative with respect to t at the neutral element e

$$X_e F\left(\boldsymbol{z}\right) = \left.\frac{d}{dt} F\left(\tilde{\boldsymbol{z}}\right)\right|_{t=0} = \sum_{i=1}^{r} (\omega^i_\mu \partial_\mu F\left(\boldsymbol{z}\right) + \omega^\phi_i \frac{d}{d\phi} F\left(\boldsymbol{z}\right) + D\phi^\nu_i \frac{d}{d\partial_\nu \phi} F\left(\boldsymbol{z}\right))\alpha_i \tag{2.49}$$

where we have

$$\omega_\mu^i(\boldsymbol{x}) = \left.\frac{d\tilde{x_\mu}}{da_i}\right|_{t=0} \qquad \omega_i^\phi(\boldsymbol{x},\phi) = \left.\frac{d\tilde{\phi}}{da_i}\right|_{t=0} \qquad \alpha_i = \left.\frac{da_i}{dt}\right|_{t=0} \tag{2.50}$$

$$D\phi_i^\nu = \frac{d\omega_i^\phi}{dx_\nu} - \sum_\mu \frac{d\omega_\nu^i}{dx_\mu}\partial_\mu\phi \tag{2.51}$$

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

$$X_{e,i} = X_e^{\Omega,i} + \omega_i^\phi\frac{d}{d\phi} + D\phi_i^\nu\frac{d}{d\partial_\nu\phi}, \quad X_e^{\Omega,i} = \omega_\mu^i\partial_\mu \tag{2.52}$$

so that X_e has the operator form

$$X_e = \sum_i \alpha_i X_{e,i} \tag{2.53}$$

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 $l_{h^\star}$ called the *push-forward*

$$Y_hF(\boldsymbol{z}) = (l_{h^\star}X_e)F(\boldsymbol{z}) = \left.\frac{d}{dt}F\left(l_h\left(g_{\gamma(t)}\right)\circ\boldsymbol{z}\right)\right|_{t=0} \tag{2.54}$$

The vector X_e operates on the function F in eq. (2.49) as a differential operator at the point $e \circ \boldsymbol{z} = \boldsymbol{z}$. The effect of $l_{h^\star}$ is that it transports the vector X_e to the vector Y_h which operates on F at the point $l_h(e)\circ\boldsymbol{z} = h\circ\boldsymbol{z}$. 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 $\mathbb{G}$

$$T\mathbb{G} = \bigcup_{h\in\mathbb{G}} T_h\mathbb{G} \tag{2.55}$$

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 $\boldsymbol{z} \in \mathcal{M}$. Similar to X_e in eq. (2.53) the vector Y_h has a coordinate representation with respect to the tangential space $T_h\mathbb{G}$

$$Y_hF(\boldsymbol{z}) = \sum_i \alpha_i' Y_{h,i} \tag{2.56}$$

$$Y_{h,i} = \omega_\mu'^i\partial_\mu + \omega_i'^\phi\frac{d}{d\phi} + D'\phi_i^\nu\frac{d}{d\partial_\nu\phi} \tag{2.57}$$

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 $g \in \mathbb{G}$

$$l_{g^\star} X_h = X_{g \cdot h} = \sum_i \alpha_i X^i_{g \cdot h} \quad \forall g \in \mathbb{G}, X_h \in \mathcal{G} \tag{2.58}$$

From eq. (2.58) we see that a consequence of left invariance is that the coordinate vector $\boldsymbol{\alpha}$ is constant under the transformation l_g. This is what is referred to as the *parallel transport* of $\boldsymbol{\alpha}$ along the transformation l_g. The Lie algebra $\mathcal{G}$ has the property that it is closed under the antisymmetric commutator $[\cdot, \cdot]$

$$[X_h, Y_h] = Z_h \in \mathcal{G} \quad \forall X_h, Y_h \in \mathcal{G} \tag{2.59}$$

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

$$A = \left\{ \phi^\star \middle| \phi^\star = \text{argmin}_\phi \left(-E^{prior} \left(\nabla\phi \right) \right) \right\} \tag{2.60}$$

to be non trivial, that is $\phi^\star \neq const$. Conditional invariance was linked to the requirement that the minimizer set A in eq. (2.60) be generated by the group $\mathbb{G}$

$$A = \{ \phi^\star | \phi^\star = g_\omega \circ \phi_0^\star \quad g_\omega \in \mathbb{G} \} \tag{2.61}$$

In eq. (2.28) 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_{ω^Ω} on the coordinate frame Ω. We motivated the introduction of deformations to the gradient operator ∇ such that the level-sets S' in eq. (2.36) have the same functional form in the transformed coordinates $\boldsymbol{x}' = g_{\omega^\Omega} \circ \boldsymbol{x}$ as in the original coordinates (see eq. (2.34)). 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 $e \in \mathbb{G}$

$$S_X = \left\{ \boldsymbol{x} \middle| X_e^\Omega \phi \left(\boldsymbol{x} \right) = 0 \right\} \tag{2.62}$$

The operator X_e^Ω is the spacial components of the vector X_e (see eq. (2.52)). Under the action of $g_{\omega_\delta} \in \mathbb{G}$ the level-set S_X transforms the following way

$$S'_X = g_{\omega_\delta} \circ S_X = \left\{ \boldsymbol{x} \middle| X^\Omega_{g_{\omega_\delta}} \phi \left(\boldsymbol{x} \right) = 0 \right\} \tag{2.63}$$

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. (2.63). 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 $p\left(\nabla\phi\right)$. Conditional invariance is the expressed by the equation

$$p\left(X_{g_{\gamma(t)}}^{\Omega,1}\phi^\star,\ldots,X_{g_{\gamma(t)}}^{\Omega,r}\phi^\star\right) = const \tag{2.64}$$

Since the basis X_e^i are the left invariant vector fields X^i evaluated at the identity e the eq. (2.64) would hold for a push-forwarded basis $X_{hg_{\gamma(t)}}^i$ for any $h \in \mathbb{G}$ since $h \circ \phi^\star \in A$ if if $\phi^\star \in A$. This is why eq. (2.64) only needs to hold for small regions $U_e \subset \mathbb{G}$ around the identity

$$\frac{d}{dt}p\left(X_{g_{\gamma(t)}}^{\Omega,1}\phi^\star,\ldots,X_{g_{\gamma(t)}}^{\Omega,r}\phi^\star\right)\Bigg|_{t=0} = V_e p\left(X_e^{\Omega,1}\phi^\star,\ldots,X_e^{\Omega,r}\phi^\star\right) = 0 \tag{2.65}$$

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

$$V_e = \sum_{i=1}^{r}\alpha_i \cdot X_e^i \tag{2.66}$$

The eq. (2.65) 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. (2.48)). However for eq. (2.65) to hold for all one parameter sub groups and thus for all $g \in \mathbb{G}$, it has to hold for all coefficient vectors $\alpha = \left(\alpha_1,\ldots,\alpha_r\right)^T$.

2.4.1 Noethers Theorems

In her original paper [5, 6] Emmy Noether handles the question: Given a model of a physical system, encoded in an action

$$E = \int_\Omega \left(\mathcal{E}\left(\boldsymbol{x},\{\phi_\rho\},\{\nabla_K\phi_\rho\}\right)\right)d^n x \tag{2.67}$$

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

$$E' = g_\gamma \circ E = E \quad \forall g_\gamma \in \mathbb{G} \tag{2.68}$$

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 will introduce 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, $\gamma = (\alpha_1, \ldots, \alpha_r)$.

- Infinite dimensional Lie groups $\mathbb{G}_\infty$, which are generalizations of the finite dimensional groups in the sense that the r parameters $\alpha_1, \ldots, \alpha_r$ 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_γ to be the smooth *infinitesimal* transformation, encoding both variations of the fields and of the coordinates

$$\phi'_\rho\left(\boldsymbol{x}\right) = \phi_\rho\left(\boldsymbol{x}\right) + \sum_{m=1}^{r} \alpha_m \omega_m^{\phi_\rho}(\boldsymbol{x}) \quad \boldsymbol{x}' = \boldsymbol{x} + \sum_{m=1}^{r} \alpha_m \boldsymbol{\omega}_m^{\Omega}(\boldsymbol{x}) \tag{2.69}$$

The functions $\omega_m^{\phi_\rho}$ and $\boldsymbol{\omega}_m^{\Omega}$ can be seen as a basis for ω^ϕ and $\boldsymbol{\omega}^\Omega$ in eqs. 2.22 and 2.25 She proved if the action S is invariant under g_ω eq. (2.68), then there exists r vectors $\boldsymbol{W}_m$ such the integral relationship

$$E - E' = \int_\Omega \sum_{m=1}^{r} \alpha_m \left[\sum_\rho \bar{\omega}_m^{\phi_\rho} \left[\mathcal{E}\right]_\rho + \operatorname{div}\left(\boldsymbol{W}_m\right)\right] = 0 \tag{2.70}$$

$$\bar{\omega}_m^{\phi_\rho} = \left(\omega_m^{\phi_\rho} - \omega_m^{\mu\Omega} \partial_\mu \phi_\rho\right) \tag{2.71}$$

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

$$\int f \cdot \partial_\mu g d^n x = \int \partial_\mu \left(f \cdot g\right) d^n x - \int \partial_\mu f \cdot g d^n x \tag{2.72}$$

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

$$\sum_\rho \bar{\omega}_m^{\phi_\rho} \left[\mathcal{E}\right]_\rho + \operatorname{div}\left(\boldsymbol{W}_m\right) = 0 \quad m = 1, \ldots, r \tag{2.73}$$

relate r of the ρ Euler-Lagrange equations $\left[\mathcal{E}\right]_\rho$ so that the physical system only has $\rho - r$ independent Euler-Lagrange equations $\left[\mathcal{E}\right]_\rho$ and thus only $\rho - r$ independent fields ϕ_ρ. In the case $\rho \leq r$ the system of equation in eq. (**??**) is overdetermined, eq. (2.70) can only hold if *all* the divergences and *all* the Euler-Lagrange equations vanish

$$[\mathcal{E}]_\rho \left(\phi_1^\star, \ldots, \phi_\rho^\star\right) = 0, \quad \mathrm{div}\left(\boldsymbol{W}_m\right)\left(\phi_1^\star, \ldots, \phi_\rho^\star\right) = 0 \tag{2.74}$$

Eq. (2.74) implies that only at the minima of the fields, $\phi_\rho^\star$ the r vectors $\boldsymbol{W}_m$ are conserved and eq. (2.68) holds.

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 derivative ∇ using the Lie algebra introduced in section 2.3. We consider the negative log-prior energy

$$E^{prior} = -\ln p\left(X_e^1\phi, \ldots, X_e^r\phi\right) = \int\limits_\Omega \mathcal{E}\left(x, \left\{X_e^i\phi\right\}\right) d^2x \tag{2.75}$$

We apply a one parameter group $g_{\gamma(t)}$ to E^{prior} and according to eq. (2.48) we can compute the vector V_e in the tangent space of $g_{\gamma(t)}$ at $t = 0$

$$\frac{d}{dt} g_{\gamma_t} \circ E^{prior}\bigg|_{t=0} = V_e E^{prior} \tag{2.76}$$

To derive the invariants $\mathbf{W}_i$ there are two components of the prior E^{prior} to consider under the action in eq. (2.76). The first is the functional $\mathcal{E}$. The energy density $\mathcal{E}$ is an element of the jet space $\mathcal{M}$, $\mathcal{E} \in \mathcal{M}$. As such the action of a one parameter group on $\mathcal{E}$ is, similar to eq. (2.49)

$$\frac{d}{dt} g_{\gamma_t} \circ \mathcal{E}\bigg|_{t=0} = V_e \mathcal{E} \tag{2.77}$$

The second component is the volume element d^2x which also changes under the transformation in eq. (2.75)

$$\frac{d}{dt} d^2\tilde{x}|_{t=0} = \frac{dv^\mu}{dx^\mu} d^2x \tag{2.78}$$

where the v^μ are the coefficient functions of V_e in Ω. Eq. (2.78) describes the deformation of volumes in Ω due to the stretching or compressing of the Cartesian coordinates. Taking both eq. (2.78) and eq. (2.77) into account eq. (2.76) proves to be (see Appendix)

$$\frac{d}{dt} g_{\gamma_t} \circ E^{prior}\bigg|_{t=0} = \int_\Omega \left(\sum_m \alpha_m \mathrm{div}\left(\boldsymbol{W}_m\right) + \tilde{v}^\phi[\mathcal{E}] \right) d^2x \tag{2.79}$$

$$\tilde{v}^\phi = v^\phi - V_e^\Omega(\phi) \tag{2.80}$$

where $[\mathcal{E}]$ are the Euler-Lagrange differentials

$$[\mathcal{E}] = \frac{\delta \mathcal{E}}{\delta \phi} - \sum_i \frac{d}{dx^\mu} \left(w_i^\mu \frac{\delta \mathcal{E}}{\delta \left(X_e^i \phi \right)} \right) \tag{2.81}$$

and the r vectors $\boldsymbol{W}_m$ have the form

$$W_m^\mu = \omega_m^\mu \mathcal{E} + \sum_i \omega_i^\mu \left(\omega_m^\phi - X_e^m(\phi) \right) \frac{\delta \mathcal{E}}{\delta \left(X_e^i \phi \right)} \tag{2.82}$$

Eq. (2.79) is the most general form of variation. It contains two components, namely one component proportional to intensity variations of the field ϕ and one component proportional to variations of the coordinate frame Ω. The transformations we are going to analyze in this thesis preserve the volume

$$\frac{dv^\nu}{dx^\nu} = 0 \tag{2.83}$$

Taking eq. (2.83) into account the divergences of the vectors $\boldsymbol{W}_m$ are easily computed

$$\mathrm{div}\left(\boldsymbol{W}_m\right) = X_e^{\Omega,m}(\phi)\,[\mathcal{E}] - \sum_i \left[X_e^{\Omega,i}, X_e^{\Omega,m} \right](\phi) \cdot \frac{\delta \mathcal{E}}{\delta \left(X_e^{\Omega,i} \phi \right)} \tag{2.84}$$

The intensity variations are encoded in the factor $\tilde{v}^\phi(\boldsymbol{x})$ in eq. (2.80). Since $v^\phi = \sum_i \alpha_i \omega_i^\phi$ and from eq. (2.50) we see that the ω_i^ϕ describe the *total* change of ϕ in the direction α_i of V_e, $\tilde{v}^\phi(\boldsymbol{x})$ encodes instantaneous intensity fluctuations of $\phi(\boldsymbol{x})$ at the point $\boldsymbol{x}$, which *do not* emerge from spatial transformations. Since the prior E^{prior} only contains one field ϕ the equation

$$\frac{d}{dt} g_{\gamma_t} \circ E^{prior}\bigg|_{t=0} = 0 \tag{2.85}$$

can only hold (by the argumentation in section 2.4.1) if the divergences of the vectors $\boldsymbol{W}_m$ in eq. (2.82) *and* the Euler-Lagrange differentials vanish

$$[\mathcal{E}]\left(\phi^\star\right) = 0, \quad \mathrm{div}\boldsymbol{W}_m = 0 \quad \forall 1 \leq m \leq r \tag{2.86}$$

which implies

$$\sum_i \left[X_e^{\Omega,i}, X_e^{\Omega,m}\right](\phi^\star) \cdot \frac{\delta \mathcal{E}}{\delta\left(X_e^{\Omega,i}\phi\right)} = 0 \tag{2.87}$$

There are three cases to consider such that eq. (2.87) 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: $\frac{\delta \mathcal{E}}{\delta(X_e^{\Omega,i}\phi)} = 0$ 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 $\frac{\delta \mathcal{E}}{\delta(X_e^{\Omega,i}\phi)}$ if non-vanishing must be orthogonal to the vector $\mathbf{M}_m$, which is a vector for fixed m defined as $(\mathbf{M}_m)_i = \left[X_e^{\Omega,i}, X_e^{\Omega,m}\right](\phi)$ over Ω

Pure Spacial Symmetries

The case of a pure spacial symmetry ($v^\phi = 0$) of the prior E^{prior}

$$V_e^\Omega E^{prior} = 0 \tag{2.88}$$

is a stronger constraint on E^{prior} then eq. (2.85). From eqs. 2.79 and 2.84 we see that in this case the Euler-Lagrange equations $[\mathcal{E}]$ cancel out

$$V_e^\Omega E^{prior} = -\int_\Omega \sum_{m=1}^{r} \alpha_m \left[\sum_i \left[X_e^{\Omega,i}, X_e^{\Omega,m}\right](\phi) \cdot \frac{\delta \mathcal{E}}{\delta\left(X_e^{\Omega,i}\phi\right)}\right] = 0 \tag{2.89}$$

It follows that if eq. (2.88) holds for any one parameter sub group $g_{\gamma(t)} \subset \mathbb{G}_\Omega$ of a more general group of spacial transformations $\mathbb{G}_\Omega$ then

$$\sum_i \left[X_e^{\Omega,i}, X_e^{\Omega,m}\right](\phi) \cdot \frac{\delta \mathcal{E}}{\delta\left(X_e^{\Omega,i}\phi\right)} = 0 \tag{2.90}$$

must hold for *any* field configuration ϕ. 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. (2.90) still holds for any field ϕ.

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 $\mathbb{R}^3$

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

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

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

The Euler-Lagrange equations are easily computed

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

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 $SO(3) \times \mathbb{R}$)

$$t' = t + \delta t \tag{2.94}$$

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

where θ_i are rotation around the x-,y- or z-axis. From Noether's theorem there exist four corresponding conserved quantities:

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

$$l_x = z\dot{y} - y\dot{z} \quad \textit{Rotation around x-axis} \tag{2.97}$$

$$l_y = z\dot{x} - x\dot{z} \quad \textit{Rotation around y-axis} \tag{2.98}$$

$$l_z = x\dot{y} - y\dot{x} \quad \textit{Rotation around z-axis} \tag{2.99}$$

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 $\mathbf{l} = (l_x, l_y, l_z)$ (Eqs. eq. (2.97) to eq. (2.99)) is called the *angular momentum* of the masses m_e and m_s as they rotate around each other The solutions to the Euler-Lagrange equations in eq. (2.93) are elliptic curves in the plane orthogonal to $\mathbf{l}$. The constancy of $\mathbf{l}$ with respect to the special orthogonal group $SO\left(3\right)$ comes the fact the plane is embedded in the euclidean coordinate space with unit metric, rather some general Riemann space.

2.5 Total Variation

The earliest attempts to optimization in computer vision all had in common, the use of isotropic priors for the regularization of the unknowns to be estimated. For example one of the earliest attempts for image de-noising involves minimizing the functional ([7])

$$E\left(u\right)=\int\left(u-u^{0}\right)^{2}dx+\frac{\lambda}{2}\int|\nabla u|^{2}dx \tag{2.100}$$

The first term in Eq. eq. (2.100) is the likelihood which states the minimizer $u^{\star}$ must be close in its intensity distribution to the given data u^{0}. The second term, the prior energy imposes smoothness on the minimizer $u^{\star}$. Both terms are quadratic in u and thus the Euler-Lagrange equations for $E\left(u\right)$ are linear in u making them computationally easy to solve. The problem with the prior $\frac{\lambda}{2}\int|\nabla u|^{2}dx$ is that it does not allow the solutions $u^{\star}$ to have discontinuities. Different approaches for anisotropic priors exist, for instance [8] introduced a quadratic prior

$$E_{p}rior=\int\left(\nabla u\right)^{T}D\left(\nabla u\right) \tag{2.101}$$

The operator D is a local 2×2 symmetric valued matrix with eigenvectors tangential to the level-sets of u^{0}. This is why D steers the direction of the gradients in Eq. eq. (2.101) in tangential direction of the level-sets, and thus also of the discontinuities of u and u^{0}. The upside is that the prior in Eq. eq. (2.101) combined with the likelihood in Eq. eq. (2.116) still lead to Euler-Lagrange equations linear in u. The downside of the prior in Eq. eq. (2.101) is that the operator field D must be precomputed on the data u^{0}, e.g with an eigenvalue analysis of the structure tensor.

In the context of shock-filtering ([9, 10, 11]) it was shown that the functional

$$E_{L_1}\left(u\right)=\int|\nabla u|dx \tag{2.102}$$

has the appealing property that it does not penalize large discontinuities However its functional derivative with respect to u is ill conditioned in the case $\nabla u \approx 0$. To alleviate the case, [9] chose the approximative prior

$$E_{L_1 approx}(u) = \int \sqrt{|\nabla u|^2 + \epsilon} dx \tag{2.103}$$

which is well behaved for $\epsilon > 0$. They were able to achieve good results with relatively sharp preserved discontinuities with data u^0 having low SNRs. Never the less in the limit $\epsilon \to 0$ 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 ([12]). To do this we need to explore the function-space the minimizers of Eq. eq. (2.102) might belong to. Smooth functions u_{smooth} are functions for which ∇u exists everywhere, thus they may be minimizers of Eq. eq. (2.102). But functions $u_{discont}$ containing discontinuities do *not* have finite L_1 norm of their gradients, $E_{L_1}(u_{discont}) = \infty$ since the gradient $\nabla u_{discont}$ does not exist at the discontinuities A generalization of Eq. eq. (2.102) is possible if one assumes ∇u 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 $\phi \in \mathcal{C}_0(\Omega)$ with compact domain, the following equality holds

$$\int_\Omega u \cdot \text{Div}\phi dx = -\int \phi d\mu < \infty \tag{2.104}$$

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

$$BV = \{u \in L_1(\Omega) | \nabla u \in \mathcal{M}(\Omega)\} \tag{2.105}$$

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. (2.104)

$$\int_\Omega u \cdot \text{Div}\phi dx \leq C\|\phi\|_\infty \tag{2.106}$$

The scalar C is the norm of the radon measure ∇u and is called the total variation of u

$$TV(u) = \sup\left\{\int_\Omega u \cdot \text{Div}\phi dx \,|\quad \|\phi\|_\infty \leq 1\right\} \tag{2.107}$$

As was discussed in [12] the functions u are geometrically piecewise smooth, meaning there exists a partitioning $\{\Omega_k\}$ of Ω such that $(\nabla u)_{\Omega_k}$ are L_1 integrable. If dl_{mk} is a line segment in the intersection $\Omega_m \cap \Omega_k$ then $TV(u)$ can be written in the form

$$TV\left(u\right)=\sum_{k}\left\|\nabla u_{\Omega_{k}}\right\|_{L_{1}}+\sum_{k<m}\int_{\Omega_{k}\cap\Omega_{m}}\left|u_{k}-u_{m}\right|dl_{km} \tag{2.108}$$

where u_k the value of u on the portion of $\partial\Omega_k$ which is interfacing with Ω_m and vice versa for u_m. The first term in eq. (2.108) penalizes the smooth parts of u (the gradients $(\nabla u)_{\Omega_k}$) while the second term penalizes the length of the section $\Omega_m \cap \Omega_k$ while maintaining the values $u_{k,m}$ and thus the *jump* $|u_k - u_m|$. 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.

2.5.1 The Mean Curvature of Total Variation

In this section we will discuss the geometrical properties of the TV norm in eq. (2.107). The sub-gradient of eq. (2.107) is equal to the set

$$\partial TV\left(u\right)=\left\{-\mathrm{Div}\sigma\,\middle|\quad\sigma\cdot\nu=0\text{ on }\partial\Omega,\ \sigma=\frac{\nabla\ u}{\left|\nabla u\right|}\text{ if }\left|\nabla u\right|\neq0\right\} \tag{2.109}$$

This set defines the set of lines $L\left(v\right)=TV\left(u\right)+\left\langle\mathrm{Div}\sigma|v-u\right\rangle$ tangential to TV at a point $u\in BV$. We define a one parameter Lie group $\gamma\left(t\right)$, such that its vector-field X fulfills the condition

$$X\cdot u\left(\Gamma^{X}\left(\boldsymbol{x}_{0},t\right)\right)=0 \tag{2.110}$$

then its integral curves $\Gamma^{X}\left(t\right)=\left(x\left(t\right),y\left(t\right)\right)$ are the level sets of u. The level sets Γ^X have a curvature κ and the standard formula for κ is

$$\kappa=\frac{1}{\left\|\dot{\Gamma}^{X}\right\|_{L_{1}}^{3}}\left(\dot{x}\cdot\ddot{y}-\dot{y}\cdot\ddot{x}\right) \tag{2.111}$$

If the vector field X is expressed by the coordinate vector $\xi\left(\boldsymbol{x}_0\right)$ then it can be shown κ is a function of the Laplacian relative to the coordinate vector $\xi\left(\boldsymbol{x}_0\right)$.

$$\kappa\left(\mathbf{x}_{0}\right)=\frac{\Delta_{\xi\xi}u\left(\mathbf{x}_{0}\right)}{\left|\nabla u\left(\mathbf{x}_{0}\right)\right|} \tag{2.112}$$

This form can easily be transformed into a divergence quantity

$$\kappa=\mathrm{Div}\left(\frac{\nabla\ u}{\left|\nabla u\right|}\right) \tag{2.113}$$

This shows us that the sub-gradient in Eq.: eq. (2.109) is equal to the curvature of the level-sets $\Gamma^X(t)$

$$\kappa = -\partial TV(u) \tag{2.114}$$

The eq. eq. (2.114) exposes the capital geometrical property of the TV norm: The TV norm penalizes the curvature of the level-sets of an image. 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.

2.5.2 Image De-noising

Image de-noising is the problem of estimating a *clean* image $u^\star$ given a noisy image u^0. The image u^0 is connected to $u^\star$ via

$$u^0 = u^\star + n \quad n \sim \mathcal{D} \tag{2.115}$$

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

$$F(u) = \frac{1}{q}\int_\Omega |u - u^0|^q dx + \lambda TV(u) \tag{2.116}$$

The degree q of the data term must be matched to the form of the distribution $\mathcal{D}$. Using the sub-gradient in eq eq. (2.109) the Euler-Lagrange equations can be calculated

$$[F](u) = \begin{cases} |u^\star - u^0|^{q-2}\left(u^\star - u^0\right) - \lambda \mathrm{Div}\phi & in\Omega \\ \phi \cdot \nu = 0 & on\partial\Omega \\ \phi = \frac{\nabla u}{|\nabla u|} & |\nabla u| - \text{nearly everywhere} \end{cases} \tag{2.117}$$

As was discussed before the term Divϕ is equal to the curvature κ on the level-set of u. Thus the parameter λ controls how strong the curvature of the level-sets are penalized. Never the less the functional eq. (2.116) still permits jumps in the image $u^\star$.

2.6 Optical Flow

A prime example of an inverse problem in computer vision is optical flow. Optical Flow labels the task of densely measuring the motion between two or

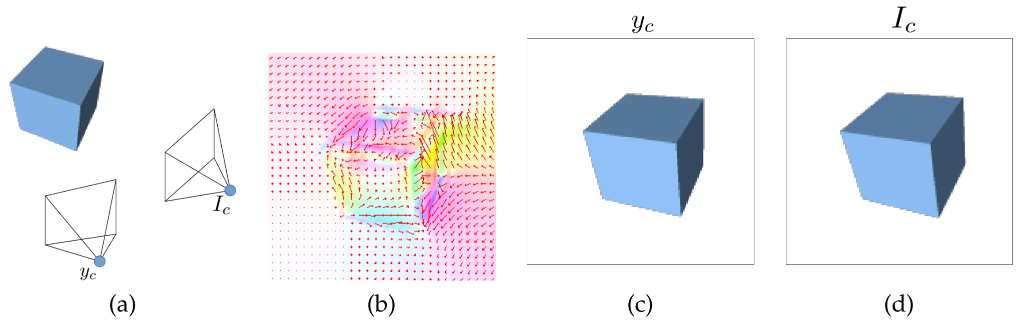
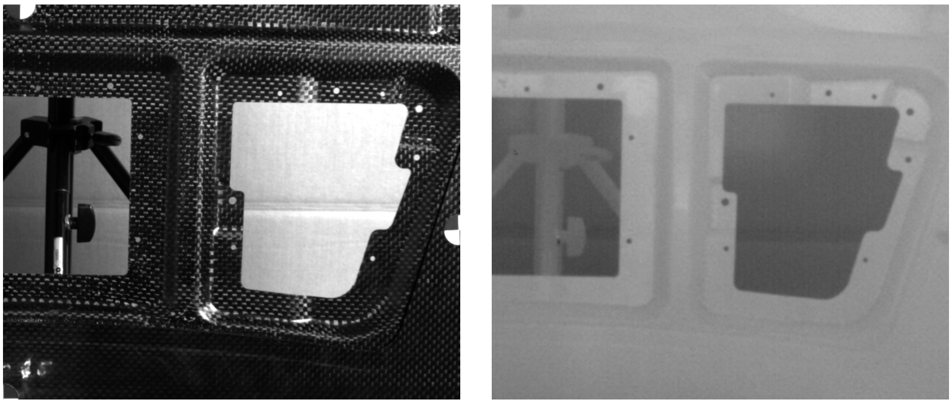


Figure 2.3: Figure 2.3a: Two cameras are shown recording a scene from different positions. The scene could could be a rigid scene or a dynamic scene with moving objects. Figure 2.3c shows the image y captured from the camera y_c and figure 2.3d the image I from the camera I_c. Figure 2.3b shows the optical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

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 (citation!!) 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 an 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 (citation!!). 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 [13] and of the human spine [14].

Optical flow models belong to the category of inverse problems ([**?**]).

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 $y(\boldsymbol{x})$ is the intensity recorded by the camera y_c at the pixel location $\boldsymbol{x} \in \Omega$ while $I(\boldsymbol{x}')$ is the intensity recorded by I_c at the location $\boldsymbol{x}' \in \Omega$. In figure 2.3a we have depicted a multi-modal setup in which the two cameras y_c and I_c are



(a) (b)

Figure 2.4

recording images (figures 2.3d and 2.3c) from different angles. In this context the optical flow field is the unknown variable $\boldsymbol{d}$ which maps the location $\boldsymbol{x}'$ in the image I to the location $\boldsymbol{x}$ in the image y

$$\boldsymbol{x} = \boldsymbol{x}' + \boldsymbol{d}\left(\boldsymbol{x}'\right) \tag{2.118}$$

The optical field $\boldsymbol{d}$ is shown in figure 2.3b as a set of vectors at every pixel $\boldsymbol{x}' \in \Omega$, whose magnitude and orientation reflect the motion of the pixel $\boldsymbol{x}'$. In an optical flow model the the latent variable X is the vector $\boldsymbol{d}$ and the data Y are the images y and I. The model is then described by the probability

$$p\left(\boldsymbol{d}|y, I\right) = p\left(y, I|\boldsymbol{d}\right) \cdot p\left(\boldsymbol{d}\right) \tag{2.119}$$

In the following we will give a short survey on the current types optical flow likelihoods $p\left(y, I|\boldsymbol{d}\right)$ and current state of the art priors $p\left(\boldsymbol{d}\right)$. We will then introduce Lie algebras and the Noether Theorem which will play a vital role the definition of our geometrical prior.

Among the earliest methods for optical flow estimation are the methods described in the seminal papers of Horn and Schunck [7] and Lukas and Kanade [15]. In [7] the following model for computing the flow between two frames of a video was proposed

$$E_{y,I}(\boldsymbol{d}) = E_{y,I}^{data}(\boldsymbol{d}) + \lambda E^{prior}(\boldsymbol{d}) \tag{2.120}$$

$$E_{y,I}^{data}(\boldsymbol{d}) = \int_{\Omega} \Big(y(\boldsymbol{x}) - I_{\boldsymbol{d}}(\boldsymbol{x}) \Big)^2 dx \quad I_{\boldsymbol{d}}(\boldsymbol{x}) = I(\boldsymbol{x} + \boldsymbol{d}(\boldsymbol{x})) \tag{2.121}$$

$$E^{prior}(\boldsymbol{d}) = \lambda \int_{\Omega} \sum_i \|\nabla d_i\|^2 \, dx \tag{2.122}$$

In eq. (2.122) the frame I is warped back to the frame y by the field $\boldsymbol{d}(\boldsymbol{x})$. The second integral in eq. (2.122) imposes an isotropic smoothness constraint on the flow field $\boldsymbol{d}$. The likelihood in eq. (2.122) makes the assumption that the brightness of the scene recorded by the camera is constant from frame to frame. This is a very strong constraint, which is rarely met in real world multi-modal setups. Figure 2.4 shows two images recorded from a visual spectrum camera (VSC, figure 2.4a) and a thermographic camera (TC, figure 2.4b). 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 figure 2.4a follow a completely different distribution then those in figure 2.4a . We need a model that can bring both images onto a common intensity space.

Furthermore the isotropic smoothness term in eq. (2.122) does not allow for discontinuities in $\boldsymbol{d}$. Several methods have been introduced which remove the assumption of isotropic flow [16, 17]. These Methods include (citation!!) TV-Regularization, anisotropic difusion guided by directional operators like the structure tensor and level set methods of the Mumford-Shah type [18]. We will introduce a methodology for the geometrical characterization of anisotropic priors in section 2.2 following a review of the TV-Regularization prior in section 2.5.

We will now discuss three statistical similarity measures (citation!!) for optical flow 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

$$\mathbb{E}(X) = \int X \cdot p(X) \tag{2.123}$$

$$\mathbb{V}ar(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2 \right) \tag{2.124}$$

2.6.1 Mutual Information

Mutual Information (MI) 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 $p\left(y,I\right)$ by

$$MI\left(y,I\right)=\int p\left(y,I\right)\ln\frac{p\left(y,I\right)}{p\left(y\right)\cdot p\left(I\right)}dydI \tag{2.125}$$

MI measures how strong the images y and I *statistically* depend on each other. In the case that y and I are statistically independent, $p\left(y,I\right)=p\left(y\right)\cdot p\left(I\right)$, then by eq. (2.125) 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 $I_{\boldsymbol{d}}$

$$E_{y,I}^{data}\left(\boldsymbol{d}\right)=-MI\left(y,I_{\boldsymbol{d}}\right) \tag{2.126}$$

However, as [19] 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.

2.6.2 Correlation Ratio

To alleviate the problems with MI, [19] 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 $I\left(\boldsymbol{x}\right)$ and $y\left(\boldsymbol{x}\right)$ 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 $p\left(\hat{y},\hat{I}\right)$, and the conditional distribution

$$p\left(\hat{y}|\,\hat{I}=I\right)=\frac{p\left(\hat{y},\hat{I}=I\right)}{p\left(\hat{I}=I\right)} \tag{2.127}$$

encodes the spacial functional relationship between y and I. They introduced the Correlation Ratio (CR)

$$\eta\left(I|y\right)=\frac{\mathbb{V}ar\left(\phi^{\star}\left(y\right)\right)}{\mathbb{V}ar\left(I\right)}\quad E_{y,I}^{data}\left(\boldsymbol{d}\right)=-\eta\left(I_{\boldsymbol{d}}|y\right) \tag{2.128}$$

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

$$\phi^\star(y) = \mathbb{E}\left(\hat{I}\middle|\,\hat{y} = y\right) = \int I p\left(I|\,y\right) dI \tag{2.129}$$

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. (2.128) is bounded between 0 and 1, 0 indicating that y and I are independent, 1 indicating a functional relationship $I = \phi^\star(y)$. The function $\phi^\star$ although not necessarily continuous, is measurable in the L_2-sense. Thus CR is a much stronger constraint then MI and has fewer, but more meaningful minima [19].

2.6.3 Cross Correlation

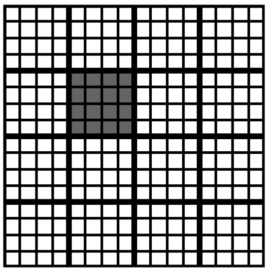
Cross Correlation (citation!!) is the strongest constrained similarity measure. It is basically an additional constraint to CR, namely that the functional relationship in eq. 2.128 must be linear. Then η reduces to

$$\eta\left(I|y\right) = \frac{Cov\left(y, I\right)}{\mathbb{V}ar\left(I\right) \cdot \mathbb{V}ar\left(y\right)} \quad I = \lambda \cdot y \tag{2.130}$$

As we will see in section **??** a measure similar to eq. (2.130) 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.

2.7 Setup of the camera rig

The data acquisition apparatus consists of a visible spectrum camera (VSC) mounted on top of a thermography camera (TC). The resolution of the VSC is 1226×1028 pixels while that of the TC is 640×512 pixels, both cameras with a focal length of 25 mm. We used a sinusoidal excitation source with a frequency of 0.1 Hz, which corresponds to a penetration depth of approximately 1.3 mm in the CFRP.



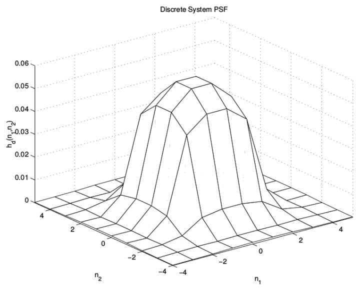


Figure 2.5: Figure 2.5a The thick grid depicts the CCD of the low resolution thermographic camera. The finer grid a virtual super-resolved version of the pixels in the TC. Figure 2.5b shows the point spread function $W_\sigma(x, y)$ of the gray pixel in figure 2.5a, taken from Hardie et al. [20]. It shows that each pixel in the TC image has a non uniform response over its surface to incoming photons.

2.8 Image Fusion

Our camera setup not only consists of two cameras with differing spectral responses, the TC and the VSC also differ in spatial resolution. However the likelihoods given introduced above have in common that they do not directly model the difference in resolution. In figure 2.5a a model of the CCD of the low resolution TC is shown overlaid with a higher resolution grid representing the VSC. The gray region in figure 2.5a 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 2.5b is the result of a theoretical model of a FLIR imager similar our TC. The model, obtained by Hardie et al. [20], 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 S of the TC image y with the help of the PSF W_σ, by stating that y is the result of the convolution of S with W_σ

$$y = W_\sigma s + n \quad n \sim \mathcal{N}(0|C_n) \tag{2.131}$$

The problem of estimating S is that there is an infinite amount of high resolution TC images $S^\star$ which relate to y via eq. (2.131) since the high spacial frequency components of S are filtered out by W_σ. In [21] Hardie suggested use of a high resolution imager I_c whose camera center is co-aligned (hence the subscript c) with the TC image y and correlated with S. The rationale behind their approach is to combine the desired features such as sharp edges and corners of I_c with the intensity spectrum of y into the super-resolved image S, while avoiding limitations such as the noise model of y. The limitation of their model is that the centers of the modalities y and I_c need to be co-linear. While this is the case in remote sensing applications, the model needs to be extended to the general case of two separated modalities. We will first outline the original model, and in chapter 3.3 we will introduce a new model for optical flow based on [21].

The key ingredient in the model of [21] is that the intensities of S and I_c are assumed to be samples drawn from the joint Gaussian $p\left(S, I_c\right)$. As I_c is already fixed as input data we can derive a conditional distribution for S via the Bayesian rule

$$p\left(S|I_c\right) = \frac{p\left(S, I_c\right)}{p\left(I_c\right)} \sim \mathcal{N}\left(\mu_{s|I_c}|C_{s|I_c}\right) \tag{2.132}$$

$$C_{s|I_c} = C_{s,s} - C^2_{s,I_c} \cdot C^{-1}_{I_c,I_c} \tag{2.133}$$

$$\mu_{s|I_c}\left(\mathbf{x}\right) = \mu_s + C_{s,I_c} \cdot C^{-1}_{I_c,I_c}\left(I_c\left(\boldsymbol{x}\right) - \mu_{I_c}\right) \tag{2.134}$$

where the variances are computed globally

$$C_{u,v} = \int_\Omega \left(u\left(\boldsymbol{x}\right) - \mu_u\right) \cdot \left(v\left(\boldsymbol{x}\right) - \mu_v\right) dx \tag{2.135}$$

We see that the mean of S conditioned on I_c, $\mu_{s|I_c}$ is linear in the values of I_c, thus in this model the intensities of S are assumed to be *globally* linearly related to the intensities of I_c. We combine eq. (2.132) with the Gaussian likelihood in eq. (2.131) to the posterior

$$p\left(S|y, I_c\right) \sim p\left(y|S\right) \cdot p\left(S|I_c\right) = \exp\left(-E\left(S\right)\right) \tag{2.136}$$

with the associated energy

$$E\left(S\right) = \frac{1}{2}\int_\Omega \Big(y\left(\boldsymbol{x}\right) - W_\sigma S\left(\boldsymbol{x}\right)\Big)^2 \cdot C_n^{-1} dx + \frac{1}{2}\int_\Omega \Big(S\left(\boldsymbol{x}\right) - \mu_{s|I_c}\left(\boldsymbol{x}\right)\Big)^2 \cdot C^{-1}_{s|I_c} dx \tag{2.137}$$

The minimization of eq. (2.137) and thus maximization of (2.136) with respect to

S gives the analytical solution [21]

$$\hat{s} = \mu_{s|I_c} + C_{s|I_c} \cdot W_\sigma^T \left(W_\sigma \cdot C_{s|I_c} \cdot W_\sigma^T + C_n\right)^{-1} \left(y - W_\sigma \mu_{s|I_c}\right) \tag{2.138}$$

Eq. (2.138) is intractable to compute due to the dense operator W_σ and the matrix-inverse operation. In [22] a computationally tractable approximation was introduced

$$\hat{s} = \mu_{s|I_c} + C_{\tilde{s}|\tilde{I}_c} \cdot \left(C_{\tilde{s}|\tilde{I}_c} + C_n\right)^{-1} \left(y - \tilde{\mu}_{s|I_c}\right) \tag{2.139}$$

$$\tilde{I}_c = W_\sigma I_c \quad \tilde{s} = W_\sigma s \approx y \tag{2.140}$$

The key issue is that this method requires both modalities, I_c and y, to be co-registered. Since we are dealing with an optical flow problem y and thus S is shifted by a disparity $\boldsymbol{d}(\boldsymbol{x})$ from I_c. This disparity has to be taken in to account by our model in chapter 3.3. The second issue is that the assumption that S and I_c are *globally* joint Gaussian is not supported by our data. However by computing $C_{s|I_c}$ in *local* sub-domains of the space Ω we can show that S and I_c 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 $\boldsymbol{v}$ 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

$$V_e^\Omega = v_x(\boldsymbol{x})\,\partial_x + v_y(\boldsymbol{x})\,\partial_y \in \mathfrak{t} \qquad \boldsymbol{v}(\boldsymbol{x})\big|_A = const \tag{3.1}$$

Consider an image $\phi(\boldsymbol{x})$. Under a one parameter transformation $g_{\gamma(t)} \in \mathbb{T}$ the vector V_e is invariant since

$$\frac{d}{dt} V_{g_{\gamma(t)}} \phi \bigg|_{t=0} = \omega_x [V_e, \partial_x]\,\phi + \omega_y [V_e, \partial_y]\,\phi \tag{3.2}$$

and the basis $\{\partial_x, \partial_y\}$ is commutative. The level-sets S_X corresponding to the vector X_e^Ω are are defined by

$$S_X = \left\{ \boldsymbol{x} \,\middle|\, \boldsymbol{v}^T \cdot \nabla\phi(\boldsymbol{x}) = 0 \right\} \tag{3.3}$$

We would like to characterize the dominant strength and the orientation of $\nabla\phi$ within the sub domain $A \subset \Omega$. In [23] it was suggested that the tangential vector $\boldsymbol{v}$ of the level sets S_X can be computed by minimizing the energy

$$J(\boldsymbol{v}) = \frac{1}{2}\int_A w(\boldsymbol{x})\,\boldsymbol{v}^T \cdot \Big(\nabla\phi(\boldsymbol{x})\,\nabla^T\phi(\boldsymbol{x})\Big)\,\boldsymbol{v} = \frac{1}{2}\boldsymbol{v}^T S \boldsymbol{v} \tag{3.4}$$

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

$$S = V^T D V \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad V = (\boldsymbol{V}_1, \boldsymbol{V}_2) \tag{3.5}$$

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 $\boldsymbol{v} = \boldsymbol{V}_2$
- $\lambda_1 \approx \lambda_2 \approx 0$: No strong gradient, image is approximately constant
- $\lambda_1 \approx \lambda_2 \gg 0$: No *linear* level sets, level sets have strong curvature

3.2 Structure Tensor Based Prior

Since the vector field V_e^{Ω} is translation invariant the structure tensor S is also translation invariant. Under the rotation group $SO\left(2\right)$ 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 $R_\theta \in SO\left(2\right)$

$$S' = R_\theta^T S R_\theta \tag{3.6}$$

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

$$E_{ST}^{prior} = \int_\Omega \mathcal{E}_{ST}\left(\nabla\phi\right) d^2x \tag{3.7}$$

$$\mathcal{E}_{ST}\left(\nabla\phi\right) = \frac{\lambda}{2}\det\left(S\right) \tag{3.8}$$

3.2.1 The Divergence-Free Vectors under $\mathbb{T} \times SO(2)$

The Lie algebra of the group $\mathbb{G} = \mathbb{T} \times SO(2)$ is the algebra $\mathcal{G} = \mathfrak{t} \times \mathfrak{so}(2)$ which has the basis $\{\partial_x, \partial_y, \partial_\theta\}$. The subset $\mathfrak{t} = \left\{X_e^{\Omega,x}, X_e^{\Omega,y}\right\}$ is the set generators of infinitesimal translations

$$X_e^{\Omega,x} = \partial_x, \quad X_e^{\Omega,y} = \partial_y \tag{3.9}$$

$\mathfrak{t}$ 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

$$X_e^{\Omega,\theta} = -y\partial_x + x\partial_y \tag{3.10}$$

From eq. (3.10) we can see that ∂_θ does not commute with t and the commutators for the basis $\left\{X_e^{\Omega,x}, X_e^{\Omega,y}, X_e^{\Omega,\theta}\right\}$ are easily computed

$$\left[X_e^{\Omega,x}, X_e^{\Omega,\theta}\right] = X_e^{\Omega,y} \quad \left[X_e^{\Omega,y}, X_e^{\Omega,\theta}\right] = -X_e^{\Omega,x} \quad \left[X_e^{\Omega,x}, X_e^{\Omega,y}\right] = 0 \tag{3.11}$$

We apply the commutators in eq. (3.11) to compute the divergences of the vectors $\boldsymbol{W}_x$, $\boldsymbol{W}_y$ and $\boldsymbol{W}_\theta$ (see eq. (2.82)) from the divergence equation in eq. (2.84) which we repeat

$$\operatorname{div}(\boldsymbol{W}_m) = \operatorname{div}(\omega_m)\,\mathcal{E} - \sum_i \left[X_e^{\Omega,i}, X_e^{\Omega,m}\right](\phi) \cdot \frac{\delta\mathcal{E}}{\delta\left(X_e^{\Omega,i}\phi\right)} \tag{3.12}$$

The representation of the basis operators $\left\{X_e^{\Omega,x}, X_e^{\Omega,y}, X_e^{\Omega,\theta}\right\}$ with respect to the Cartesian coordinate frame Ω

$$\boldsymbol{\omega}_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \boldsymbol{\omega}_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \boldsymbol{\omega}_\theta = \begin{pmatrix} -y \\ x \end{pmatrix} \tag{3.13}$$

We see that the divergences of all vectors in eq. (3.13) vanish. Since the structure tensor S is only a function on $X_e^{\Omega,x}(\phi)$ and $X_e^{\Omega,y}(\phi)$ the divergences for the vectors $\boldsymbol{W}_x$ and $\boldsymbol{W}_y$ vanish for any field configuration ϕ

$$\operatorname{div}(\boldsymbol{W}_x) = 0, \quad \operatorname{div}(\boldsymbol{W}_y) = 0 \tag{3.14}$$

The divergence of the rotational vector $\boldsymbol{W}_\theta$ is

$$\operatorname{div}(\boldsymbol{W}_\theta) = -\left[X_e^{\Omega,x}, X_e^{\Omega,\theta}\right](\phi) \cdot \frac{\delta\mathcal{E}}{\delta\left(X_e^{\Omega,x}\phi\right)} - \left[X_e^{\Omega,y}, X_e^{\Omega,\theta}\right](\phi) \cdot \frac{\delta\mathcal{E}}{\delta\left(X_e^{\Omega,y}\phi\right)} \tag{3.15}$$

$$= -\left(X_e^{\Omega,x}\phi, X_e^{\Omega,y}\phi\right) \cdot \boldsymbol{M}_\theta \cdot \boldsymbol{P}, \quad \boldsymbol{P} = \begin{pmatrix} \frac{\delta\mathcal{E}}{\delta\left(X_e^{\Omega,x}\phi\right)} \\ \frac{\delta\mathcal{E}}{\delta\left(X_e^{\Omega,y}\phi\right)} \end{pmatrix} \tag{3.16}$$

where $\boldsymbol{M}_\theta$ is the Pauli matrix

$$\boldsymbol{M}_\theta = \begin{pmatrix} 0, & -1 \\ 1, & 0 \end{pmatrix} \tag{3.17}$$

The Pauli matrix $\boldsymbol{M}_\theta$ is the matrix corresponding to infinitesimal rotations. It rotates canonical momentum $\boldsymbol{P}$ to the orthogonal vector $\boldsymbol{P}^\perp$

$$\boldsymbol{P}^\perp = \begin{pmatrix} -\frac{\delta\mathcal{E}}{\delta\left(X_e^{\Omega,y}\phi\right)} \\ \frac{\delta\mathcal{E}}{\delta\left(X_e^{\Omega,x}\phi\right)} \end{pmatrix} = \boldsymbol{M}_\theta \cdot \boldsymbol{P} \tag{3.18}$$

For the structure tensor prior in eq. (3.8) the orthogonal canonical momentum is

$$\boldsymbol{P}^\perp = S \cdot \langle\nabla\phi\rangle^\perp, \quad \langle\nabla\phi\rangle^\perp = \boldsymbol{M}_\theta \cdot \langle\nabla\phi\rangle \tag{3.19}$$

where $\langle\nabla\phi\rangle$ is the mean of the gradient $\nabla\phi$ in the sub domain $A \subset \Omega$

$$\langle\nabla\phi\rangle = \int_A w\left(\boldsymbol{x}\right) \nabla\phi\left(\boldsymbol{x}\right) d^2x \tag{3.20}$$

One can show that $\langle\nabla\phi\rangle$ is an eigenvector to the structure tensor S in the direction of the maximal eigenvalue λ_{max}. It follows that the perpendicular vector $\langle\nabla\phi\rangle^\perp$ is also an eigenvector to S but in the direction of the minimal eigenvalue λ_{min}

$$\boldsymbol{S} \cdot \langle\nabla\phi\rangle^\perp = \lambda_{min}\langle\nabla\phi\rangle^\perp = \boldsymbol{P}^\perp \tag{3.21}$$

so that the perpendicular canonical momentum $\boldsymbol{P}^\perp$ is also an eigenvector in the direction of λ_{min}

$$\boldsymbol{S} \cdot \boldsymbol{P}^\perp = \lambda_{min}\boldsymbol{P}^\perp \tag{3.22}$$

The divergence of the vector $\boldsymbol{W}_\theta$ with the canonical momentum in eq. (3.21) is then

$$\mathrm{div}\left(\boldsymbol{W}_\theta\right) = -\lambda_{min}\nabla^T\phi \cdot \langle\nabla\phi\rangle^\perp \tag{3.23}$$

The scalar product $\nabla^T\phi\langle\nabla\phi\rangle^\perp$ and λ_{min} in eq. (3.23) are non-zero for noise inflicted fields ϕ. However eq. (3.23) shows that the divergence operation is the first order Taylor approximation of ϕ under the spacial transformation

$$\boldsymbol{x} \rightarrow \boldsymbol{x} - \tau\lambda_{min}\langle\nabla\phi\rangle^\perp \tag{3.24}$$

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. [21]. We will

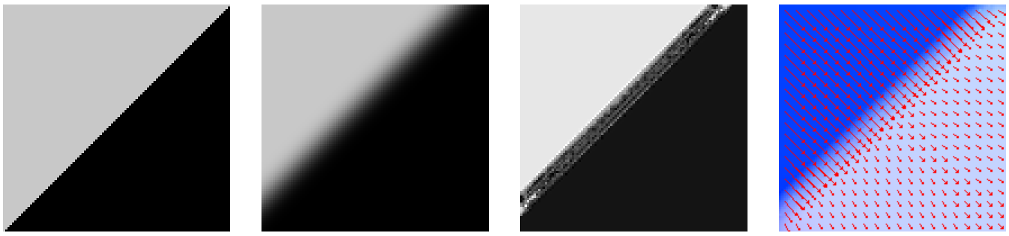


Figure 3.1: Figure 3.1a shows a synthetic high resolution image I^{syn}. In figure 3.1b we show a low resolution image y^{syn}. y^{syn} is computed by convolution of I^{syn} with Gaussian G_σ with standard deviation $\sigma = 5$ and translated by 10 pixels relative to I^{syn}. Figure 3.1d shows the flow $\boldsymbol{d}$ computed with the model in eq. (2.122), which does not incorporate knowledge of the scale difference between y^{syn} and I^{syn} and figure 3.1c show the warped image $I^{syn}_{\boldsymbol{d}}$

address the two issues outlined in section 2.8, namely that the images y and I (Figures **??**b and **??**d) 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 $\boldsymbol{d}(\boldsymbol{x})$ 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.

Problem a: The images y and I have different intensity distributions, since the TC and the VSC are sensitive to different spectra.

Problem b: The images y and I have different resolutions.

Problem 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 $\boldsymbol{d}$ can only be estimated with a likelihood $p\,(y, I|\boldsymbol{d})$ which measures how similar the images y and I are given $\boldsymbol{d}$. However a likelihood that measures the similarity of the intensities of y and I like the one in eq. (2.122) 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 $\boldsymbol{d}$ 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 3.1. I^{syn} in figure 3.1a shows a sharp linear boundary and y^{syn} (figure 3.1b) 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

$$E\left(\boldsymbol{d}^{syn}\right) = \frac{1}{2}\int_\Omega \Big(y^{syn}\left(\boldsymbol{x}\right) - I^{syn}_{\boldsymbol{d}^{syn}}\left(\boldsymbol{x}\right)\Big)^2 dx + \frac{\lambda}{2}\sum_i \int_\Omega \left\|\nabla d^{syn}_i\left(\boldsymbol{x}\right)\right\|^2 dx \quad (3.25)$$

(see eq. (2.122)) to compute the optical flow $\boldsymbol{d}^{syn}$ mapping I^{syn} to y^{syn} (see figure 3.1d). Figure 3.1c shows the image $I^{syn}_{\boldsymbol{d}^{syn}}\left(\boldsymbol{x}\right) = I^{syn}(\boldsymbol{x} + \boldsymbol{d}^{syn}\left(\boldsymbol{x}\right))$. We can see that the optical flow $\boldsymbol{d}$ 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 3.1b).

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 $p\left(y, S|\boldsymbol{d}\right)$ can measure how similar the images y and S are given $\boldsymbol{d}$.

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

In section 2.8 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. (2.132))

$$p\left(S|y, I_c\right) = p\left(y|S\right) \cdot p\left(S|I_c\right) \quad (3.26)$$

$$-\ln\left(p\left(y|S\right)\right) = \frac{1}{2}\int_\Omega \Big(y\left(\boldsymbol{x}\right) - W_\sigma S\left(\boldsymbol{x}\right)\Big)^2 \cdot C_n^{-1} dx \quad (3.27)$$

$$-\ln\left(p\left(S|I_c\right)\right) = \frac{1}{2}\int_\Omega \Big(S\left(\boldsymbol{x}\right) - \mu_{s|I_c}\left(\boldsymbol{x}\right)\Big)^2 \cdot C_{s|I_c}^{-1} dx \quad (3.28)$$

with the conditional variance and mean

$$C_{S|I_c} = C_{S,S} - C^2_{S,I_c} \cdot C^{-1}_{I_c,I_c} \quad (3.29)$$

$$\mu_{S|I_c} = \mu_S + C_{S,I_c} \cdot C^{-1}_{I_c,I_c}\left(I - \mu_I\right) \quad (3.30)$$

In the conditional prior $p\left(S|I_c\right)$ in eq. (3.28) 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 $\boldsymbol{d}(\boldsymbol{x})$,

$$I_c\left(\boldsymbol{x}\right) = I\left(\mathbf{x} + \mathbf{d}\left(\mathbf{x}\right)\right) = I_{\mathbf{d}}\left(\mathbf{x}\right) \quad (3.31)$$

Substituting eq. (3.31) into eq. (3.26) and following, we obtain the posterior

$$p\left(S|y,I,\mathbf{d}\right)=p\left(S|y,I_{\mathbf{d}}\right) \tag{3.32}$$

with the energy

$$E_{post}\left(S,\boldsymbol{d}\right)=\frac{1}{2}\int_{\Omega}\Big(y\left(\boldsymbol{x}\right)-W_{\sigma}S\left(\boldsymbol{x}\right)\Big)^{2}\cdot C_{n}^{-1}dx+\frac{1}{2}\int_{\Omega}\Big(S\left(\boldsymbol{x}\right)-\mu_{s|I_{\boldsymbol{d}}}\left(\boldsymbol{x}\right)\Big)^{2}\cdot C_{s|I_{\boldsymbol{d}}}^{-1}dx \tag{3.33}$$

While keeping $\boldsymbol{d}$ fixed we minimize $E_{post}\left(S,\boldsymbol{d}\right)$ with respect to S and obtain similar to eq. (2.139) a closed form solution for S

$$\hat{S}=\mu_{s|I_{\boldsymbol{d}}}+C_{\tilde{s}|\tilde{I_{\boldsymbol{d}}}}\cdot\Big(C_{\tilde{s}|\tilde{I_{\boldsymbol{d}}}}+C_{n}\Big)^{-1}\Big(y-\tilde{\mu}_{s|I_{\boldsymbol{d}}}\Big) \tag{3.34}$$

We insert the simplified closed form expression for $\hat{S}$ from eq. (3.34) into E_{post} and obtain an energy measuring the similarity between y and $\tilde{I}_{\boldsymbol{d}}=W_{\sigma}I_{\boldsymbol{d}}$

$$E_{data}\left(\boldsymbol{d}\right)=E_{post}\Big(\hat{S},\boldsymbol{d}\Big) \tag{3.35}$$

$$=\frac{1}{2}\int_{\Omega}\Big(y\left(\boldsymbol{x}\right)-f\cdot\tilde{I}_{\boldsymbol{d}}\left(\boldsymbol{x}\right)\Big)^{2}\cdot C_{s|\tilde{I}_{\boldsymbol{d}}}\Big(C_{s|\tilde{I}_{\boldsymbol{d}}}+\lambda C_{n}\Big)^{-2} \tag{3.36}$$

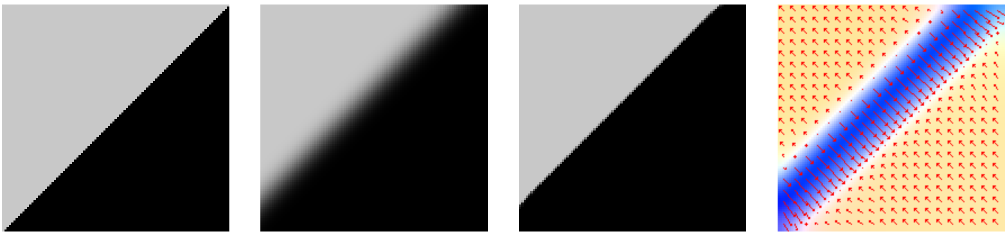
$$f=C_{y,\tilde{I}_{\boldsymbol{d}}}C_{\tilde{I}_{\boldsymbol{d}},\tilde{I}_{\boldsymbol{d}}}^{-1} \tag{3.37}$$

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

$$p\left(y,I|\boldsymbol{d}\right)=\exp\left(-E_{data}\left(\boldsymbol{d}\right)\right) \tag{3.38}$$

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. (3.38) 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 $p\left(y|S\right)$ in eq. (3.27) to the TC image y. The prior $p\left(S|I\right)$ in eq. (3.28) couples S to the high resolution image I. As a result $I_{\boldsymbol{d}}$ in E_{data} in eq. (3.42) 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 $\tilde{I}_{\boldsymbol{d}}$ to a range similar to that of y so that E_{data} is a measure for the similarity between y and $f\cdot\tilde{I}_{\boldsymbol{d}}$.

To demonstrate that our likelihood E_{data} in eq. (3.38) 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 3.1. 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 aproximately the same. The



(a) (b) (c) (d)

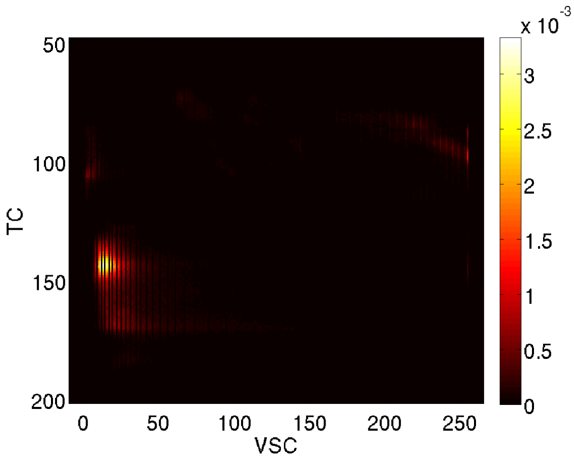
Figure 3.2: Figure 3.2a shows a synthetic high resolution image I^{syn}. In figure 3.2b we show a low resolution image y^{syn}. y^{syn} is computed by convolution of I^{syn} with Gaussian G_σ with standard deviation $\sigma = 5$ and translated by 10 pixels relative to I^{syn}. Figure 3.2d shows the flow $\boldsymbol{d}$ computed with the model in eq. (3.38), which incorporates knowledge of the scale difference between y^{syn} and I^{syn} and figure 3.2c show the warped image $I^{syn}_{\boldsymbol{d}}$

image $I^{syn}_{\boldsymbol{d}}$ is convolved with W_σ. The resulting image $\tilde{I}^{syn}$ has the same scale as y^{syn}. The resulting optical flow $\boldsymbol{d}^{syn}$ is shown in figure 3.2d. Notice the blurred boundary $\boldsymbol{d}^{syn}$ around the linear feature in I^{syn} (figure 3.2a). This is the result of E_{data} in eq. (3.38) measuring the difference between y^{syn} and the blurred image $\tilde{I}^{syn}_{\boldsymbol{d}} = W_\sigma I^{syn}_{\boldsymbol{d}}$. In eq. (3.2c) we see $I^{syn}_{\boldsymbol{d}}$. The linear boundary has been warped by $\boldsymbol{d}^{syn}$ without being corrupted like in figure 3.1c.

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 **??** the assumption of linearity fails. In figure 3.3 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 3.4 we have evaluated the histogram within local region of interests. The histograms in figures 3.4c and 3.4f 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. (2.135)

$$C_{u,v}(\boldsymbol{x}_0) = \int_\Omega \omega(\boldsymbol{x} - \boldsymbol{x}_0)(u(\boldsymbol{x}) - \mathbb{E}(u, \boldsymbol{x}_0)) \cdot (v(\boldsymbol{x}) - \mathbb{E}(v, \boldsymbol{x}_0)) \tag{3.39}$$



(a)

Figure 3.3: 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

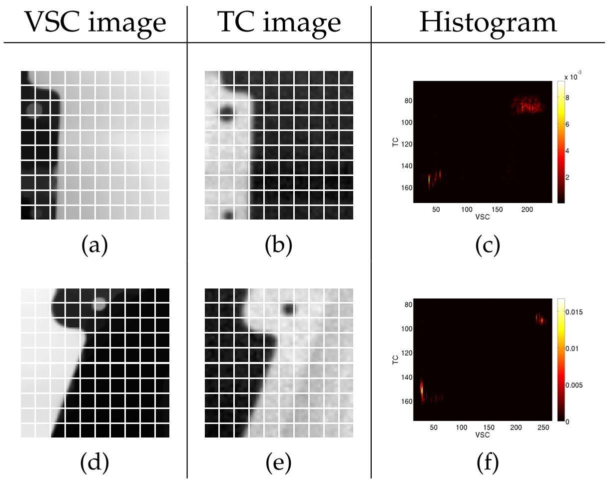
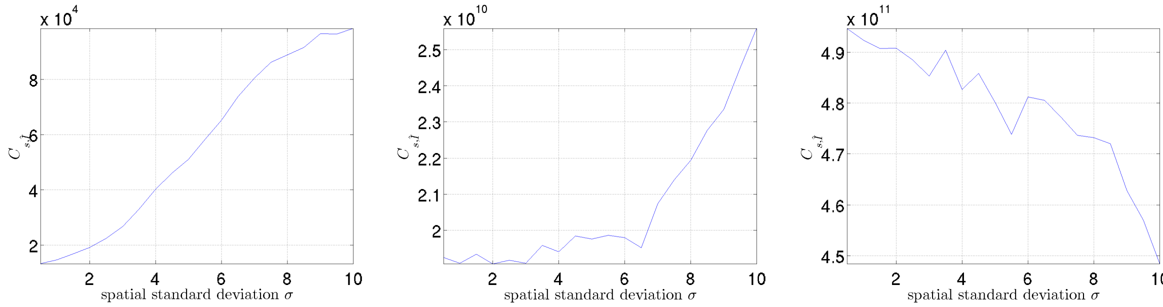


Figure 3.4: Different roi's and their joint histograms. A grid is shown in the VSC and the TC image to emphasize the disparity between them. The gridsize is 10 pixels. In the histograms we see there is a linear relationship between the VSC and the TC roi's



(a) (b) (c)

Figure 3.5: Median conditional variance $\hat{C}^{\sigma,a}_{s|I}$ for $a = 5$ (figure 3.5a), $a = 23$ (figure 3.5b) and $a = 33$ (figure 3.5c). We can see that for small values of a $\hat{C}^{\sigma,a}_{s|I}$ has a minimum at $\sigma < 2$, and for larger values of a the profile changes so that the minimum of $\hat{C}^{\sigma,a}_{s|I}$ is at $\sigma \geq 10$

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

$$\omega(\boldsymbol{x}) = \begin{cases} \frac{1}{|W|-1} & 0 \leq x, y \leq a \\ 0 & else \end{cases} \tag{3.40}$$

Then $C^{\sigma,a}_{s|I}(\boldsymbol{x})$ becomes a local meassure that meassures 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 loose the local linearity between the TC and the VSC image. In figure 3.5 we have plotted the median conditional variance

$$\hat{C}^{\sigma,a}_{s|I} = \text{median}\left(C^{\sigma,a}_{s|I}(\boldsymbol{x})\right) \tag{3.41}$$

as a function of σ for three fixed values of the window size a. In figure 3.5a ($a = 5$) $\hat{C}^{\sigma,a}_{s|I}$ has a minimum for $\sigma < 2$, and in figure 3.5c ($a = 33$) it is minimal for $\sigma \geq 10$. The profile of $\hat{C}^{\sigma,a}_{s|I}$ changes from monotonic increasing to monotonic decreasing for small to large values of a. Since we know the value for the scale parameter σ, $\sigma^\star = 2$ from the ccd resolutions of the cameras, the idea find the optimal value $a^\star$ such that $\hat{C}^{\sigma,a^\star}_{s|I}$ is minimal at $\sigma = \sigma^\star$. For $a = 23$ this is the case as we see in figure 3.5b. Thus for the data in figure **??** $a^\star = 23$ is the optimal value so that $\hat{C}^{\sigma,a^\star}_{s|I}$ has physically meaningful minimum $\sigma^\star = 2$. The local data term E_{data} now has the form

$$E_{data}\left(\boldsymbol{d}\right) = \frac{1}{2}\int_{\Omega}\left(y\left(\boldsymbol{x}\right) - f(\boldsymbol{x})\cdot\tilde{I}_{\boldsymbol{d}}\left(\boldsymbol{x}\right)\right)^2 \cdot C^{\sigma^\star,a^\star}_{s|\tilde{I}_{\boldsymbol{d}}}\left(\boldsymbol{x}\right)\left(C^{\sigma^\star,a^\star}_{s|\tilde{I}_{\boldsymbol{d}}}\left(\boldsymbol{x}\right) + \lambda C_n\right)^{-2} \tag{3.42}$$

$$f(\boldsymbol{x}) = C_{y,\tilde{I}_{\boldsymbol{d}}}(\boldsymbol{x})C^{-1}_{\tilde{I}_{\boldsymbol{d}},\tilde{I}_{\boldsymbol{d}}}(\boldsymbol{x}) \tag{3.43}$$

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

$$E\left(\boldsymbol{d}\right) = E_{data}\left(\boldsymbol{d}\right) + \frac{\lambda}{2}\left(\sum_i \mathrm{Det}\left(S\left(d_i\right)\right)\right) \tag{3.44}$$

The matrix $S\left(d_i\right)$ is the structure tensor (see eq. (3.4)) acting on each component of the optical flow $\boldsymbol{d}$. 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 3.44 and obtain the optimum flow field $\hat{\boldsymbol{d}}$ we deploy a simple newton scheme with a nested linearization of 3.44. 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 3.44 consists of one part comming 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 3.44 with respect to $\boldsymbol{d}$. At small k its eigenvalues are small due to the initial guess $\boldsymbol{d}_0 = 0$. The matrix Q_k^{data} is the second derivative of the likelihood in eq. (3.44). 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 3.44 to include an L_2 prior on the flow field $\boldsymbol{d}$ but with a small lagrange multiplier

Algorithm 1 Optical Flow with Structure Tensor prior

Initialize $\boldsymbol{d}_0 = 0$
Set $\boldsymbol{r}_0 = \frac{\delta E(d)}{\delta \boldsymbol{d}} (\boldsymbol{d}_0)$
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 $\|\boldsymbol{r}\| > \epsilon$ or $k < N$ **do**
    set $\boldsymbol{d}_{k+1} = \boldsymbol{d}_k + \alpha \boldsymbol{\delta}$
    expand $E(\boldsymbol{d}_{k+1}) = E(\boldsymbol{d}_k) + \alpha \boldsymbol{b}_k^T \boldsymbol{\delta} + \frac{\alpha^2}{2} \boldsymbol{\delta}^T Q_k \boldsymbol{\delta}$
    solve $Q_k \boldsymbol{\delta} = \boldsymbol{b}_k$ for δ with conjugate gradients and suitable preconditioning
    compute $\boldsymbol{d}_{k+1} = \boldsymbol{d}_k + \alpha \boldsymbol{\delta}$, $k \to k+1$
  end while
  upsample d_N, set $d_0 = d_N$, $k = 0$
  $s = s - 1$
end while

λ_2

$$E(\boldsymbol{d}) = \int (y - \hat{s}_{I,\boldsymbol{d}})^2 \cdot C_{s|I_{\boldsymbol{d}}} + \frac{\lambda}{2} \left(\sum_i (\text{Det}(S(d_i)) + \lambda_2 \|\nabla d_i\|) \right) \tag{3.45}$$

With the L_2 prior in 3.45 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

$$e_{EPE} = \|\boldsymbol{v} - \boldsymbol{v}_{gt}\| \tag{3.46}$$
$$e_{AE} = \cos(\sphericalangle(\boldsymbol{v}, \boldsymbol{v}_{gt})) \in \{-1, 1\} \tag{3.47}$$

The EPE e_{EPE} meassures how well the computed optical flow $\boldsymbol{v}$ fits the true optical flow $\boldsymbol{v}_{gt}$. In cases where $\boldsymbol{v}$ does not match $\boldsymbol{v}_{gt}$ 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
	7	$2.36, 0.01, 7.24$	$0.42, -1.00, 1.00$
figure 3.8a	9	$1.32, 0.00, 6.02$	$0.87, -1.00, 1.00$
	11	$1.15, 0.00, 6.45$	$0.91, -1.00, 1.00$
	7	$0.84, 0.01, 13.35$	$0.87, -1.00, 1.00$
figure 3.8f	9	$0.46, 0.01, 8.23$	$0.97, -1.00, 1.00$
	11	$0.40, 0.00, 8.25$	$0.98, -1.00, 1.00$
	7	$0.47, 0.01, 5.22$	$0.97, -0.96, 1.00$
figure 3.9a	9	$0.28, 0.00, 3.71$	$0.99, -1.00, 1.00$
	11	$0.25, 0.00, 2.50$	$0.99, -1.00, 1.00$
	7	$0.44, 0.00, 2.73$	$0.98, -1.00, 1.00$
figure 3.9f	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 3.8a to 3.9f). 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 3.8) generally have higher EPE values than the roi's with linear or constant geometry (figure 3.9)

In figure 3.6 the rubber whale sequence of the middleburry data set is shown, and in figure 3.6b the corresponding ground truth $\boldsymbol{v}_{gt}$. In figure 3.6d the computed flow-field $\boldsymbol{v}$ is shown for a filter size of 11, while in figure 3.6c the resulting flow for the TV model is shown. Figures 3.8 and 3.9 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 3.10 and 3.11 show the same for the TV model. We can observe from the comparison between figures 3.6d and 3.6c that the TV

Figure	Median, Min, Max EPE	Median, Min, Max AE
figure 3.10a	1.38, 0.00, 5.83	0.92, −1.00, 1.00
figure 3.10f	0.20, 0.00, 3.34	1.00, −1.00, 1.00
figure 3.11a	0.04, 0.00, 2.08	1.00, −1.00, 1.00
figure 3.11f	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 3.10a to 3.11f). 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 3.12a	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 3.12f	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 3.7a are all curvilinear there is no correlation between the filtersize and the EPE.

model produces smoother results which are closer to the ground truth (figure 3.6b). 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 obeserve 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 3.9 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 3.8. The roi with the largest discrepancy from the group of linear motions is figure 3.8a which depicts a rotating wheel. Since the wheel is largely free of texture, the motion field (figure 3.8d) is penalized by the structure tensor prior in such a way that it aquires spurious linear motion boundaries. This is the reason for its high EPE value. The roi in figure 3.8f shows another case of a motion field violating

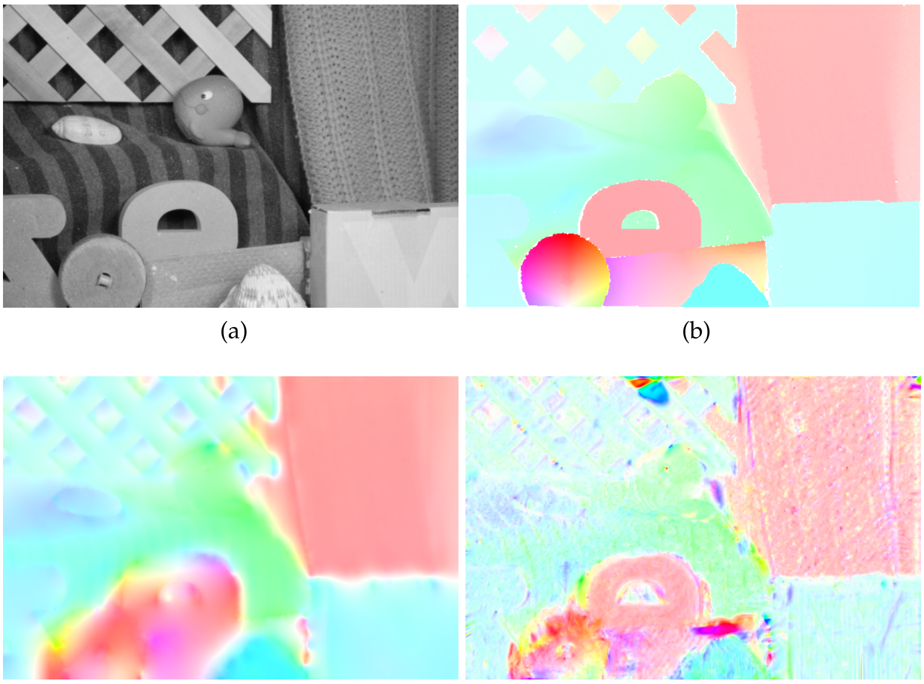
Figure	Median, Min, Max EPE	Median, Min, Max AE
figure 3.13a	0.44, 0.00, 6.12	1.00, −1.00, 1.00
figure 3.13f	0.12, 0.01, 7.38	1.00, 0.00, 1.00

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

the assumption of linear motion boundaries. In the ground truth roi in figure 3.8j 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 3.10d 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 3.10i shows the resulting flow of the TV model at the two junctions in figure 3.10f. 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 3.9 and 3.11). In figure 3.9d we see that the structure tensor model inflicts more of the texture from the underlying data (figure 3.9a) on the estimated flow then the TV model (see figure 3.11d for the result of the TV model) thus leading to a slightly higher EPE value (table 3.1). Figure 3.9i 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 3.9i, we see that the structure tensor model produces sharper lineat motion boundaries.

In figure 3.7 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 3.7d), see the EPE and AE values in figure 3.12 and table 3.3.



(a) (b)

(c) (d)

Figure 3.6: Rubberwhale 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

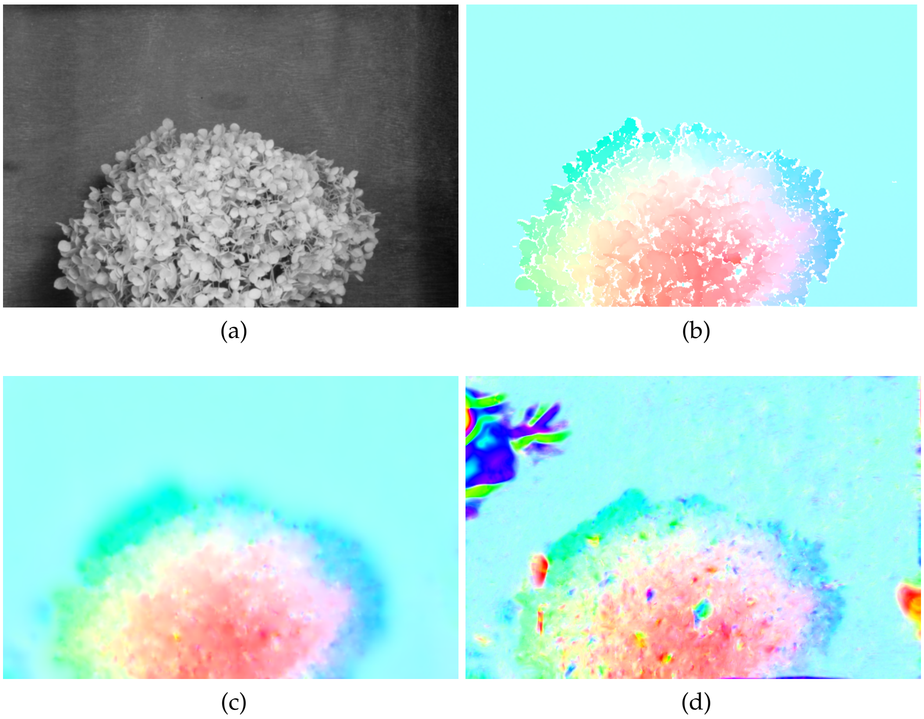


Figure 3.7: Hydrangea Sequence

Figure 3.7a shows one frame of the sequence. figure 3.7d shows the estimated optical flow, figure 3.7c the result of the TV model and figure 3.7b shows the provided ground truth

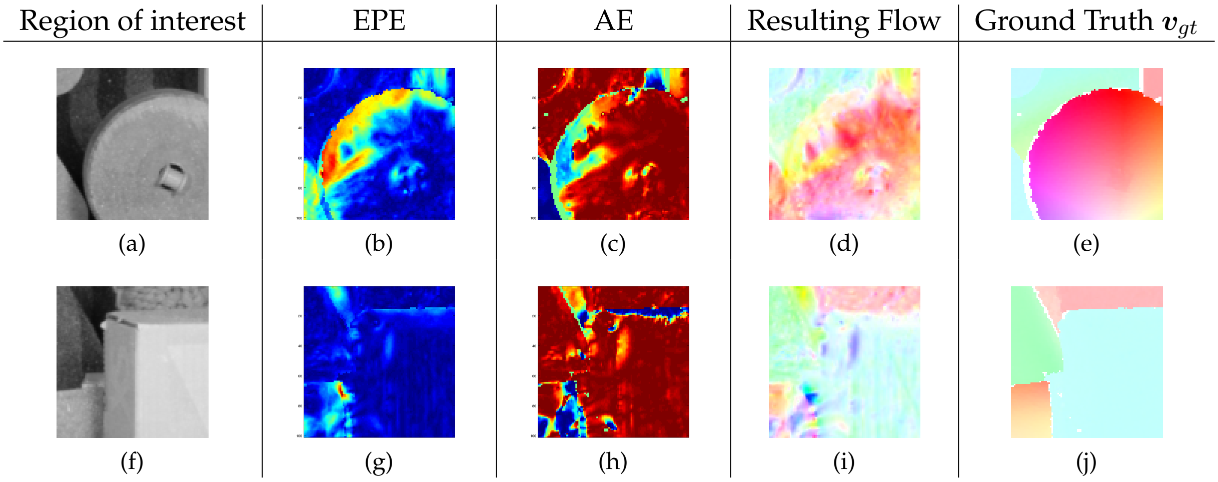


Figure 3.8: Error Analysis ST model: This figure shows two examples of motion field with nonlinear boundaries. In figure 3.8c we see that along the circumference of the wheel the EPE has the largest values and in figure 3.8h the is largest the junction point where three objects ar moving against each other.

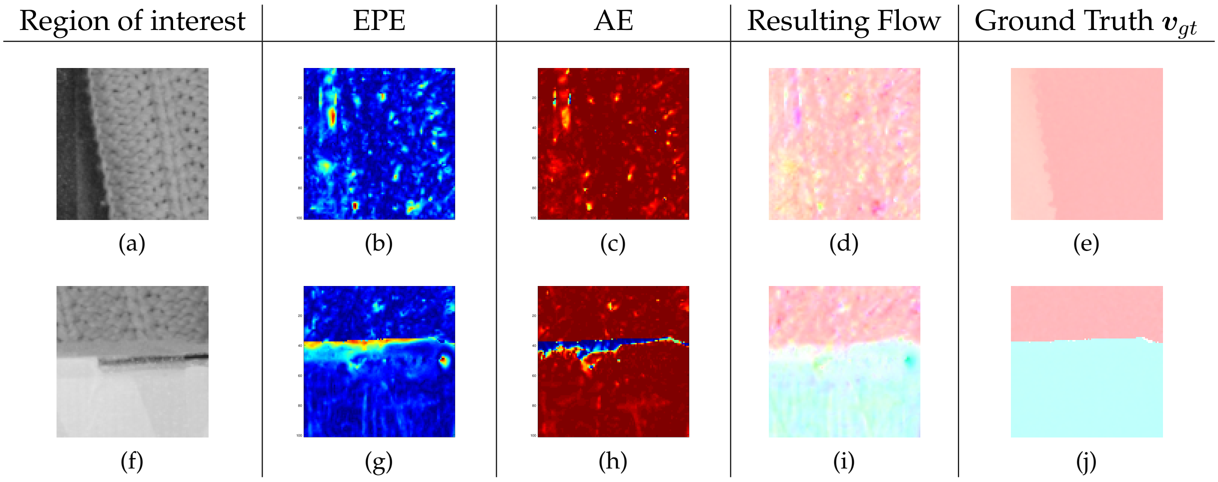


Figure 3.9: Error Analysis ST: This figure shows two examples of motion fields with linear boundaries. In figures 3.9d and 3.9i we can see that the resulting flow with texture inflicted from the data. Nevertheless the EPE values are nearly homogenous and small (see figures 3.9c and 3.9h)

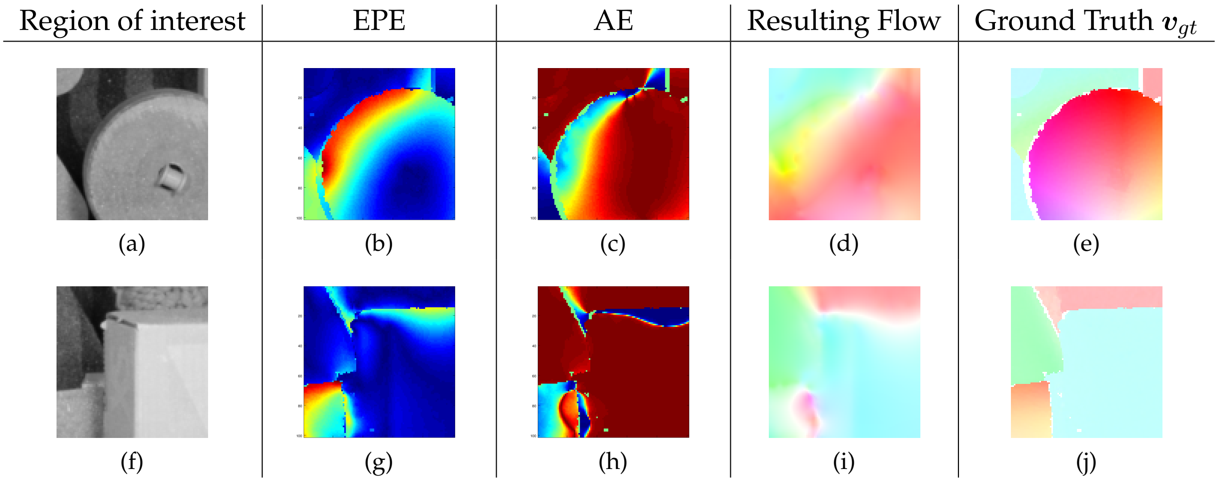


Figure 3.10: Error Analysis TV model: This figure shows two examples of motion field with nonlinear boundaries. In figure 3.10c we see that along the circumference of the wheel the EPE has the largest values and in figure 3.10h the is largest the junction point where three objects ar moving against each other.

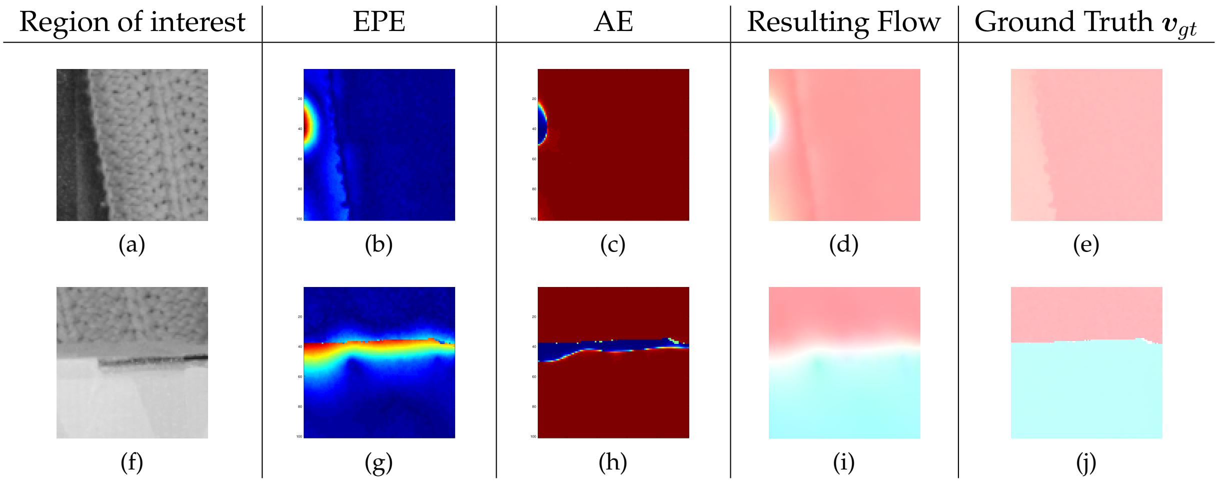


Figure 3.11: Error Analysis TV: This figure shows two examples of motion fields with linear boundaries. In figures 3.11d and 3.11i we can see that the resulting flow with texture inflicted from the data. Nevertheless the EPE values are nearly homogenous and small (see figures 3.11c and 3.11h)

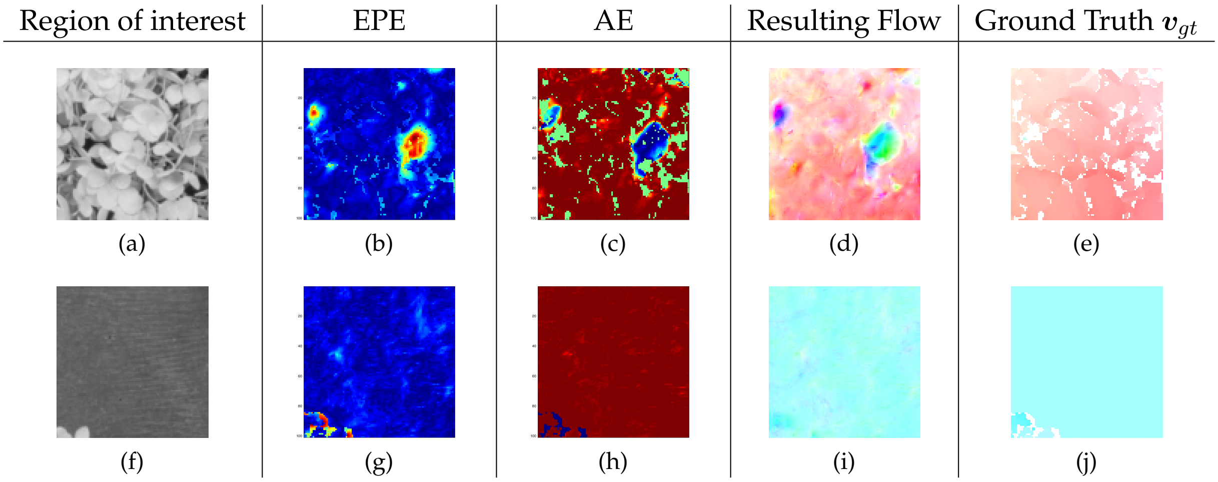


Figure 3.12: Error Analysis:
Second Column: Endpoint Error, Third Column: Angular Error.

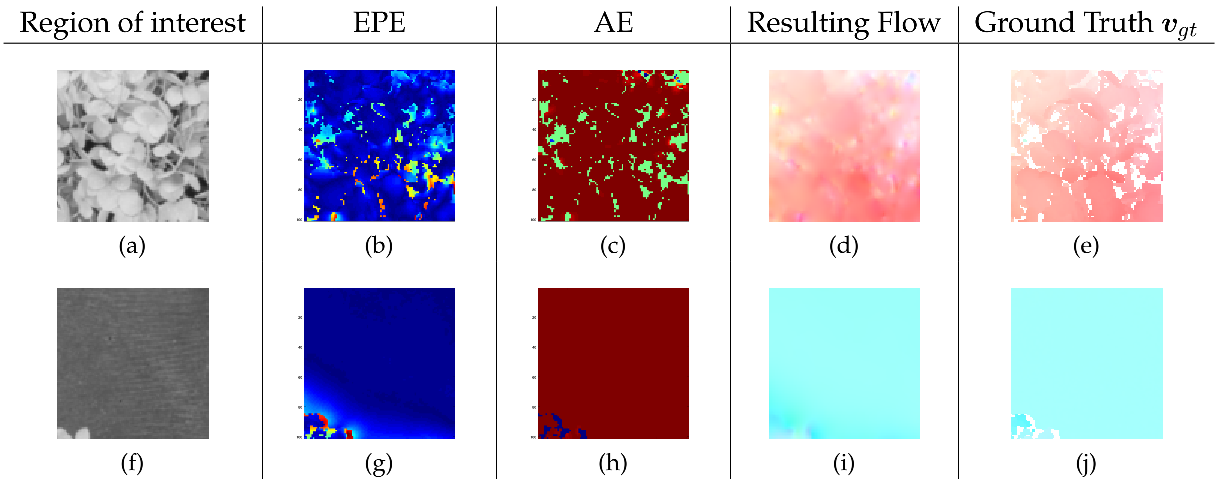


Figure 3.13: Error Analysis:
Second Column: Endpoint Error, Third Column: Angular Error.

3.9 Eigenvalue analysis and the stabilization parameter λ_2

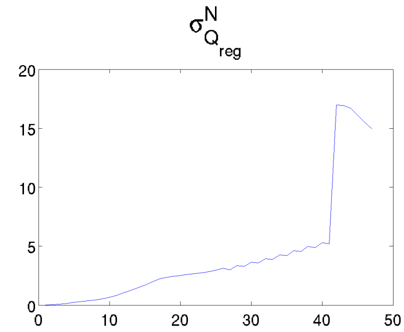


Figure 3.14: $\lambda_2 = 10^{-3}$

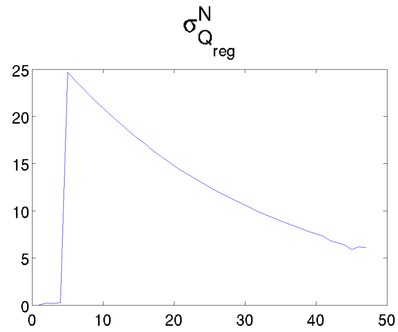


Figure 3.15: $\lambda_2 = 10^{-6}$

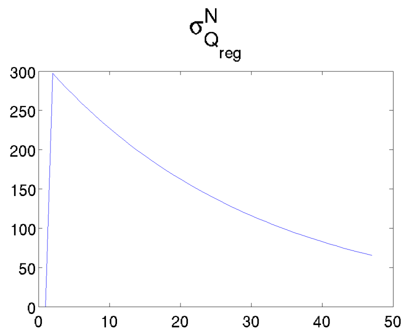


Figure 3.16: $\lambda_2 = 10^{-9}$

Figure 3.17: Analysis of the largest eigenvalue σ_Q^i of Q^{reg}

In chapter 3.3 we stated that the L_2 term in eq. (3.45) is needed to support the numerical stability of the model. We will back this statement now. Figures 3.14, 3.15 and 3.16 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. (3.45) governs the regularization. As the number of iterations increases the structure tensor determinant gets more weight, until the point where its influence over weighs 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 3.18, 3.19 and 3.20 show the vector $\boldsymbol{b}$, that is the Euler-Lagrange equation vector for different values of λ_2. Comparing the magnitude of $\boldsymbol{b}$ in Figures 3.18, 3.19 and 3.20 we see that for

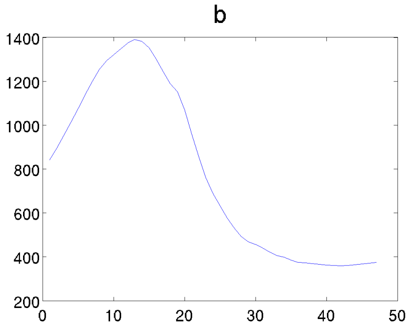


Figure 3.18: $\lambda_2 = 10^{-3}$

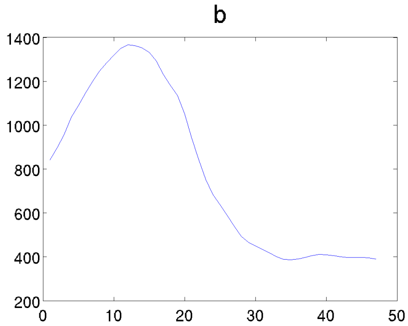


Figure 3.19: $\lambda_2 = 10^{-6}$

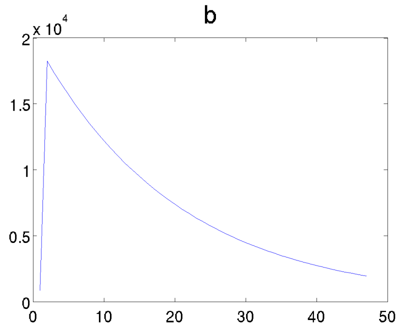


Figure 3.20: $\lambda_2 = 10^{-9}$

Figure 3.21: Analysis of the Euler-Lagrange vector $\boldsymbol{b}$

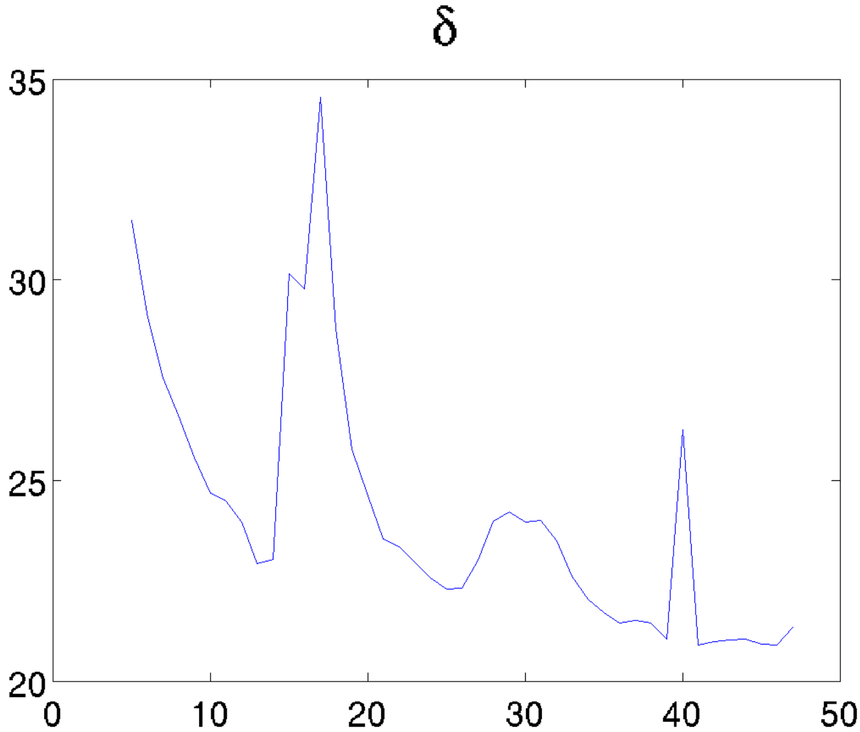
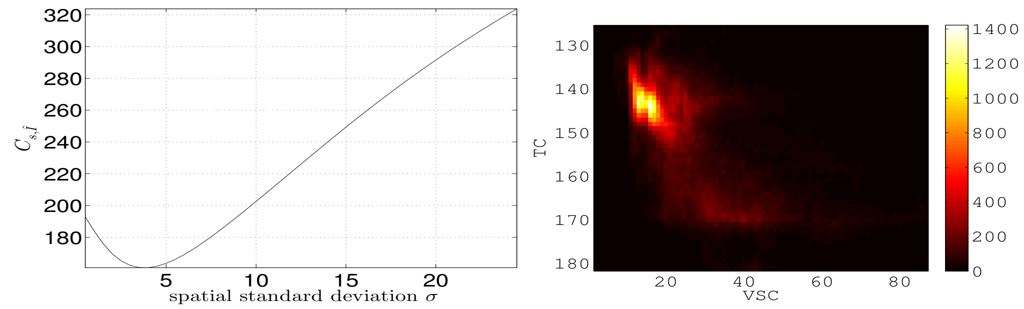


Figure 3.22: Analysis of the Euler-Lagrange vector δeq. (**??**)



(a) (b)

Figure 3.23: 3.23a: Dependence of $C_{\tilde{s}|\tilde{I}}$ on the scaling parameter σ. 3.23b: Joint Histogram $p\,(y, I)$ of the TC and smoothed VSC image pair y and $\tilde{I}$ at the optimum $\sigma^\star = 4$, the scale at which y and $\tilde{I}$ are maximally linear.

$\lambda_2 = 10^{-9}$ b is several orders of magnitude larger then 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 $\boldsymbol{b}$ is of the same order of magnitude as for $\lambda_2 = 10^{-3}$ but as we see in figure 3.15 the structure tensor only needs 4 iterations untils 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 figure **??** the update vector $\boldsymbol{\delta}$ gets unstable after 15 iterations.

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 emmisivity of carbon is approximately 0.98 (see [**?**]). It is in this case that in the amplitude image in figure **??**b the CFRP has a uniform amplitude. In the visual spectrum domain (figure **??**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. (2.131). The local conditional variance $C_{\tilde{s}|\tilde{I}}(\boldsymbol{x})$ in eq. (**??**) is a meassure 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 $C_{\tilde{s}|\tilde{I}}(\boldsymbol{x})$ has two parameters we need to estimate: the scale parameter σ from the likelihood in eq. (2.131) and the window size a of the window function ω. Since $C_{\tilde{s}|\tilde{I}}(\boldsymbol{x})$ is varies spacially we compute its median value $\hat{C}_{s|\tilde{I}}$. In figure **??** we have plotted for various window sizes a the median conditional variance $\hat{C}_{s|\tilde{I}}$ over the filter size σ. We can see that for window sizes $a \leq 23$ $\hat{C}_{s|\tilde{I}}$ has minima at $\sigma \approx 0$ while for larger window sizes $a \geq 31$ it tends to be minimal at filtersizes $\sigma > 6$. Figure **??** show their optimum $\sigma^\star$ plotted over the window size a. We see that window sizes $a < 21$ and $A > 31$ lead to unrealistic scale differences $\sigma^\star \approx 0$ and $\sigma^\star \geq 6$, 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 $C_{\tilde{s}|\tilde{I}}(\boldsymbol{x})$ is smaller compared to the case $a = 27$.

In figure **??** we show the resulting optical flow for different region of interests (roi). Figures 3.24a and 3.24f show the resulting optical flow $\boldsymbol{d}$ which match the corresponding VSC image I and TC image y in the table. Fow each roi we computed the joint histogram $p(y, I)$ (Figures 3.24b and 3.24g). In figure 3.24b $p(y, I)$ has two isolated maxima which is sufficient for for a linear relationship

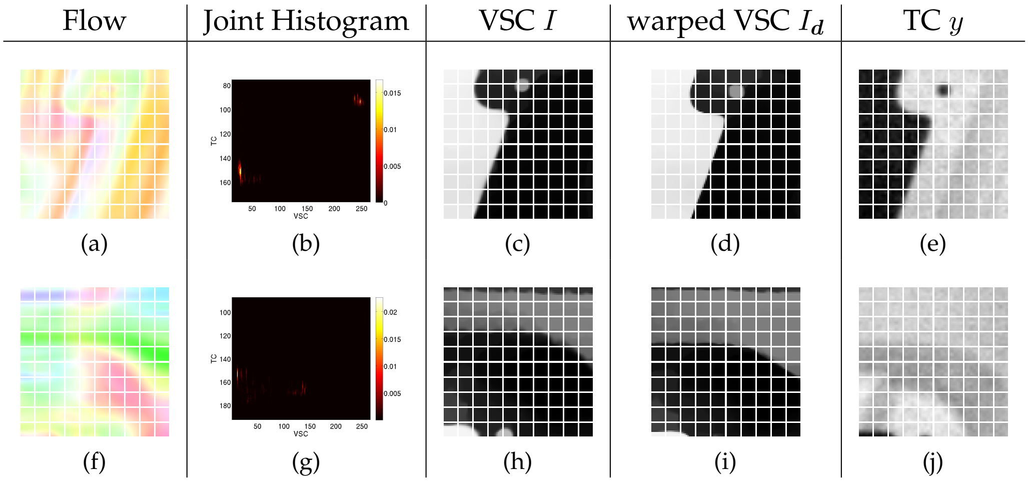


Figure 3.24: Multimodal Optical Flow: The resulting flow $\boldsymbol{d}$, VSC image I, the warped VSC $I_{\boldsymbol{d}}$, the TC image y as well as the joint histogram $p\,(y, I)$ are shown for different region of interests. We can observe that the boundaries of the flow are blurred. This comes from 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, $I_{\boldsymbol{d}}$ and y with 10 pixels per element to visually asses the quality of the flow. We can see the larger features are correctly matched, while smaller features are matched in a suboptimal fashion

??

between y and I. In figure 3.24g the linearity is obstructed to a minor degree since the TC image in figure 3.24j has a slight structural difference in the lower left corner compared to figure 3.24h.

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

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

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

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

$$\phi' = \phi^{\star} + \tau v^{\phi} \tag{4.3}$$

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

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

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

The vector V_e^{ϕ} is obtained from eqs. 2.52 and 2.53 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

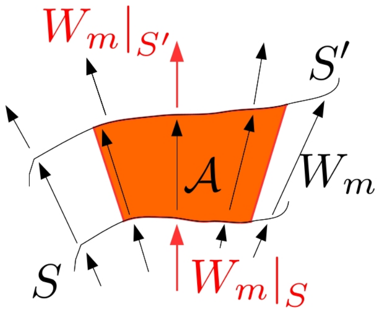


Figure 4.1: This figure shows a transformation of the level-set S to S' along the vector $\boldsymbol{W}_m(\boldsymbol{x})$. The region $\mathcal{A} \subset \Omega$ is the region a section of S traverses as it is shifted along $\boldsymbol{W}_m$ to the end position S'. If the divergence of $\boldsymbol{W}_m$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

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

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

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. (2.22))

$$\boldsymbol{x}' = \boldsymbol{x} + \tau^\Omega \boldsymbol{v}(\boldsymbol{x}) \tag{4.7}$$

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

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

The exact form of the vector $\boldsymbol{b}(\boldsymbol{x})$ will soon be deduced, now we wish to give an intuitive idea of $\boldsymbol{b}(\boldsymbol{x})$ 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 $\phi^\star$. 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. (2.79) simplifies to

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

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

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

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

$$\mathrm{div}\left(\boldsymbol{W}_m\right) = X_e^{m,\Omega}\left(\phi\right)\left[\mathcal{E}_{tot}\right] \tag{4.11}$$

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

$$\int_{\mathcal{A}} \mathrm{div}\left(\boldsymbol{W}_m\right) d^2 x = \int_{\partial\mathcal{A}} \boldsymbol{W}_m d\boldsymbol{S} \tag{4.12}$$

thus from eq. (4.11) we have

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

In figure 4.1 we have depicted the situation where a level-set S is shifted by the m-th one parameter Lie group g^{a_m} generating $\boldsymbol{W}_m$. $S' = g^{a_m} \circ S$ is the result of the shift and $\mathcal{A}$ is the region traversed by the shift of a section of S. The boundary $\partial\mathcal{A}$ consists of two lines tangential to $\boldsymbol{W}_m$ besides the sections of S and S'. Since the flux over the tangential lines vanishes we have

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

From Eq. (4.14) we see that the Euler-Lagrange differentials $[\mathcal{E}_{tot}]$ act as a *source* for transformations $\boldsymbol{b}_m^{\|}$ of the level-sets S *parallel* to $\boldsymbol{W}_m$. On the other hand transformations $\boldsymbol{b}_m^{\perp}$ which are perpendicular to $\boldsymbol{W}_m$ *must* change E^{prior} if they themselves are not symmetries of E^{prior}. Hence the vector $\boldsymbol{b}$ can only be a linear combination of vectors $\boldsymbol{b}_m^{\perp}$ which *perpendicular* to the $\boldsymbol{W}_m$

$$\boldsymbol{b}\left(\boldsymbol{x}\right) = \sum_{m=1}^{r} \alpha_m \boldsymbol{b}_m^{\perp}\left(\boldsymbol{x}\right), \quad \boldsymbol{b}_m^{\perp} \perp \boldsymbol{W}_m \tag{4.15}$$

Since $\boldsymbol{W}_m$ only depends on the prior E^{prior} (see eq. (2.84)) and not on the data term E^{data} the vector $\boldsymbol{b}$ in eq. (4.15) is guided only by the geometric constraints in the prior E^{prior}.

4.2 The Generalized Newton Algorithm

Our motivated procedure in eq. (**??**) depends on the assumption that from the r equations $\operatorname{div}(\boldsymbol{W}_m) = 0$ the vector $\boldsymbol{b}$ can be constructed from a basis of r vectors $\boldsymbol{b}_m$ which are perpendicular to the $\boldsymbol{W}_m$. To demonstrate this we remember how the energy E transforms under an arbitrary sub group $g_{\gamma(t)} \subset \mathbb{G}$ (eq. (2.79))

$$\frac{d}{dt} g_{\gamma_t} \circ E\Big|_{t=0} = \int_\Omega \left(\sum_i \left[V_e^\Omega, X_e^{\Omega,i} \right] (\phi) \cdot \boldsymbol{P}_i + v^\phi [\mathcal{E}_{tot}] \right) d^2x = 0 \tag{4.16}$$

$$= \int_\Omega \left(\sum_m \alpha_m \sum_i (\boldsymbol{M})_{mi} \cdot \boldsymbol{P}_i + v^\phi [\mathcal{E}_{tot}] \right) d^2x = 0 \tag{4.17}$$

$$(\boldsymbol{M})_{mi} = \left[X_e^{\Omega,m}, X_e^{\Omega,i} \right] (\phi), \quad \boldsymbol{P}_i (\phi(\boldsymbol{x})) = \frac{\delta \mathcal{E}_{prior} (\phi(\boldsymbol{x}))}{\delta \left(X_e^{\Omega,i} \phi \right)} \tag{4.18}$$

From eq. (4.17) we can deduce the exact form of the steepest descent updates in eq. (4.19) and eq. (**??**). For instance the Euler-Lagrange equations $[\mathcal{E}_{tot}]$ provide the gradient step in eq. (4.6). Nearly all methods for minimizing the energy E are of the basic Newton type in algorithm 2 which is based around the concept of steepest decent

$$\phi' = \phi - \tau^\phi \cdot [\mathcal{E}_{tot}] \tag{4.19}$$

The step-size parameter τ^ϕ is either chosen to be constant or, in more advanced algorithms (citation!!!) adjusted dynamically, for example on the conjugate gradients algorithm ([**?**]). The term $\boldsymbol{M} \cdot \boldsymbol{P}$ arrives from the spacial operator V_e^Ω thus it is the term we need to construct the vector $\boldsymbol{b}$ in eq. (4.8). Since the commutator matrix $\boldsymbol{M}$ in eq. (**??**) is an element of the Lie algebra $\mathcal{G}$ it can be represented in terms of the basis X_e^i

$$(\boldsymbol{M})_{mi} = \sum_l C_{mi}^l X_e^{\Omega,l} (\phi) \tag{4.20}$$

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

$$X_e^{\Omega,l} = \omega_l^\mu (\boldsymbol{x}) \, \partial_\mu \tag{4.21}$$

the product $\boldsymbol{M}\cdot\boldsymbol{P}$ in eq. (**??**) takes the form

$$
\begin{aligned}
\left(\boldsymbol{M}\cdot\boldsymbol{P}\right)_m =& B_m\phi\left(\boldsymbol{x}\right) & (4.22)\\
B_m =& b_m^\mu\left(\boldsymbol{x}\right)\partial_\mu, \quad b_m^\mu\left(\boldsymbol{x}\right) = \sum_i P_i(\phi\left(\boldsymbol{x}\right))C_{mi}^l\omega_l^\mu\left(\boldsymbol{x}\right) & (4.23)
\end{aligned}
$$

We would like to discuss the operator B_m. The invariance of E with respect to the group $\mathbb{G}$ requires that $\boldsymbol{M}\cdot\boldsymbol{P}$ vanishes. By eq. (4.22) this means the operator B_m must satisfy the level-set equation

$$
B_m\phi = b_m^\mu\partial_\mu\phi = 0 \tag{4.24}
$$

From eq. (4.24) we see that the vector $\boldsymbol{b}_m$ is tangential to the level-sets of $\phi(\boldsymbol{x})$. 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 $\boldsymbol{M}$ being left invariant and the canonical momentum $\boldsymbol{P}$, which is the functional derivative of $\mathcal{E}^{prior}$ and as such also left invariant. The consequence of the left invariance of B_m is that eq. (4.24) holds for any coordinate transformation of an arbitrarily chosen reference frame Ω_0

$$
B_m\phi\left(g\circ\boldsymbol{x}_0\right) = 0 \quad \forall\boldsymbol{x}_0, g \quad g\in\mathbb{G}, \quad \boldsymbol{x}_0\in\Omega_0 \tag{4.25}
$$

We interpret the operator B_m as the left-invariant vector-field for the one parameter Lie group $g_t^B\subset\mathbb{G}$. We use it to define a diffusion equation for the coordinate frame Ω

$$
\boldsymbol{x}\left(t\right) = g_t^B\circ\boldsymbol{x}, \quad \left.\frac{d\boldsymbol{x}\left(t\right)}{dt}\right|_{t=0} = \sum_{m=1}^r\alpha_m B_m\boldsymbol{x} = \boldsymbol{B}\cdot\boldsymbol{x} \tag{4.26}
$$

The operator $\boldsymbol{B}$ is a linear combination of the r operators B_m from eq. (4.23) and the coefficient vector $\boldsymbol{\alpha}$ comes from the decomposition of V_e in eq. (4.17). The diffusion process in eq. (4.26) is a non-linear process since the coefficient vector $\boldsymbol{b}_m(\boldsymbol{x})$ itself (eq. (4.23)) is a function of the coordinates $\boldsymbol{x} = g\circ\boldsymbol{x}_0$. It is guided along the components B_m which do not vanish due to trivial symmetries. To understand the dynamical properties of the diffusion equation in eq. (4.26) 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. (4.24) the first order derivative of the data term $\mathcal{E}^{data}$ under the action of g_t^B vanishes

$$
\frac{d}{dt}\left.\left(g_t^B\circ\mathcal{E}^{data}\left(\boldsymbol{x}\right)\right)\right|_{t=0} = \frac{\delta\mathcal{E}^{data}}{\delta\phi}\cdot\boldsymbol{B}\phi = 0 \tag{4.27}
$$

This means the diffusion process in eq. (4.26) is not impeded by the data term $\mathcal{E}^{data}$. The action of g_t^B on the prior $\mathcal{E}^{prior}$ allows for a geometrical explanation of the diffusion in eq. (4.26) proceeds. We let g_t^B act on $\mathcal{E}^{prior}$ and compute the derivative with respect to t

$$\frac{d}{dt}\left(g_t^B \circ \mathcal{E}^{prior}\right)\Big|_{t=0} = \sum_i P_i \cdot \left[\boldsymbol{B}, X_e^i\right] \phi \tag{4.28}$$

Eq. (4.28) is proportional to the angle between $\boldsymbol{P}$ and the commutator $\left[\boldsymbol{B}, X_e^i\right] \phi$ and allows for the following geometrical interpretation: this angle is a measure for the curvature. The commutator $\left[\boldsymbol{B}, X_e^i\right] \phi$ is the change the gradient $X_e^i \phi$ undergoes under the translation $\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{b}$. Under the diffusion process in eq. (4.26) minimizes $\mathcal{E}^{prior}$ and thus bends the level-sets of ϕ such that $\left[\boldsymbol{B}, X_e^i\right] \phi$ is at a right angle with $\boldsymbol{P}$.

Algorithm 2 Basic Newton Method

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

Algorithm 3 Generalized Newton Method

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

The step parameter τ^ϕ Explain $M \cdot P$ and $\boldsymbol{x} \to x + \alpha M \cdot P$, hessian of image-> curvature of level-set along $M \cdot P$ goes to zero If we set $v_\mu = 0$ then V_e represents the first variation

5 The Generalized Prior

In [24] a generalization of the structure tensor was introduced. The generalization is based on the introduction of the canonical coordinates $\xi(\boldsymbol{x})$ and $\eta(\boldsymbol{x})$ which pose a deformation of the Cartesian coordinate space Ω. The prime example is the transformation from Cartesian to polar coordinates $(x, y) \to (r, \theta)$. The gradient with respect the new coordinates can be expressed with the Cartesian coordinates via the Jacobian matrix J

$$\begin{pmatrix} \partial_\xi \\ \partial_\eta \end{pmatrix} = J^{-1} \cdot \begin{pmatrix} \partial_x \\ \partial_y \end{pmatrix} \quad J = \begin{pmatrix} \xi_x & \eta_x \\ \xi_y & \eta_y \end{pmatrix} \tag{5.1}$$

The differential operators $\{\partial_\xi, \partial_\eta\}$ also form the basis for the algebra $\mathcal{H}$ of the general Lie group $\mathbb{H}$, that is $[\partial_\xi, \partial_\eta] = 0$ if and only if the following conditions hold

$$\partial_x \xi = -\partial_y \eta \tag{5.2}$$
$$\partial_y \xi = \partial_x \eta \tag{5.3}$$

The eqs. eq. (5.3) are the famous Cauchy-Riemann differential equations and their combination give the separate wave equations

$$\Delta \xi = 0 \tag{5.4}$$
$$\Delta \eta = 0 \tag{5.5}$$

A solution of eq. (5.4) implies that there must also exist a solution for eq. (5.5). This is why one calls the pair $\{\xi, \eta\}$ a pair of *conjugate functions*. Within the coordinate frame (ξ, η) the operators $\{\partial_\xi, \partial_\eta\}$ obey the natural conditions

$$\begin{aligned} \partial_\xi \xi = 1, \quad \partial_\xi \eta = 0 \\ \partial_\eta \xi = 0, \quad \partial_\eta \eta = 1 \end{aligned} \tag{5.6}$$

The integral curves Γ^X generated by the operators $\{\partial_\xi, \partial_\eta\}$ can be written as an exponential Taylor series

$$\Gamma^X(s) = \exp(s \cdot X) \quad X = \omega_\xi \partial_\xi + \omega_\eta \partial_\eta \tag{5.7}$$

Level set functions ϕ satisfying

$$\frac{d}{ds}\phi\left(\Gamma^{X}(s)\right) = 0 \tag{5.8}$$

may exist if and only if ([24]) the exponential series in eq. (5.7) separates which according to the Baker-Hausdorff-Campbell formula is only the case when the operators $\{\partial_\xi, \partial_\eta\}$ commute

$$\exp(s \cdot X) = \exp(s \cdot \omega_\xi \partial_\xi) \cdot \exp(s \cdot \omega_\eta \partial_\eta) \Leftrightarrow [\partial_\xi, \partial_\eta] = 0 \tag{5.9}$$

On the other side if the coordinate functions (ξ, η) satisfy the Cauchy-Riemann equations (eqs. (**??**)) then one verifies that $[\partial_\xi, \partial_\eta] = 0$

The Group $O(2)$

We will now show an example on the group $O(2)$, the group of rotations and dilations. The transformation of the Cartesian coordinate system to the polar coordinates in 2 dimensions is given by the equations

$$x = r \cdot cos(\theta) \tag{5.10}$$

$$y = r \cdot sin(\theta) \tag{5.11}$$

Using the expression in eq. (5.1) the Jacobian J is easily calculated

$$J = \frac{1}{r}\begin{pmatrix} x & -y \\ y & x \end{pmatrix} \tag{5.12}$$

so that the derivative operators $\{\partial_\xi, \partial_\eta\}$ may be expressed in the Cartesian domain

$$\partial_\xi = x\partial_x + y\partial_y \quad \partial_\eta = -y\partial_x + x\partial_y \tag{5.13}$$

The coordinates ξ and η that satisfy eq. (5.6) with the derivative operators in eq. (5.13) are functions of the Cartesian coordinates

$$\xi(x, y) = \log(r) \quad r = \sqrt{x^2 + y^2} \tag{5.14}$$

$$\eta(x, y) = \arctan\left(\frac{y}{x}\right) \tag{5.15}$$

The function $\eta(x, y)$ is easily recognizable as the angle θ to the x-axis while the function $\xi(x, y)$ is not the radius r. This is because the operators $\{\partial_\xi, \partial_\eta\}$ must both have the same dimension, that means they must be invariant dilations $r \to \lambda \cdot r$.

The level sets of the algebra spun by the operators $\{\partial_\xi, \partial_\eta\}$ are linear with respect to the coordinates (ξ, η) by virtue of eq. (5.6). Thus by arguments similar those following eq. (**??**) , the authors in [24] introduced the *generalized structure tensor* (GST)

$$S_{\xi,\eta} = \int \begin{pmatrix} (\partial_\xi \phi)^2 & \partial_\xi \phi \partial_\eta \phi \\ \partial_\xi \phi \partial_\eta \phi & (\partial_\eta \phi)^2 \end{pmatrix} d\xi d\eta \tag{5.16}$$

As $S_{\xi,\eta}$ is a symmetric matrix there exists a decomposition

$$S_{\xi,\eta} = V^T D V \quad D = \begin{pmatrix} \lambda_\xi & 0 \\ 0 & \lambda_\eta \end{pmatrix} \quad V = (\boldsymbol{v}_1, \boldsymbol{v}_2) \tag{5.17}$$

The rotation matrix V acts in the (ξ, η) coordinate space. It does not necessarily correspond to rotations in the Cartesian coordinates (x, y). See [24] for a discussion on the steer-ability of the GST.

5.0.1 The Transformation Properties of the GST

Within the generalized coordinate frame (ξ, η) the action of the group $\mathbb{G}$ manifests itself as a translation

$$g_{\epsilon_1,\epsilon_2} \circ \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} \xi + \epsilon_1 \\ \eta + \epsilon_2 \end{pmatrix} \tag{5.18}$$

As mentioned before the basis operators $\{\partial_\xi, \partial_\eta\}$ and as a consequence all elements of the Lie algebra $\mathcal{G}$ *commute* with $\mathbb{G}$ (expand on left invariance!!)

$$g_{\epsilon_1,\epsilon_2} \circ X = X \circ g_{\epsilon_1,\epsilon_2} \forall X \in \mathcal{G} \tag{5.19}$$

Another important fact is that under the transformation in eq. (5.18) the volume element $d\xi d\eta$ is invariant. The consequence is that the GST in eq. (5.16) is invariant with respect to the generalized translation in eq. (5.18)

$$g_{\epsilon_1,\epsilon_2} \circ S_{\xi,\eta} = S_{\xi,\eta} \tag{5.20}$$

The GST has another interesting transformation property. As eq. (5.17) indicates, there exists an action of the rotation group $SO(2)$ on the generalized coordinate frame (ξ, η)

$$\xi' = \cos(\theta)\,\xi + \sin(\theta)\,\eta \tag{5.21}$$
$$\eta' = -\sin(\theta)\,\xi + \cos(\theta)\,\eta \tag{5.22}$$

The action of the transformation in eq. (5.22) yields a basis transformation of the Lie algebra $(\partial_\xi, \partial_\eta)$

$$\partial'_\xi = \cos(\theta)\,\partial_\xi + \sin(\theta)\,\partial_\eta \tag{5.23}$$
$$\partial'_\eta = -\sin(\theta)\,\partial_\xi + \cos(\theta)\,\partial_\eta \tag{5.24}$$

Under the change of basis in eq. (5.24) the GST transforms like a tensor

$$S'_{\xi,\eta} = R^T S_{\xi,\eta} R \quad R = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \tag{5.25}$$

It is important to note (see [24]) that the transformation in eq. (5.22) is not necessarily connected to the rotations on the Cartesian space Ω the functions $\xi(\boldsymbol{x})$ and $\eta(\boldsymbol{x})$ are embedded in. In fact it deforms the level sets corresponding to $(\partial_\xi, \partial_\eta)$ in a highly non linear manner.

5.1 Generalized Structure Tensor Based Prior

Our objective is to construct a prior $P(\nabla\phi)$ which is invariant to the transformations in eq. (5.25). We want to define a finite set of Lie groups $\mathbb{G}_i$ for which the classes of level sets $A_{\mathbb{G}_i}$ are minimizer sets for $P(\nabla\phi)$ (see eq. (**??**)). The methodology goes as follows: We define energy $E(\nabla\phi)$ (the negative log of $P(\nabla\phi)$) as a *product of the determinants* of the corresponding GSTs S_{ξ_i,η_i}

$$E(\nabla\phi) = \prod_i \mathrm{Det}(S_{\xi_i\eta_i}) \tag{5.26}$$

The energy in eq. (5.26) inherits the translation invariance (only when $\phi \in A_{\mathbb{G}_i}$??) of the GST in eq. (5.20), which we will show now. Due to the rotation invariance of the determinants in eq. (5.26) we can write the individual determinants in terms of their eigenvalues

$$\mathrm{Det}(S_{\xi_i\eta_i}) = \lambda_{\xi_i}\lambda_{\eta_i} \tag{5.27}$$

We can write the eigenvalues λ_{ξ_i} and λ_{η_i} as the squares of the orthogonal operators X_{ξ_i} and X_{η_i} which constitute a rotation of the basis $(\partial_{\xi_i}, \partial_{\eta_i})$

$$\text{Det}\left(S_{\xi_i\eta_i}\right) = \left(X_{\xi_i}\left(\phi\right)\right)^2 \left(X_{\eta_i}\left(\phi\right)\right)^2 \tag{5.28}$$

Under the adjoint action of the group $\mathbb{G}_i$ the operators X_{ξ_i} and X_{η_i} are invariant

$$\frac{d}{\epsilon_1}\left(g_{\epsilon_1,\epsilon_2} \circ X_\xi \circ g^{-1}_{\epsilon_1,\epsilon_2}\right)\bigg|_{\epsilon_1,\epsilon_2=0} = [X_\xi, \partial_\xi] = 0 \tag{5.29}$$

$$\frac{d}{\epsilon_2}\left(g_{\epsilon_1,\epsilon_2} \circ X_\xi \circ g^{-1}_{\epsilon_1,\epsilon_2}\right)\bigg|_{\epsilon_1,\epsilon_2=0} = [X_\xi, \partial_\eta] = 0 \tag{5.30}$$

eq. (5.30) also holds for X_{η_i}. It is evident that under the adjoint action the determinant $\text{Det}\left(S_{\xi_i\eta_i}\right)$ remains invariant. Since the determinant $\text{Det}\left(S_{\xi_i\eta_i}\right)$ vanishes when $\phi \in A_{\mathbb{G}_i}$ the whole energy in eq. (5.26) is invariant to any of the adjoint actions of the $\mathbb{G}_i$ if ϕ is locally in any of the sets $A_{\mathbb{G}_i}$

The open question which remains to be answered is, how does the energy eq. (5.26) transform when ϕ is not locally contained in any of the $A_{\mathbb{G}_i}$. The basis elements of two different $\mathcal{G}_i$ and $\mathcal{G}_j$ do not necessarily commute

$$\left[\partial_{\xi_i}, \partial_{\xi_j}\right] \neq 0 \tag{5.31}$$

$$\left[\partial_{\xi_i}, \partial_{\eta_j}\right] \neq 0 \tag{5.32}$$

$$\left[\partial_{\eta_i}, \partial_{\eta_j}\right] \neq 0 \tag{5.33}$$

$$\tag{5.34}$$

The question arises that when ϕ is locally within the vicinity of a particular $\phi_i^\star \in A_{\mathbb{G}_i}$, $\|\phi - \phi_i^\star\| \leq \delta$ will it be brought further away from $\phi_i^\star$ under the action of $\mathcal{G}_j$, that is $\|g_j \circ \phi - \phi_i^\star\| \leq \delta'$, $\delta' > \delta$? Or can the groups $\mathcal{G}_i$ and $\mathcal{G}_j$ be related to each other such that $\delta' < \delta$? To answer this we look at the product algebra $\mathcal{G}_i \times \mathcal{G}_j$ spun by the basis elements X_l. These operators may not be commutative but may be in involution

$$[X_i, X_j] = \sum_l C^l_{i,j} X_l \tag{5.35}$$

If this is the case then the commutator $[X_i, X_j]$ is also an element of the joint algebra $\mathcal{G}_i \times \mathcal{G}_j$. (if ϕ is roughly linear, then a dilation ∂_r will stretch the level sets to a line thus resulting in a level-set of the translation group)

5.1.1 Analysis of the Eigenvalues of the Rotation Dilation Group

We will now turn our focus on the eigenvalues of the Rotation Dilation GST. We use the polar coordinates from eq. (5.11). The integration window of the GST is

$$\xi_0 - \epsilon_\xi < \xi_0 < \xi_0 + \epsilon_\xi \tag{5.36}$$

$$\phi_0 - \epsilon_\phi < \phi_0 < \phi_0 + \epsilon_\phi \tag{5.37}$$

where $\xi_0 = \ln(r_0)$ so that eq. (5.37) translates to a region around the curvature radius r_0 and the angle ϕ_0. The level sets parameterized by the polar coordinates in the region in eq. (5.37) are the sectional curves of constant curvature $k_0^{-\epsilon_\xi} < k_0 < k_0^{\epsilon_\xi}$, $k_0 = \frac{1}{r_0}$. The Rotation Dilation GST from eq. (5.16) can be written in Cartesian coordinates

$$S_{\xi,\eta} = \int_{y_0-\epsilon}^{y_0+\epsilon} \int_{x_0-\epsilon}^{x_0+\epsilon} \begin{pmatrix} (\partial_\xi \phi)^2 & \partial_\xi \phi \partial_\eta \phi \\ \partial_\xi \phi \partial_\eta \phi & (\partial_\eta \phi)^2 \end{pmatrix} \frac{1}{r^2} M(x,y)\, dxdy \tag{5.38}$$

The mask $M(x,y)$ enforces the conditions eq. (5.37). The orientations of the level sets in this domain are fixed and thus This is why the minimum eigenvalue of the GST is zero only for level sets matching the curvature k_0 and the orientation ϕ_0. Since the integration space of $S_{\xi,\eta}$ is centered around ϕ_0, $S_{\xi,\eta}$ is not rotation invariant. In figure 5.1a the image function $I(x,y) = -r^2$ is depicted. The GST was evaluated for $r_0 = 30$ and $\phi_0 = \frac{\pi}{4}$ (Figures 5.1b to 5.1d) and $\phi_0 = \frac{\pi}{2}$ ((Figures 5.1e to 5.1g)). The eigenvalue corresponding to X_η, the derivative in angular direction is denoted by λ_2 (Figures 5.1d to 5.1g). It is observed that λ_2 has a minimum at the radius $r_0 = 30$ and the angles $\phi_0 = \frac{\pi}{4}$ (figure 5.1d) and $\phi_0 = \frac{\pi}{2}$ (figure 5.1g). As a result the determinant of the GST is only minimal at exactly those values (see figures 5.1b and 5.1e).

5.2 Image Denoising with the GST

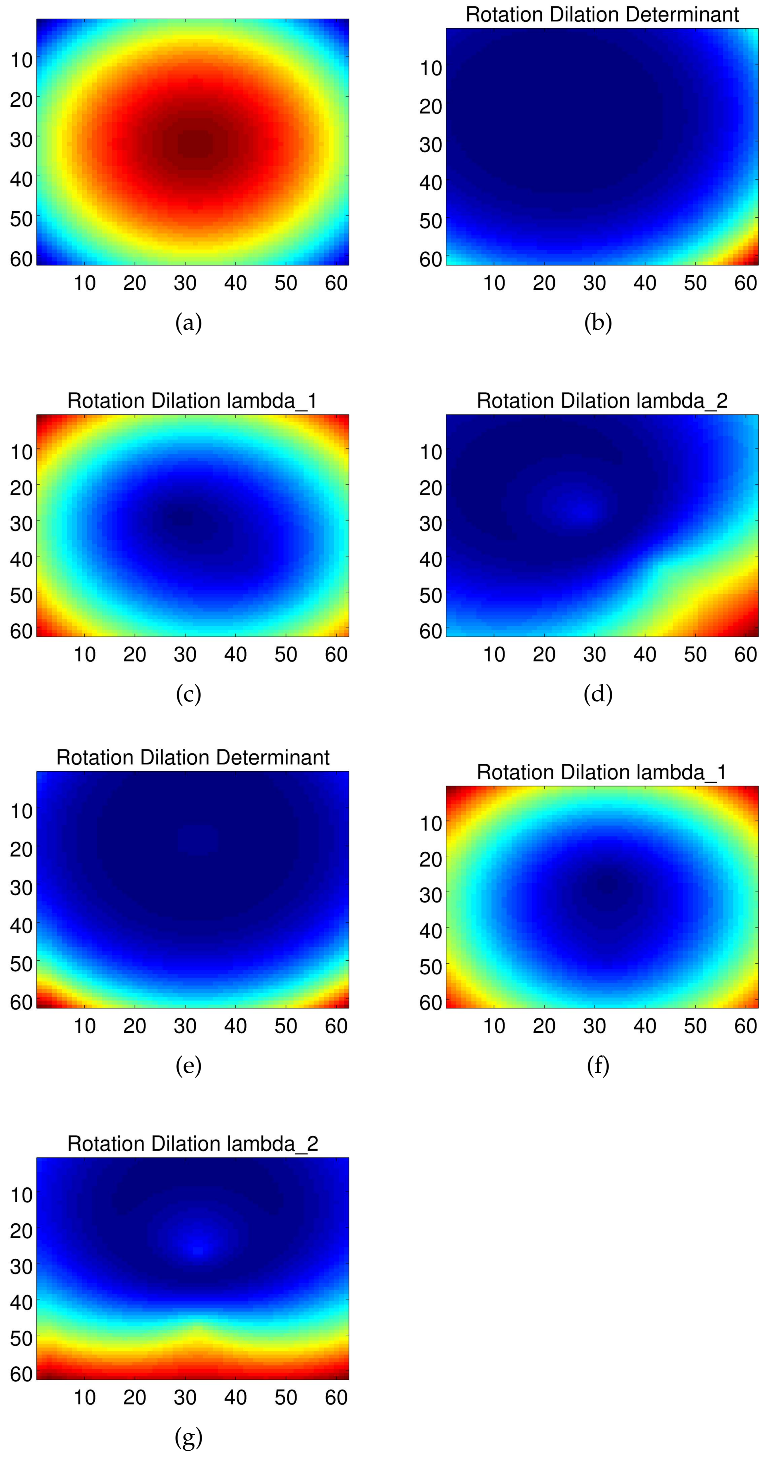


Figure 5.1: Figure 5.1a is the image $I(r) = -r^2$, figures 5.1b to 5.1d are the determinant, higher and lower eigenvalue of the GST for $r_0 = 30$ and $\phi_0 = \frac{\pi}{4}$. Figures 5.1e to 5.1g are same for $\phi_0 = \frac{\pi}{2}$.

6 Conclusions

Bibliography

[1] Paul Fieguth, *Statistical image processing and multidimensional modeling*, Springer Science & Business Media, 2010. 8

[2] Alexander A Kirillov, *An introduction to Lie groups and Lie algebras*, vol. 113, Cambridge University Press Cambridge, 2008. 15

[3] Elizabeth Louise Mansfield, *A practical guide to the invariant calculus*, vol. 26, Cambridge University Press, 2010. 15

[4] 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. 15

[5] E. Noether, "Invariante variationsprobleme," *Nachrichten Der Koeniglichen Gesellschaft Der Wissenschaften Zu Goettingen*, vol. Mathematische-Physikalische Klasse, pp. 235–257, 1918. 19

[6] E. Noether, "Invariant variation problems," *Transport Theory and Statistical Problem*, vol. 1 (3), pp. 183–207, 1971. 19

[7] B. K.P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, issues 1-3, pp. 185–203, 1981. 25, 30

[8] 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. 25

[9] 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. 25, 26

[10] 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. 25

[11] Leonid I Rudin, "Images, numerical analysis of singularities and shock filters," 1987. 25

[12] Kristian Bredies and Dirk Lorenz, "Mathematische bildverarbeitung," *Vieweg+ Teubner*, vol. 4, no. 6, pp. 12, 2011. 26

[13] 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. 29

[14] 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. 29

[15] 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. 30

[16] 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. 31

[17] 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, 31

[18] D. Mumford and J. Shah, "Optimal approximation by piecewise smooth functions and associated varia- tional problems," *Comm. Pure and Applied Mathematics*, vol. 42, pp. 577–685, 1989. 31

[19] Alexis Roche, Grégoire Malandain, Nicholas Ayache, and Xavier Pennec, "Multimodal image registration by maximization of the correlation ratio," 1998. 32, 33

[20] 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. 34

[21] 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. 35, 36, 40

[22] Y. Zhang, "Spatial resolution enhancement for hyperspectral image based on wavelet baysian fusion," *4th International Congress on Image and Signal Processing*, 2011. 36

[23] 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. 37

[24] J. Bigun, *Vision with Direction A Systematic Introduction to Image Processing and Computer Vision*, Springer, 2006. 69, 70, 71, 72