Optic Flow from Dynamic Anchor Point Attributes

a feasibility study

B.J. Janssen

Master of Science thesis

Project period: November 2003 – January 2005
Report Number: 05A/01
Commissioned by:

Supervisors:
Dr. L.M.J. Florack (Dep. of Biomedical Engineering, TU/e)
Prof.dr.ir. B.M. ter Haar Romeny (Dep. of Biomedical Engineering, TU/e)

Additional Commission members:
Prof.dr.ir. P.P.J. van den Bosch (Dep. of Electrical Engineering, TU/e)
Prof.dr. M. Nielsen (Dep. of Innovation, ITU)

The Department of Electrical Engineering of the Eindhoven University of Technology accepts no responsibility for the contents of M.Sc. theses or practical training reports
Abstract

Optic flow describes the apparent motion that is present in an image sequence. We show the feasibility of obtaining optic flow from dynamic properties of a sparse set of so called anchor points. Singular points of a Gaussian scale space image are identified as feasible anchor point candidates and analytical expressions describing their dynamic properties are presented. The advantage of approaching the optic flow estimation problem using these anchor points is that in these points the notorious aperture problem does not manifest itself.

The proposed optic flow estimation algorithm heavily depends on stationary reconstruction, which aims for a reconstruction of a static image from a set of features that is visually close to the image from which the features are extracted. Degrees of freedom that are not fixed by the constraints are disambiguated with the help of a so-called prior (i.e. a user defined model). A linear reconstruction framework is proposed that generalises a previously proposed scheme. As an example we propose a specific prior and apply it to the reconstruction from singular points. The reconstruction is visually more attractive and has a smaller $L_2$-error than the previously proposed linear methods.

The proposed optic flow estimation method succeeds in finding a dense vector field that approaches the optic flow field from a sparse set of inherent multi scale anchor points. As opposed to classical optic flow estimation schemes the proposed method accounts for an explicit scale component of the vector field, which could encode a hitherto unknown dynamic property.
Contents

1 Introduction 5

2 Anchor Points for Dynamic Analysis 7
   2.1 Singular Points .................................................. 7
   2.2 Flow ........................................................................ 8
   2.3 Error Measure for Anchor Point Localisation & Tracking ....... 10
   2.4 1D Analytical Example ............................................. 10
   2.5 Conclusions ............................................................ 12

3 Reconstruction of Stationary Images 13
   3.1 Introduction ............................................................ 13
   3.2 Theory .................................................................... 14
   3.3 Reconstruction from Singular Points ......................... 16
      3.3.1 Prior Selection .................................................. 16
      3.3.2 Implementation ................................................. 17
      3.3.3 Richer Features ................................................. 18
   3.4 Evaluation .............................................................. 18
      3.4.1 Qualitative Evaluation ....................................... 19
      3.4.2 Quantitative Evaluation .................................... 21
   3.5 Conclusions & Recommendations ............................. 21

4 Flow Reconstruction 23
   4.1 Proposed Vector Field Retrieval Method .................... 23
   4.2 Richer Features ...................................................... 24
   4.3 Evaluation ............................................................ 25
      4.3.1 Qualitative Evaluation ....................................... 28
      4.3.2 Quantitative Evaluation .................................... 29
   4.4 Conclusions & Recommendations ............................. 31

5 Conclusions & Recommendations 35
   5.1 Conclusions ............................................................ 35
   5.2 Recommendations ................................................... 36

A Alternative Approach to Theorem 33 37

B Parallel Mathematica Source Code 39
   B.1 Documentation of the makeParallelGramm Module ........ 39
   B.2 Source Code of the makeParallelGramm Module ............ 40
Chapter 1

Introduction

Large academic attention to the area of optic flow estimation can be attributed to the vast number of application areas of optic flow. Among these application areas one can find medical image registration, video processing, video coding and robotics.

Current variational optic flow estimation algorithms are based on the brightness constancy assumption that was initially proposed by Horn & Schunck [18]. This constraint is insufficient to determine optic flow unambiguously since constant brightness occurs on surfaces of codimension one (curves in 2D, surfaces in 3D, etc.). The intrinsic ambiguity has become known as the aperture problem. The ambiguity is typically resolved by adding an extra regularisation term to the optic flow constraint equation. Horn & Schunck used a quadratic regulariser. Ever since many alternative regularisation schemes have been proposed, essentially following the same rationale.

Next to variational methods [6, 29, 37] that are similar to the method proposed by Horn & Schunck, correlation-based [2, 33], frequency-based [17] and phase-based methods [9] were proposed. In order to cover large displacements several coarse-to-fine strategies of these techniques were proposed [35, 40]. Werkhoven et al., Florack et al. and Suinesiaputra et al. [13, 34, 38] developed biologically inspired variational optic flow estimation methods incorporation “optimal” local scale selection. The work by Florack et al. shows that taking notion of scale may lead to superior performance compared to the optic flow estimation algorithms evaluated by Barron et al. [4]. Recent work by Brox et al. [4] shows impressive results of current state of the art optic flow estimation techniques.

Despite these impressive results the aperture problem remains essentially unsolved. To shed new light on the aperture problem we investigate flow of so called anchor points in Gaussian scale space. These anchor points could be any type of isolated points in Gaussian scale space. In these points the aperture problem is nonexistent and therefore the flow can be unambiguously measured. If these points carry “sufficiently rich” dynamic information one may hypothesise that any high resolution dense optic flow field that is consistent with the constraints posed by the anchor points and their dynamic features will be a reasonable representative of the “underlying” optic flow field one aims to extract. This has the additional advantage that optic flow definition becomes independent of image resolution, as opposed
to “constant brightness” paradigms. The proposed optic flow estimation method which is depicted in Figure 1.1 is of a multiscale nature since each anchor point lives at a certain scale. This leads to a method in which automatic scale selection is manifest. Chapter 2 discusses the subject of anchor points.

Figure 1.1: An overview of the proposed optic flow estimation algorithm. First anchor points are selected. The second step is the flow estimation that is done on the sparse set of selected anchor points that are present in the scale space of a single frame. Finally this sparse multi-scale vector field is converted to a dense vector field at grid scale.

The sparse multi-scale vector field constructed by the motion of the anchor points has to be converted to a dense high resolution vector field at scale $s = s_0$, in which $s_0 > 0$ is related to grid scale. To this extent we generalise the reconstruction from information of singular points as proposed by Nielsen and Lillholm [30]. This generalisation of the stationary reconstruction problem is discussed in Chapter 3. A paper covering this subject has been accepted for oral presentation at the Scale Space 2005 conference [19].

In Chapter 4 the extension of the reconstruction algorithm to vector valued images and results of its application to flow reconstruction are discussed.
Chapter 2

Anchor Points for Dynamic Analysis

A Gaussian scale space representation \(u(x; s)\) in \(n\) spatial dimensions is obtained by convolution of a raw image \(f(x)\) with a normalised Gaussian:

\[
\begin{align*}
  u(x; s) &= (f \ast \varphi_s)(x), \\
  \varphi_s(x) &= \frac{1}{\sqrt{4\pi s^n}} e^{-\frac{|x|^2}{4s}}.
\end{align*}
\] (2.1)

We aim to identify and determine the velocity of anchor points in the scale space of an image sequence. The scale space of an image sequence is a concatenation of the scale spaces of the images of the sequence. This means time scale is not taken as an explicit dynamic parameter. The type of anchor point that is discussed in this thesis is the so called singular point. Next to these points any type of generically isolated point that is intrinsic in the scale space of an image could have been considered. An introduction to scale space theory can be found in a tutorial book by ter Haar Romeny [15] and a book by Lindeberg [27].

2.1 Singular Points

A singular point is a non-Morse critical point of a Gaussian scale space image. Scale \(s\) is taken as a control parameter. This type of point is also referred to in the literature as a degenerate spatial critical point or as a toppoint or catastrophe.

**Definition 1 (singular point).** A singular point \((x; s) \in \mathbb{R}^{n+1}\) is defined by the following equations.

\[
\begin{align*}
  \nabla u(x; s) &= 0, \\
  \det \nabla^T u(x; s) &= 0.
\end{align*}
\] (2.2)

Here \(\nabla\) denotes the spatial gradient operator.

It is clear that these points in scale space are truly singular. The behavior near singular points is the subject of catastrophe theory. Damon studied the applicability of established
catastrophe theory in a scale space context [7]. Florack and Kuijper have given an overview of the established theory in their paper about the topological structure of scale space images for the generic case of interest, and investigated geometrical aspects of generic singularities [11]. More on catastrophe theory in general can be found in a monograph by Gilmore [14].

Singular points can be found by following critical paths, which are curves of vanishing gradient. A generic singular point manifests itself either as a creation or annihilation event. In most cases critical paths connect to the ground plane, \( s = 0 \), but not always as is shown in Figure 2.1. In the \( n = 1 \) dimensional case only annihilation events are generic.

![Figure 2.1: A creation-annihilation loop. The points denote singular points and the lines show the critical paths.](image)

2.2 Flow

When tracking anchor points over time we have one extra state variable. Spatial non-Morse critical points follow a path through scale spacetime in the scale space of an image sequence. Such a path can be described by a parameterised curve \( p(t) \) as long as it is transversal to time frames. Along this curve, of which a visualisation can be found in Figure 2.2, the properties of the point as described in equation (2.2) do not change.

**Definition 2 (Flow Vector).** The motion of a singular point is described by a flow vector \( \mathbf{v} \in \mathbb{R}^{n+1} \),

\[
\mathbf{v} = \begin{bmatrix} \dot{x} \\ \dot{s} \end{bmatrix}.
\]

Here \( \dot{x} \in \mathbb{R}^n \) denotes a spatial vector describing the flow’s spatial components and \( \dot{s} \in \mathbb{R} \)
2.2 Flow

Figure 2.2: The path \( p(t) \) that a singular point travels through time. In the frames the zero crossings of the image gradient are depicted as a function of scale \( s \). These lines are called critical paths.

describes the scale component. The time component\(^1\) of this vector is implicit and taken as 1. In other words it is assumed that time \( t \) is a valid parameter of the flow vector’s integral curve. A dot is shorthand for \( \frac{\partial}{\partial t} \).

**Theorem 1.** The velocity of a singular point at time \( t \) is given by

\[
\begin{bmatrix}
\dot{x} \\
\dot{s}
\end{bmatrix} = - \begin{bmatrix}
H(t; x, s) & \nabla \frac{\partial u(t; x, s)}{\partial s} \\
\nabla^T \det H(t; x, s) & \ldots & \frac{\partial \det H(t; x, s)}{\partial s}
\end{bmatrix}^{-1} \begin{bmatrix}
\nabla \frac{\partial u(t; x, s)}{\partial t} \\
\frac{\partial \det H(t; x, s)}{\partial t}
\end{bmatrix}.
\tag{2.4}
\]

\( H(t; x, s) \) is the \( n \times n \) matrix with components \( \frac{\partial^2 u(t; x, s)}{\partial x^a \partial x^b} \) where \( a \) and \( b \) index the spatial dimensions.

**Proof.** We ignore boundaries of the image sequence. By definition of scale space and the fact that we are studying natural image sequences it can be assumed that the studied function is at least piecewise continuously differentiable (even analytical with respect to spatial variables). By definition the properties of the tracked point along the curve,

\[
\begin{bmatrix}
\nabla u(t; x, s) \\
\det H(t; x, s)
\end{bmatrix} = 0,
\tag{2.5}
\]

do not change. If the Jacobian determinant does not degenerate,

\[
\det \begin{bmatrix}
H(t; x, s) & \nabla \frac{\partial u(t; x, s)}{\partial s} \\
\nabla^T \det H(t; x, s) & \ldots & \frac{\partial \det H(t; x, s)}{\partial s}
\end{bmatrix} \neq 0,
\tag{2.6}
\]

the implicit function theorem \([1]\) can be applied to obtain the flow along the parameterised curve \( (x(t), s(t)) \) (see Figure 2.2) through time. Assuming that equation (2.6) holds, due to the implicit function theorem, the movement of the anchor point in a sufficiently small neighborhood of the initial position is given by inversion of

\[
\begin{bmatrix}
H(t; x, s) & \nabla \frac{\partial u(t; x, s)}{\partial s} \\
\nabla^T \det H(t; x, s) & \ldots & \frac{\partial \det H(t; x, s)}{\partial s}
\end{bmatrix} \begin{bmatrix}
\dot{x} \\
\dot{s}
\end{bmatrix} = - \begin{bmatrix}
\nabla \frac{\partial u(t; x, s)}{\partial t} \\
\frac{\partial \det H(t; x, s)}{\partial t}
\end{bmatrix}.
\tag{2.7}
\]

\( \frac{\partial}{\partial t} \)\(^1\)This assumption is called the “temporal gauge” and reflects the transversality assumption.
This follows straightforward by setting
\[
\frac{d}{dt} \begin{bmatrix} \nabla u(t; x, s) \\ \det H(t; x, s) \end{bmatrix} = 0 ,
\]
in which \( \frac{d}{dt} \) denotes the total time derivative along the flow.

Problems arise when the point that is followed disappears. At such an event the point in scale spacetime changes to a non-Morse critical point and the implicit function theorem does not hold anymore. This can be detected by checking if the Jacobian determinant degenerates. As such this can be used to evaluate the validity of the estimated flow vector.

### 2.3 Error Measure for Anchor Point Localisation & Tracking

Because of noise the resulting vector of equation (2.7) will not point to the exact position of the anchor point in the next frame. A refinement of the solution can be made by using a Taylor expansion around the estimated position of the anchor point as proposed by Florack and Kuijper [11]. Evaluation of
\[
\begin{bmatrix} x' \\ s' \end{bmatrix} = \begin{bmatrix} - \det H(t; x, s) H^{-1}(t; x, s) \frac{\partial u(x,s)}{\partial s} \\ \det H(t; x, s) \end{bmatrix}
\]
near a singular point result in a vector pointing to the actual position of the singular point. The prime in \( x' \) and \( s' \) denotes differentiation with respect to a path parameter \( p \) implicitly defined by \( \frac{ds}{dp} = \det H(t; x, s) \). For a detailed derivation of equation (2.9) we refer to the original article. Notice that the so-called “cofactor matrix”
\[
\tilde{H}(t; x, s) \overset{\text{def}}{=} \det H(t; x, s) H^{-1}(t; x, s)
\]
is well defined even in the limit where \( \det H(t; x, s) \to 0 \). This gives us an error measure that can be incorporated in tracking mechanisms. Note that one can march this vector field in order to obtain a higher accuracy.

### 2.4 1D Analytical Example

Consider a 1D signal that propagates through time \( \Psi(x, t) = e^{-2(x-1-2t)^2} + 2e^{-2(x-3-t)^2} \). This signal is depicted in Figure 2.3a. A Gaussian blurred version of this signal,
\[
\Psi(x, t, \sigma) = \frac{1}{\sqrt{1 + 8\sigma}} \left( e^{-\frac{2(-1-2t-x)^2}{1+8\sigma}} + 2e^{-\frac{2(3+t-x)^2}{1+8\sigma}} \right),
\]
can be analysed to find its spatial critical point. One can follow the zero-crossings of the first derivative of the signal through scale, these paths are called the critical paths of the signal. The critical paths of equation (2.11) are depicted in Figure 2.3b. The critical scale can be calculated by following the critical paths and find out if the Hessian matrix \((u_{xx} \text{ in this case}) \)
To obtain the velocity of the singular point that resides at \((x, \sigma) = (1.35146, 0.164695)\) at time \(t = 0\) the following equation has to be solved (recall equation (2.7)):

\[
\begin{pmatrix}
  u_{xx} & u_{xxx} \\
  u_{xxx} & u_{xxxx}
\end{pmatrix}
\begin{pmatrix}
  \dot{x} \\
  \dot{\sigma}
\end{pmatrix}
= -
\begin{pmatrix}
  u_{xt} \\
  u_{xxt}
\end{pmatrix}.
\] (2.12)

The solution of equation (2.12) can be determined by taking the semi-inverse (note that, usually, plain inversion is not stable enough) of the matrix appearing in the equation. After finding the estimated position of the spatial critical point at the next point in time the actual position can be found by using the same technique as demonstrated in the beginning of this example. The result of the estimation and the actual position for a “time-step” of \(\Delta t = 0.1\) is shown in Figure 2.5. The position of point \(b\) is, because of the fact that it is impossible to find the position of the spatial critical point analytically, also an estimated position (like the position of the initial point that is labelled \(a\)). The difference between point \(b\) and point \(c\) is small compared to its velocity times time-step and therefore we can conclude that, in this case, our scheme works.

Similar results can be obtained with a 1D signal in which smoothness in the time direction is enforced by means of a blur scale \(\tau\) in the time direction. Using this time scale \(\tau\) the Gaussian
blurred signal is defined by

$$
\Psi(x, t, s, \tau) = e^{-\frac{2(-1-2t+x)^2}{1+8\tau+16s^2}} + \frac{2(-3-t+x)^2}{2\sqrt{1+8\tau+4s^2}}. \tag{2.13}
$$

At a fixed time-scale $\tau$ the spatial critical point of the signal can be found in the same manner as demonstrated in the previous example. The movement of the point’s position by slightly changing the time-scale can be found by applying equation (2.9).

### 2.5 Conclusions

Singular points are natural candidates as anchor points for dynamic analysis. Their dynamical properties can be described analytically and their localisation error can be reduced with the help of a scale space measure. An example has been given showing the feasibility of anchor point tracking in 1D.
Chapter 3

Reconstruction of Stationary Images

In order to obtain a dense flow field from singular points endowed with their dynamic attributes a so called dense flux field is generated that is consistent with these features. In this chapter we will study the essence of this so-called “reconstruction” algorithm in the simplified context of stationary scalar images. Optic flow will be discussed in the next chapter.

3.1 Introduction

We describe a general method for reconstruction from scale space interest points and their differential attributes. Using the reconstruction the information content of these points can be investigated (Nielsen and Lillholm [30]).

Lillholm, Nielsen and Griffin [26, 30] have put emphasis on a “sparse” constraint set and the role of different priors. In general their priors are not given in terms of an inner product. The disadvantage of their approach is that the reconstruction algorithm is not linear and therefore slow and somewhat cumbersome to implement. Kanters et al. [23] use the assumption of a “sufficiently rich” set of constraints. The role of the prior is less significant so they chose for a standard L2-norm. We shall refer to this as the standard linear reconstruction scheme. Advantages of his approach are that the reconstruction algorithm is linear and analytical results can be found. The disadvantage is that if the set of constraints is not sufficiently rich then this method is qualitatively outperformed by nonlinear reconstruction.

We propose a general reconstruction framework which can be applied to a large set of priors. Any prior that can be described by a norm formed by an inner product can be mapped to this framework. Our method overcomes the disadvantages of the standard linear reconstruction scheme [23] while retaining linearity. This is done by replacing the L2-inner product by an inner product of Sobolev type. To verify the proposed method we apply it to the reconstruction from singular points. A prior that smoothes the reconstructed image while not violating the constraints is selected. This aims for a reconstruction that has as few additional singular
points as possible under the constraints. Also the features are enriched by taking higher order derivatives into account.

3.2 Theory

Definition 3. The $L_2$-inner product for $f, g \in L_2(\mathbb{R}^2)$ is given by

$$ (f, g)_{L_2} = \int_{\mathbb{R}^2} f(x) g(x) dx . $$ (3.1)

This is the standard inner product used in previous work [23, 26, 30].

The reconstruction problem boils down to the selection of a representative of the metameric class consisting of $g \in L_2(\mathbb{R}^2)$ such that

$$ (\psi_i, g)_{L_2} = c_i , \quad (i = 1 \ldots N) $$ (3.2)

with $\psi_i$ denoting the distinct localised filters that generate the $i^{th}$ filter response $c_i = (\psi_i, f)_{L_2}$. For an alternative description of this class see appendix A.

The selection of $g$ is done by minimising a prior subject to the constraints of equation (3.2). A distinction can be made between priors (global constraints) that are constructed by a norm formed by an inner product and those that are constructed by a norm that is not formed by an inner product. In the former case it is possible to translate the reconstruction problem to a linear projection. This maps the reconstruction problem onto straightforward linear algebra. To this end we propose a generalisation of Definition 3 as follows.

Definition 4 (A-Inner Product). Let $A \in B(L_2(\mathbb{R}^2))$, i.e. a continuous linear operator on $L_2(\mathbb{R}^2)$. Then

$$ (f, g)_A = (f, g)_{L_2} + (Af, Ag)_{L_2} . $$ (3.3)

Note that we may write

$$ (f, g)_A = \left( f, (I + A^t A) g \right)_{L_2} . $$ (3.4)

For an image $f \in L_2(\mathbb{R}^2)$ we consider a collection of filters $\psi_i \in L_2(\mathbb{R}^2)$ and filter responses $c_i, i = 1, \ldots, N$, given by

$$ c_i = (\psi_i, f)_{L_2} . $$ (3.5)

Thus the a priori known features are given in terms of an $L_2$-inner product. In order to express these features relative to the new inner product we seek an effective filter, $\kappa_i$ say, such that

$$ (\kappa_i, f)_A = (\psi_i, f)_{L_2} $$ (3.6)

for all $f$. We will henceforth refer to $\psi_i$ as an “$L_2$-filter” and to $\kappa_i$ as its corresponding “$A$-filter”.

Lemma 2 (A-Filters). Given \( \psi_i \in L^2(\mathbb{R}^2) \), its corresponding A-filter is given by

\[
\kappa_i = (I + A^\dagger A)^{-1}\psi_i .
\]  

(3.7)

Proof. Applying Definition 4 using the fact that \( (I + A^\dagger A) \) is self adjoint,

\[
(k_i, f)_A = \left( (I + A^\dagger A)(I + A^\dagger A)^{-1}\psi_i, f \right)_{L_2} = (\psi_i, f)_{L_2} .
\]  

(3.8)

We aim to establish a reconstruction \( g \) that satisfies equation (3.2) and simultaneously minimises

\[
E(g) = \frac{1}{2} (g, g)_A .
\]  

(3.9)

Since \( g \) satisfies equation (3.2) we may as well write

\[
E(g) = \frac{1}{2} ((g, g)_A + (Ag, Ag)_{L_2}) - \lambda^i ((\psi_i, g)_{L_2} - c_i) .
\]  

(3.10)

in other words

\[
E(g) = \frac{1}{2} ( (g, g)_A + (Ag, Ag)_{L_2} ) - \lambda^i ((\psi_i, g)_{L_2} - c_i) .
\]  

(3.11)

Summation convention applies to upper and lower feature indices \( i = 1\ldots N \). The first term in equation (3.10) is referred to as the prior. The remainder consists of a linear combination of constraints, equation (3.2), with Lagrange multipliers \( \lambda^i \).

Theorem 3. The solution to the Euler-Lagrange equations for equation (3.10) can be found by A-orthogonal projection of the original image \( f \) on the linear space \( V \) spanned by the filters \( \kappa_i \), i.e.

\[
g = P_V f = (\kappa^i, f)_A \kappa_i .
\]  

(3.12)

Here we have defined \( \kappa^i \defeq G_{ij} \kappa_j \) with Gramm matrix \( G_{ij} = (\kappa_i, \kappa_j)_A \) and \( G^{ik} G_{kj} = \delta^i_j \).

Proof. The functional derivative of equation (3.10) with respect to the image \( g \) is given by

\[
\frac{\delta E(g)}{\delta g} = (I + A^\dagger A)g - \lambda^i \psi_i .
\]  

(3.13)

The solution to the corresponding Euler-Lagrange equations is formally given by

\[
g = \lambda^i (I + A^\dagger A)^{-1}\psi_i = \lambda^i \kappa_i .
\]  

(3.14)

So the filter responses can be expressed as

\[
c_i = (\psi_i, g)_{L_2} = \lambda^i (\psi_i, \kappa_j)_{L_2} = \lambda^i (\kappa_i, \kappa_j)_A .
\]  

(3.15)

Consequently \( \lambda^i = G^{ij} c_j \). Applying this to equation (3.14) leads to

\[
g = \lambda^i \kappa_i = G^{ij} c_j \kappa_i = G^{ij} (\kappa_j, f)_A \kappa_i = (\kappa^i, f)_A \kappa_i .
\]  

(3.16)

This completes the proof of Theorem 3.

Theorem 3 is written in a Euler-Lagrange formalism to comply with previous work on this subject [23, 26, 30]. The authors do notice the linear reconstruction problem can be approached in a simpler and more elegant way. This approach is sketched in appendix A.
3.3 Reconstruction from Singular Points

The theory of the previous section is applicable to any set of linear features. Here we are particularly interested in feature attributes of so-called singular points in Gaussian scale space. For the remainder of this thesis we use the following convention for the continuous Fourier Transform

\[
\mathcal{F}(f)(\omega) = \hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int e^{-i\omega x} f(x) dx
\]

\[
\mathcal{F}^{-1}(f)(x) = f(x) = \frac{1}{\sqrt{2\pi}} \int e^{i\omega x} \hat{f}(\omega) d\omega.
\]  

Notice that with this definition Fourier transformation becomes a unitary transformation.

3.3.1 Prior Selection

Johansen showed \cite{20, 21} that a one dimensional signal is defined up to a multiplicative constant by its singular points. This is probably not the case for two dimensional signals (images). It was conjectured that these points endowed with suitable attributes do contain enough information to be able to obtain a reconstruction that is visually close to the initial image \cite{23, 26, 30}.

Figure 3.1: The image on the right hand side shows the standard linear reconstruction, taking up to second order differential structure into account, as proposed by Kanters et al. \cite{23}, from 63 singular points of “Lena’s eye”. The original image, from which the singular points are taken is shown on the left hand side.

As can be seen in Figure 3.1 the standard linear reconstruction proposed by Kanters et al. \cite{23}, which is based on the standard $L_2$-inner product, is far from optimal. The problem can be identified by determining the number of additional singular points that appear in the reconstructed image while strictly insisting on the features to hold. In case of a perfect reconstruction the number of singular points would be equal for the reconstructed and original image. In the case of Figure 3.1 the actual number of singular points is higher, causing spurious structure. The number of singular points in the reconstructed image can be reduced by
3.3. Reconstruction from Singular Points

smoothing the image. Therefore a prior derived from the following inner product is proposed:

\[(f, g)_A = (f, g)_{L^2} + (-\gamma \sqrt{-\Delta f}, -\gamma \sqrt{-\Delta g})_{L^2} = (f, g)_{L^2} - (f, \gamma^2 \Delta g)_{L^2}\]

This prior introduces a smoothness constraint to the reconstruction problem. The degree of smoothness is controlled by the parameter \(\gamma\). When \(\gamma\) vanishes the projection equals the one from standard linear reconstruction \[23\]. Note that this is a standard prior in first order Tikhonov regularisation \[10, 36\].

3.3.2 Implementation

Using the inner product of equation (3.18) the A-filter equals,

\[\kappa_i = (I - \gamma^2 \Delta)^{-1} \psi_i = \mathcal{F}^{-1} \left( \frac{1}{1 + \gamma^2 ||\omega||^2} \mathcal{F}(\psi_i)(\omega) \right)\]

The filter shape in the spatial domain is somewhat harder to obtain. For two dimensions \((n = 2)\) the convolution filter that represents the linear operator \((I - \gamma^2 \Delta)^{-1}\) equals

\[\phi_\gamma(x, y) = \frac{1}{2\pi \gamma^2} K_0\left(\frac{\sqrt{x^2 + y^2}}{\gamma}\right)\]

with \(K_0\) representing the zeroth order modified Bessel function of the second kind. This was also noted by Florack, Duits and Bierkens \[10\] who worked on Tikhonov regularization and its relation to Gaussian scale space. The singularity of \(\phi_\gamma(x)\) in the origin gives rise to numerical problems. The Fourier representation of the operator is sampled and after that a discrete inverse Fourier transform is applied to it.

The calculation of the Gramm matrix \(G_{ij}\) is the computationally hardest part of the reconstruction algorithm. An analytic expression for this matrix is not available (unless \(\gamma = 0\)). Therefore the inner products \((\kappa_i, \kappa_j)_A\) have to be found by numerical integration. By the Parseval theorem we have (recall equations (3.19) and (3.20))

\[(\kappa_i, \kappa_j)_A = \left( \frac{1}{1 + \gamma^2 ||\omega||^2} \hat{\psi}_i, \hat{\psi}_j \right)_{L^2} = \left( \frac{1}{1 + \gamma^2 ||\omega||^2} \hat{\psi}_i \ast \hat{\psi}_j \right)_{L^2} = (\phi_\gamma, \psi_i \ast \psi_j)_{L^2}\]

At this point we have not yet specified the \(\psi_i\) filters. Since we are interested in the properties of singular points in Gaussian scale space we define the filters as follows.

**Definition 5** \((\psi_i)\). A filter \(\psi_i\) is a localised derivative of the Gaussian kernel, recall equation (2.1), at a certain scale. Given \(x, y, \xi, \eta \in \mathbb{R}\) and \(m, n \in \mathbb{N}_0\)

\[\psi_i(x) \overset{\text{def}}{=} \varphi_{m,n}^{\sigma,\xi,\eta}(x, y) = \frac{\partial^{m+n}}{\partial x^m \partial y^n} \varphi_{\sigma}(x, y) \bigg|_{x \rightarrow x - \xi, y \rightarrow y - \eta} \]

with \(i \overset{\text{def}}{=} (m, n, \xi, \eta, \sigma) \in \mathbb{N}_0^2 \times \mathbb{R}^2 \times \mathbb{R}_+\). It is understood that \(s = \frac{1}{2\sigma^2}\).

\(^1\)The operational significance of the fractional operator \(-\sqrt{-\Delta}\), which is the generator of the Poisson scale space, is explained in detail by Duits et al. \[8\]. In Fourier space it corresponds to the multiplicative operator \(-||\omega||\).
Applying Definition 5 to equation (3.21) reveals that the inner products in the Gramm matrix can be expressed as a Gaussian derivative of the spatial representation of $\phi_\gamma$. Note that this can be exploited for any operator that one chooses to use as a regulariser.

At this point the Gramm matrix can be constructed. Inversion of this matrix is done by means of Singular Value Decomposition. The projection onto the filters can be done in either the frequency or the spatial domain. The image in the Fourier domain can be obtained by projecting on the Fourier representations of the filters,

$$ g = G^{ij} c_j \kappa_i = \mathcal{F}^{-1} [G^{ij} c_j \hat{\kappa}_i] $$  \hspace{1cm} (3.23)

This avoids problems with the singularity of the Bessel function. An Inverse Discrete Fourier Transform of the sampled reconstruction function results in the desired image.

### 3.3.3 Richer Features

Obtaining a visually appealing reconstruction from singular points can be achieved by selecting an “optimal” space for projection. This approach is discussed above. Another way to enhance the quality of the reconstruction is by using more information about the points that are used for reconstruction. In the standard linear reconstruction scheme only up to second order differential structure was used. In our experiments also higher order differential properties of the singular points were taken into account. This has the side effect that the Gramm matrix will be harder to invert when more possibly dependent properties are used.

### 3.4 Evaluation

To evaluate the newly proposed reconstruction scheme reconstructions from singular points of different images are performed. The singular points are obtained using ScaleSpaceViz [22], which is based on a zero-crossings method. After the singular points are found the unstable ones are filtered out by applying a threshold on the amount of structure that is present around a singular point. The amount of structure can be found by calculating the “differential total variation norm” or “deviation from flatness”

$$ tv = \sigma^4 \text{Tr} (H^2) $$  \hspace{1cm} (3.24)

that was proposed by Platel et al. [32]. $H$ represents the Hessian matrix and $\sigma$ represents the scale at which the singular point appears. The reconstruction algorithm is implemented in Mathematica [39].

The images that are chosen to evaluate the performance of the reconstruction algorithm are those used by Kanters et al. and Lilhholm et al. for the evaluation of their reconstruction algorithms [26] [23], Lena’s eye and MR brain. The size of the former image is $64 \times 64$ pixels and the size of the latter image is $128 \times 128$ pixels. The pixel values of these images are integer valued ranging from 0 to 255.
3.4. Evaluation

Figure 3.2: Reconstruction from 31 singular points of Lena’s eye with up to second order features. The upper row shows the original image and reconstructions with $\gamma = 0$ and $\gamma = 5$. The second row shows reconstructions with $\gamma = 22$, $\gamma = 50$ and $\gamma = 250$. The first image in the second row shows the reconstruction with the lowest relative $L_2$-error.

3.4.1 Qualitative Evaluation

First we study reconstruction from singular points taking into account up to second order derivatives of the image at the locations of the singular points. Figure 3.2 shows the reconstruction from 31 singular points of Lena’s eye. These points are selected using a tv-norm of 32. Note that the tv-norm scales with the square of the image range. The first image in the upper row displays the image from which the singular points were obtained. Successive images are reconstructions from these points with an increasing $\gamma$. The second image in the first row shows a reconstruction with $\gamma = 0$, which equals the reconstruction by Kanters et al. [23], and the first image in the second row depicts the reconstruction with a minimal relative $L_2$-error. The same convention is used in the reconstruction from 55 singular points of MR brain that is displayed in Figure 3.3. The singular points of this image were acquired using a tv-norm of 128.

Figures 3.2 and 3.3 show the “fill-in effect” of the smoothing prior. The reconstruction with the smallest relative $L_2$-error is visually more appealing than the images with a smaller $\gamma$. A reconstruction with $\gamma = 250$ lacks details that were visible in the other reconstructions. This happens because the Gramm matrix is harder to invert when dependent basis functions are used. With an increasing $\gamma$ the kernels become wider and thus more dependent on one another. The reconstructions of MR brain show “leaking” edges. Because the prior smoothes the image the very sharp edges of this image are not preserved and consequently the leaking effect appears.

To investigate the influence of enrichment of the features the same experiments are repeated but now up to fourth order derivatives are taken into account in the features. The results for the reconstruction from the singular points of Lena’s eye can be found in Figure 3.4 and the results for the reconstruction from the singular points of MR brain are depicted in Figure 3.5.
3.4. Evaluation

Figure 3.3: Reconstruction from 55 singular points of *MR brain* with up to second order features. The upper row shows the original image and reconstructions with $\gamma = 0$ and $\gamma = 3$. The second row shows reconstructions with $\gamma = 7$, $\gamma = 50$ and $\gamma = 250$. The first image in the second row shows the reconstruction with the lowest relative $L_2$-error.

In both cases the images show more detail and are visually more appealing than their second order counter parts. The reconstruction of the *MR brain* image still shows leaking but this effect is reduced when compared to second order reconstruction.

Figure 3.4: Reconstruction from 31 singular points of *Lena’s eye* with up to fourth order features. The upper row shows the original image and reconstructions with $\gamma = 0$ and $\gamma = 4$. The second row shows reconstructions with $\gamma = 19$, $\gamma = 50$ and $\gamma = 250$. The first image in the second row shows the reconstruction with the lowest relative $L_2$-error.
3.5. Conclusions & Recommendations

Figure 3.5: Reconstruction from 55 singular points of MR brain with up to fourth order features. The upper row shows the original image and reconstructions with $\gamma = 0$ and $\gamma = 4$. The second row shows reconstructions with $\gamma = 8$, $\gamma = 50$ and $\gamma = 250$. The first image in the second row shows the reconstruction with the lowest relative $L_2$-error.

3.4.2 Quantitative Evaluation

In order to verify the quality of the reconstructions of both images under a varying $\gamma$ the relative $L_2$-error,

$$L_2\text{-error} = \frac{||f - g||_{L_2}}{||f||_{L_2}},$$

(3.25)

of the reconstructed images is calculated. Figure 3.6 shows four graphs depicting this error for both second order and fourth order reconstruction of Lena’s eye and MR brain. All graphs show that an optimal value exists for the $\gamma$ parameter. This can be explained by the fact that the Gramm matrix is harder to invert with increasing $\gamma$ due to increasing correlation among the filter cf. equation (3.21). Because of that dependent equations will be removed during the SVD. This leads to a reconstruction with less detail and thus a larger $L_2$-error. The reconstructions of the MR brain image show an increasing $L_2$-error with an increasing $\gamma$. This error becomes even larger than the $L_2$-error of the reconstruction with $\gamma = 0$. This can be attributed to the sharp edges of the head that are smoothened and thus show leaking into the black surroundings of the head. The background clearly dominates the contribution to the $L_2$-error. Lena’s eye does not suffer from this problem because of its smoothness.

3.5 Conclusions & Recommendations

We have proposed a linear reconstruction method that leaves room for selection of arbitrary priors as long as the prior is a norm of Sobolev type. This greatly reduces the complexity of the reconstruction algorithm compared to non-linear methods.
3.5. Conclusions & Recommendations

Figure 3.6: The relative $L_2$-error of the reconstructions from 31 singular points of *Lena’s eye* (upper row) and 55 singular points of *MR brain* (lower row). The first column shows the $L_2$-error for varying $\gamma$ when second order reconstruction is used, i.e. up to second order derivatives are taken into account in the features. The second column displays fourth order reconstruction.

We have selected one possible prior characterised by a free parameter $\gamma$ that aims for a smooth reconstruction. This provides a control parameter for selecting different metameric reconstructions, i.e. reconstructions all consistent with the prescribed constraints. Comparisons with standard linear reconstruction as done by Kanters et al. [23] show that it is possible to improve the reconstruction quality while retaining linearity. Reconstruction from a selection of singular points of the *MR brain* image proves to be more difficult than reconstruction of smoother images like *Lena’s eye*. The problem, that shows up as “leaking” edges, is reduced by taking higher order differential structure into account in the reconstruction algorithm. When the $\gamma$ parameter is increased basis functions get more dependent on each other. This leads to a harder to invert Gramm matrix and consequently to a reduction of detail in the reconstruction.

Both taking $\gamma > 0$ and taking higher order features into account lead to visually more appealing images and a smaller $L_2$-error compared to standard linear reconstruction. It remains an open question how to select an optimal $\gamma$.

Future work will include the use of anisotropic basis functions that depend on the local image orientation and investigation of an adaptive $\gamma$ parameter. Next to these improvements to the algorithm other priors that fit in the proposed framework will be investigated.
Chapter 4

Flow Reconstruction

The reconstruction algorithm that is proposed in the previous chapter will be used to obtain a dense flux field that explicitly contains the desired flow information. The velocity of the anchor points (recall Theorem 1) is encoded in a multi-scale flux field $J$.

4.1 Proposed Vector Field Retrieval Method

**Definition 6 (Dense Flux Field).** The dense flux field $j \in L^m_2(\mathbb{R}^n)$ at scale $s = s_0$, in which $s_0 > 0$ is related to the grid scale, is defined by

$$j = f \mathbf{v}$$

(4.1)

with $f \in L_2(\mathbb{R}^n)$ the image intensity and $\mathbf{v} \in L^m_2(\mathbb{R}^n)$ the dense optic flow field introduced in Definition 2.

In order to find a dense flux field from flux measurements in scale space the minimisation with respect to $k \in L^m_2(\mathbb{R}^n)$ of the following Euler-Lagrange formalism is proposed

$$E[k] = \frac{1}{2} \int_{\mathbb{R}^n} k \cdot k \, dV + \sum_{p \in P} c_p \cdot \left[ \int_{\mathbb{R}^n} k \phi_p \, dV - J_p \right].$$

(4.2)

With $k$ denoting the estimated dense flux field (gray value times scale space velocity) at scale $s = s_0$. The filter responses for each point $p \in P$ are

$$J_p = \int_{\mathbb{R}^n} j \phi_p \, dV.$$

(4.3)

These measurements of the image flux in scale space can only be performed reliably in well defined anchor points in scale space. We restrict ourselves to the anchor points that were introduced in Chapter 2 namely singular points.
Theorem 4. The Euler-Lagrange equation of equation (4.2) is decoupled for each spatial and scale component. Hence the minimising flux field equals

$$k = \mathcal{P}_V j,$$

(4.4)

with $V$ the span of the filters $\{\phi_p\}_{p \in P}$, in which $P$ denotes the collection of anchor points, and $\mathcal{P}_V$ the component-wise linear projection onto $V$.

Proof. Making the vector inner product in equation (4.2) explicit we note that

$$E[k] = \sum_{a=1}^{m} \left[ \frac{1}{2} (k_a, k_a)_{L_2} + \sum_{p \in P} \alpha_p^a \left( (k_a, \phi_p)_{L_2} - J_p^a \right) \right] \overset{\text{def}}{=} \sum_{a=1}^{m} E[k_a].$$

(4.5)

This is a sum of positive convex energies and can therefore be minimised for each component of $k$ separately. With the equivalence relation

$$j \sim k \iff j_a - k_a \in V^\perp \ \forall \ a = 1 \ldots m$$

(4.6)

the flux fields are called equivalent if their known set of features (equation (4.3)) do not differ. Because the features in the reconstructed flux field and the source flux field do not differ it suffices to minimise $\sum_a (k_a, k_a)$. By the same token as before for the static case, cf. Appendix A we have,

$$\min_{k \in [j]} \sum_{a=1}^{m} (k_a, k_a)_{L_2} = \min_{k \in [j]} \sum_{a=1}^{m} \left[ \|k_a - \mathcal{P}_V j_a\|^2 + \|\mathcal{P}_V J_a\|^2 \right] = \min_{k \in [j]} \sum_{a=1}^{m} \left[ \|k_a - \mathcal{P}_V j_a\|^2 + \|\mathcal{P}_V J_a\|^2 \right],$$

(4.7)

with $[j] = \{k \in L_2(\mathbb{R}^n) \mid j \sim k\}$ the equivalence class of $j$. Note that, because $k \in [j]$, we have $\mathcal{P}_V k_a = \mathcal{P}_V j_a$ whence $\mathcal{P}_V (k_a - \mathcal{P}_V j_a) = 0$, i.e. $(k_a - \mathcal{P}_V j_a) \in V^\perp$, so indeed the Pythagoras Theorem can be applied. It is clear that equation (4.7) is minimised for $k_a = \mathcal{P}_V j_a$. So

$$k = \mathcal{P}_V j,$$

(4.8)

which completes the proof. □

4.2 Richer Features

When more information is taken into account one can expect to obtain more accurate results. This can be done in a similar fashion as is described in the previous chapter, taking higher order derivatives of the flux into account. For first order differential information this leads to

$$\frac{\partial J(x; s)}{\partial x^a} = -\int_{\mathbb{R}^n} j(x) \frac{\partial \phi_a(x)}{\partial x^a} dx.$$ 

(4.9)

Note that these derivatives are taken with respect to space, not time. When singular points are tracked this simplifies to

$$\frac{\partial J(x_p; s_p)}{\partial x^a} = u(x_p; s_p) \frac{\partial v(x_p; s_p)}{\partial x^a} \ \forall \ p \in P,$$

(4.10)
since the gradient of $u(x; s)$ vanishes at these points. So we need to measure first order differential information of the motion of anchor points. For notational convenience we write $v = A^{-1}b$, with

$$A = \begin{bmatrix} H & \nabla \frac{\partial u}{\partial s} \\ \nabla^T \det H & \ldots & \frac{\partial \det H}{\partial s} \end{bmatrix}$$

(4.11)

and

$$b = \begin{bmatrix} \frac{\partial^2 u}{\partial x \partial t} \\ \frac{\partial \det H}{\partial t} \end{bmatrix}.$$  

(4.12)

$\frac{\partial v}{\partial x^a}$ can be obtained by inversion of (recall Theorem 1)

$$A \frac{\partial v}{\partial x^a} = -\left( \frac{\partial b}{\partial x^a} + \frac{\partial A}{\partial x^a}v \right).$$

(4.13)

This can be repeated for any order of differential information as long as the derivatives that appear in the equations are sufficiently accurate [5, 16, 12].

### 4.3 Evaluation

The results of the optic flow estimation are evaluated quantitatively by means of the angular error and the error in the magnitude of the estimated vector field. Test sequences that are used for evaluation are the famous Yosemite Sequence and the Translating Tree Sequence. Qualitative evaluation is performed by analysis of the flow field of the above mentioned sequences and the Hamburg Taxi Sequence. The latter sequence lacks the availability of ground truth, which is the reason why it is not included in the quantitative evaluation. These image sequences and their ground truth are obtained from the University of Western Ontario (ftp://ftp.csd.uwo.ca in the directory /pub/vision/) as mentioned in a paper by Barron et al. [4]. The algorithm is implemented in Mathematica [39] using the Parallel Computing Toolkit. A part of the source code of the algorithm can be found in Appendix B. The experiments were carried out on the Mathematica cluster that was recently obtained by the biomedical imaging group where this study was performed.

The properties of the different sequences, of which a visualisation of the center frames can be found in Figure 4.1, are as follows.

**Yosemite Sequence** The magnitude of the velocity field is strictly positive everywhere in the image. In the upper part of the image sequence a discontinuity in the vector field manifests itself. This synthetic sequence possess a divergent optic flow field in the upper right corner and a horizontally translating optic flow field in the upper part of the image. The lower left part shows relatively large optic flow vectors of approximately 4.0 pixels/frame. This sequence is regarded hard due to the wide variety in the optic flow field, the occluding edges between the mountains and the horizon and the severe aliasing in the lower portion of the image sequence. The reference flow field for the center image of this sequence is presented in the left part of Figure 1.2.
4.3. Evaluation

**Translating Tree Sequence** The velocity is constant with about 2.0 pixels/frame everywhere in the image. This image sequence is acquired by translating a camera in front of a picture. The reference flow field of the center image of this sequence is presented in the right part of Figure 4.2.

**Hamburg Taxi Sequence** Four moving objects are present in the image sequence, notably a white taxi turning around the corner in the center of the image, a black car moving to the right in the lower left corner, a black van moving to the left in the lower right corner and a pedestrian walking in the upper left part of the image sequence. The velocity of the pedestrian is approximately 0.3 pixels/frame and the velocities of the cars are about 1.0, 3.0 and 3.0 pixels/frame respectively.

![Figure 4.1: Center frames of the test sequences. From left to right, the Yosemity Sequence, the Translating Tree Sequence and the Hamburg Taxi Sequence.](image)

![Figure 4.2: Reference flow fields of the Yosemity Sequence and the Translating Tree Sequence.](image)
Initial guesses of the positions of the singular points are obtained using ScaleSpaceViz \[22\]. The locations of the points returned by this program are refined by iteration over equation \[2.9\] until the estimated error is below $10^{-3}$ pixels. In case a singular point position can not be refined, which can be caused by a poor initial guess of the singular point position, the point is not taken into account. Furthermore only stable singular points are selected for the evaluation of the algorithm. The stability measure used is based on the extremal variances of a singular point of an image that is distorted with pixel uncorrelated noise as described by Baltnachnova et al. \[3\]. If the maximum of the extremal variances is below the displacement of 1 pixel the singular point is selected. Otherwise the singular point will not be taken into account. Furthermore only singular points above a certain scale are selected to ensure the numerical noise on the higher order derivatives is low. The setting for our experiments is $s = 1.75$ pixels\[^2\]. At this scale evaluation of the $4^{th}$ order derivatives that are present in the computation of both the velocity of a singular point and the employed stability measure can be computed reliably.

The dense flux field is estimated for each spatial component separately as described in Theorem \[4\]. The scale component is, for the sake of simplicity, ignored. More on this subject will be discussed in section 4.4.

For each of the test sequences the velocities of the singular points are estimated as proposed in Theorem \[1\]. When the velocity of a singular point is unstable the point is not taken into account in the algorithm. The velocity of a singular point is considered stable as long as its stability measure (recall Theorem \[1\] equation \[2.6\]) satisfies

$$\det \begin{bmatrix} \mathbf{H}(t; \mathbf{x}, s) & \nabla \frac{\partial u(t; \mathbf{x}, s)}{\partial s} \\ \nabla^T \det \mathbf{H}(t; \mathbf{x}, s) & \ldots & \frac{\partial \det \mathbf{H}(t; \mathbf{x}, s)}{\partial s} \end{bmatrix} > \epsilon . \quad (4.14)$$

In the experiments $\epsilon$ is set to $10^{-5}$. This settings of $\epsilon$ is an educated guess and should be subject of further research.

The time derivatives in equation \[2.7\] are either taken as two-point central derivatives or as Gaussian derivatives with fixed time scale $\tau$, where $\tau$ is equivalent to $\sigma^2/2$ in a Gaussian distribution. In case of discrete derivatives the derivatives that do not possess an explicit temporal component are taken as sheer spatial Gaussian derivatives. Otherwise the Gaussian derivatives are taken with an anisotropic scale component in the time direction, notably $\tau$. The performance of the reconstruction depends on the setting of the introduced parameter $\gamma$. Since the goal of this section is to evaluate the feasibility of the proposed algorithm only a fixed sensible setting of this parameter, $\gamma = 16$ pixels, is taken into account. Definition \[6\] is applied to resolve the desired dense optic flow field. To overcome numerical problems while resolving the optic flow from the estimated dense flux field (recall Definition \[6\]) the input image sequences, ranging from 0 to 255, are shifted up to range from 1 to 256.
4.3. Evaluation

Figure 4.3: The estimated flow field of the Yosemite Sequence using time scale $\tau = 2$ frames. Subfigure (a) shows the estimated vector field with velocity encoded in the length of the vectors and subfigure (b) shows the same vector field without the velocity information. Thin headed arrows denote vectors that possess an angular error that is larger than 20°.

4.3.1 Qualitative Evaluation

Evaluation is commenced by analysis of the estimated flow field of the Yosemite Sequence. Figure 4.3 shows the estimated vector field with time scale $\tau = 2$ frames. The thin headed arrows indicate vectors with an angular error larger than 20°. A projection of the singular points, that were used in the reconstruction step of the algorithm, on the central image of the sequence is depicted Figure 4.4. Erroneous vectors appear mostly on the edges and in the occlusion area in the top of the image. A closer look at the vector field reveals a blob like structure in the magnitude of the velocity field. This can be explained by the fact that only 303 singular points are used in the reconstruction of a $316 \times 252$ flow field. The reconstruction strives for a smooth result with a minimal $L_2$-norm. This means the velocity is pulled to zero in places that are not sufficiently covered by the singular points that are used in the algorithm. The smoothness that circumvents this problem is embodied in the $\gamma$ parameter which, in this case, was apparently set too low. This problem that shows up as an error in the velocity already appeared in a similar fashion in the previous chapter, viz. in the reconstruction of the MR brain image. The reconstruction results of that MR image for several settings of $\gamma$ can be found in Figure 3.3. In the previous chapter it was also noted that a larger $\gamma$ can make the system less stable which in turn can lead to loss of features. The vectors showing extreme errors in its magnitude are caused by unstable singular point flow estimation. If $\epsilon$ was set to a larger value these errors would not occur. Increasing $\epsilon$ does however reduce the number of singular points that can be used in the reconstruction step of the algorithm which can lead to overall worse performance.

A visualisation of the angular component of the estimated flow vectors belonging to the Translating Tree Sequence with time scale set to $\tau = 3$ frames is presented in Figure 4.5. In this case problems occur at the boundaries. However only the edges that possess appearing
4.3. Evaluation

Figure 4.4: The projection of a set of singular points on the central image of the Yosemite Sequence. These singular points were used in the reconstruction step of the algorithm in order to obtain the results that are displayed in Figure 4.3.

and disappearing singular points show problems. The results could be influenced negatively by the low number of features (53 singular points with an image size of $150 \times 150$ pixels) that were used in the algorithm.

The optic flow field of the Hamburg Taxi Sequence superimposed on the image on which the optic flow estimation is performed is shown in Figure 4.6. It shows all vectors that possess a magnitude larger than 0.2 pixels. Most notable is the fact that the pedestrian, present in the upper left of the image sequence, is detected by our algorithm. Also some false vectors are present between the white taxi and the black van. This could be caused by the tree that causes both occlusion and additional distortion in the image sequence. Boundary problems are visible at the location where the van enters the sequence.

4.3.2 Quantitative Evaluation

A more objective approach to the evaluation of the quality of the estimated flow field are the average angular error and the average error of the vector magnitude. The results of the experiments conducted on the Yosemite Sequence are summarised in Table 4.1. It shows the angular error, standard deviation of the angular error, the average error in the 2-norm and the number of singular points used in the reconstruction step for different time scales $\tau$. Table 4.2 shows the results for the experiments conducted on the Translating Tree Sequence.

The tables clearly show the method using discrete central derivatives and no time scale component is outperformed by methods that do employ a time scale component. Especially bad results in the average magnitude error of this method stand out. Overall this error is large as was already noticed and explained in the previous section. However outliers, as visible in Figure 4.3 let the error seem more severe than it actually is. Nevertheless we may notice the
30 4.3. Evaluation

Figure 4.5: The estimated flow field of *Translating Tree Sequence* using time scale $\tau = 3$ frames. Thin headed arrows denote flow vectors possessing an angular error above $20^\circ$.

<table>
<thead>
<tr>
<th>Time Scale</th>
<th>Average Angular Error</th>
<th>Standard Deviation of Angular Error</th>
<th>Average Error in Magnitude</th>
<th># Stable Singular Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d$</td>
<td>$28.53^\circ$</td>
<td>$34.77^\circ$</td>
<td>$269%$</td>
<td>$538$</td>
</tr>
<tr>
<td>1</td>
<td>$24.93^\circ$</td>
<td>$36.00^\circ$</td>
<td>$99%$</td>
<td>$352$</td>
</tr>
<tr>
<td>2</td>
<td>$19.19^\circ$</td>
<td>$34.45^\circ$</td>
<td>$61%$</td>
<td>$303$</td>
</tr>
<tr>
<td>3</td>
<td>$19.97^\circ$</td>
<td>$34.30^\circ$</td>
<td>$72%$</td>
<td>$271$</td>
</tr>
<tr>
<td>4</td>
<td>$22.94^\circ$</td>
<td>$42.61^\circ$</td>
<td>$77%$</td>
<td>$251$</td>
</tr>
<tr>
<td>5</td>
<td>$25.00^\circ$</td>
<td>$39.50^\circ$</td>
<td>$91%$</td>
<td>$226$</td>
</tr>
</tbody>
</table>

Table 4.1: Summary of the performance of the proposed optic flow retrieval method applied to the *Yosemite Sequence*. Time scale $d$ means discrete central derivative.

estimation of the vector magnitudes is far from optimal. When the time scale is increased less singular points can be taken into account in the reconstruction step of the algorithm. The setting of the time scale is a tradeoff between the number of features that are noise robustness. The performance of the algorithm is worse for the *Translating Tree Sequence* than for the *Yosemite Sequence*. This could be caused by the fact that in the *Translating Tree Sequence* relatively few singular points are available compared to the *Yosemite Sequence*.

A histogram of the angular errors of the estimated optic flow in the *Yosemite Sequence* using a time scale of $\tau = 2$ frames is shown in Figure 4.7. The figure shows that the number of erroneous vectors decreases rapidly with the magnitude of the angular error, which is a promising result. Similar results are obtained for both sequences using different time scales. Still many vectors that point to the opposite direction are present. This increases the standard deviation of the angular error that is presented in Table 4.1 and Table 4.2. Since we did no develop a confidence measure it is unjustified to leave these incorrect vectors out.
Figure 4.6: The estimated flow field of *Hamburg Taxi Sequence* using time scale $\tau = 2$ frames superimposed on the image on which the calculations were applied. Only vectors that denote a velocity larger than 0.2 pixels/frame are displayed.

To give the reader a grasp of the performance of other methods Table 4.3, showing the performance of several alternative optic flow estimation methods on the *Yosemite Sequence*, is included. All methods that are presented in Table 4.3 produce a 100% dense vector field. The results are obtained from Barron et al. [4] and Brox et al. [6]. Unfortunately no results from the scale space method by Florack et al. [13] were available for this sequence. One should notice this is a feasibility study that is neither completely developed nor tuned for optimal performance.

### 4.4 Conclusions & Recommendations

A rough approximation of the optic flow of an image sequence can be obtained using anchor points in a Gaussian scale space. General reconstruction theory can be applied to optic flow retrieval from anchor points by encoding the dynamic properties of these points in a so called flux field. In order to obtain a dense flux field from the flux of multi scale anchor points, which is related to the optic flow of an image sequence, general reconstruction theory can be applied to each dimension independently. Theory covering the inclusion of higher order flow information in the algorithm is presented but evaluation is not included. Possible disadvantages of inclusion of higher order features is the loss of accuracy due to the high order of derivatives that are needed in the computation of the dynamic properties of singular
Time Scale | Average Angular Error | Standard Deviation of Angular Error | Average Error in Magnitude | # Stable Singular Points |
--- | --- | --- | --- | --- |
$d$ | $34.56^\circ$ | $37.95^\circ$ | $299\%$ | $130$ |
1 | $34.54^\circ$ | $52.24^\circ$ | $52\%$ | $79$ |
2 | $32.72^\circ$ | $51.21^\circ$ | $51\%$ | $73$ |
3 | $27.11^\circ$ | $43.03^\circ$ | $53\%$ | $53$ |
4 | $30.69^\circ$ | $40.54^\circ$ | $55\%$ | $48$ |
5 | $31.86^\circ$ | $42.89^\circ$ | $55\%$ | $42$ |

Table 4.2: Summary of the performance of the proposed optic flow retrieval method applied to the *Translating Tree Sequence*. Time scale $d$ means discrete central derivative.

![Figure 4.7](image)

Figure 4.7: Histogram of the angular error of the estimated flow field belonging to the *Yosemite Sequence*.

points. First order multi scale flux in 2D for example requires $5^{th}$ order spatial derivatives. For coarse scales this need not be a problem.

The evaluation of zeroth order flow estimation, for which up to zeroth order dynamic features are taken into account, shows reasonable results considering the sparseness of features. All parameters are set to a fixed setting and evaluation is conducted on distinct image sequences possessing essentially different properties that are meaningful for the evaluation of optic flow. Therefore we can state that the proposed method is feasible.

Problems that showed up are errors in the estimated velocity that are mainly caused by too sparse a set of features. Other causes of error are instability of the anchor point flow.
4.4. Conclusions & Recommendations

<table>
<thead>
<tr>
<th>Technique</th>
<th>Average Angular Error</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horn &amp; Schunck (original)</td>
<td>32.43°</td>
<td>30.28°</td>
</tr>
<tr>
<td>Our Method</td>
<td>19.19°</td>
<td>34.45°</td>
</tr>
<tr>
<td>Singh Step 1</td>
<td>18.24°</td>
<td>17.02°</td>
</tr>
<tr>
<td>Anandan</td>
<td>15.84°</td>
<td>13.46°</td>
</tr>
<tr>
<td>Singh Step 2</td>
<td>13.16°</td>
<td>12.07°</td>
</tr>
<tr>
<td>Nagel</td>
<td>11.71°</td>
<td>10.59°</td>
</tr>
<tr>
<td>Horn &amp; Schunck (modified)</td>
<td>11.26°</td>
<td>16.41°</td>
</tr>
<tr>
<td>Uras et al.</td>
<td>10.44°</td>
<td>15.00°</td>
</tr>
<tr>
<td>Weickert et al.</td>
<td>5.18°</td>
<td>8.68°</td>
</tr>
<tr>
<td>Brox et al. (2D)</td>
<td>2.46°</td>
<td>7.31°</td>
</tr>
<tr>
<td>Brox et al. (3D)</td>
<td>1.94°</td>
<td>6.02°</td>
</tr>
</tbody>
</table>

Table 4.3: Performance of different methods from literature on the Yosemite Sequence. All methods produce a dense vector field.

...estimation, which is in general caused by occlusion and the image boundary, and numerical instability in the inversion step that is present in the reconstruction algorithm.

One of the unique properties of the proposed algorithm is that the velocity vectors possess a scale component. This can potentially be useful for segmentation and camera motion estimation. We neglected the presence of this scale component in the reconstruction process. It does give more information about the local structure of the vector field and can therefore be used as an extra constraint for the estimation of the spatial components of the optic flow field. A “zoom”, for example, will cause singular points to translate and move downwards in scale. A method to add this extra property in the estimation of the optic flow field has yet to be proposed.

The parameter $\epsilon$ is set to a fixed value and caused large errors in the estimated optic flow field. Although “tuning” will work for individual sequences a more fundamental understanding of the subject of anchor point flow stability should be developed.

Taking more anchor points into account will probably boost the performance of the algorithm. Possible candidates are singular points of the laplacian of an image, so called blob points as proposed by Lindenberg [28], scale space saddles as proposed by Kuijper [24, 25] or other singular or critical points. One can use all these point types simultaneously which makes the method comparable to phase based methods, yet still taking into account the inherent multi scale nature of the considered image sequence.

A more sophisticated reconstruction method may result in better performance of the entire algorithm. Possible improvements of the applied reconstruction method were proposed in the final section of Chapter [3].
Chapter 5

Conclusions & Recommendations

5.1 Conclusions

Two main goals have been achieved:

1. significant improvement of linear stationary reconstruction from singular point features

2. demonstration of the feasibility of optic flow estimation from dynamic anchor point attributes

The central part of the algorithm depends on a technique called stationary reconstruction. It aims for a reconstruction from a sparse set of features evaluated at so-called anchor points that is, in some precise sense, similar to the image from which the features were extracted. We developed a linear reconstruction framework that generalises a previously proposed scheme. The quality of the reconstruction is visually more attractive and has a smaller $L_2$-error than the previously proposed linear methods.

Singular points of a Gaussian scale space of an image are identified as feasible candidates for dynamic analysis. In addition to an error measure analytical expressions describing the dynamic properties of these anchor points are presented.

In order to estimate the optic flow of an image sequence the dynamic properties of the anchor points are encoded in a multi scale flux field. We show that the reconstruction algorithm can be applied to each dimension independently, generating a dense flux field that is related to the optic flow of an image sequence. A unique property of the obtained vector field is the scale component that arises from the “vertical” movement of anchor points in scale space.

Evaluation of the newly developed optic flow estimation algorithm shows promising results considering the small number of anchor points. Small objects that are not always detected by other methods presented in the literature do show up using our method. However the generated vector field does not compete with results of state of the art methods. Problems show up in the estimated magnitude of the vectors. This can be explained by the relatively small number of anchor points that are taken into account during the evaluation. One should
keep in mind that this is a feasibility study of a method that is still in its infancy. Current results are comparable to or better than the initially proposed optic flow estimation method by Horn & Schunck. Although no striking performance is presented we can conclude the newly proposed method is feasible and does appear promising, but more research should be conducted concerning this subject as suggested.

5.2 Recommendations

Improvement of the stationary reconstruction method can lead to a better performance of the optic flow estimation method. To this extent research covering anisotropic basis functions that depend on local image structure should be considered. Also other priors can be considered in order to achieve better performance in both reconstruction quality and computational complexity. A dynamic selection of the smoothness parameter $\gamma$ that is present in the current reconstruction method can also lead to an increase of performance. Current research already covers the investigation of alternative inversion algorithms.

A better understanding of anchor point flow stability can circumvent the large errors in the magnitude that showed up in the results of the experiments. In the current method the error measure, as presented in Section 2.3 is not exploited. One could for example use this measure to achieve more reliable vectors or remove incorrect vectors. Taking more anchor points into account can also lead to better performance of the algorithm. Possible candidates are singular points of the laplacian of the image, so called blob points as proposed by Lindenberg [28], scale space saddles as proposed by Kuijper [24, 25] or other singular or critical points. One can use all these point types simultaneously, which makes the method comparable to phase based methods, yet still taking into account the inherent multi scale nature of the considered image sequence. Taking higher order differential structure of anchor points into account may also lead to better performance. One should notice this also puts a higher demand on the derivatives used in the features and therefore could cause numerical errors at fine scales. To overcome this problem one could for instance only take into account higher order features at appropriate/sufficiently coarse scales.

A method for inclusion of the scale component in the estimation of the optic flow field has yet to be proposed. Olver [31] wrote a monograph covering symmetries of differential equations. His work can be a starting point for addressing this problem.

Several methods in literature also incorporate a confidence measure. In our case this could be produced by evaluation of the quality of stationary reconstruction from the static properties of the anchor points. Further research should be conducted concerning this subject.
Appendix A

Alternative Approach to Theorem 3

Recall that $V$ is the span of the filters $\kappa_i$. Then

$$V^\perp = \{ f \in L_2(\mathbb{R}^2) \mid (\kappa_i, f)_A = 0 \ \forall \ i = 1, \ldots, N \} \quad (A.1)$$

On the space of images $L_2(\mathbb{R}^2)$ we define the following equivalence relation:

$$f \sim g \iff f - g \in V^\perp, \quad (A.2)$$

Notice that the set of equivalence/metameric classes is given by

$$L_2(\mathbb{R}^2)/\sim \overset{\text{def}}{=} \{ [f] \mid f \in L_2(\mathbb{R}^2) \} = \{ f + V^\perp \mid f \in L_2(\mathbb{R}^2) \} \quad (A.3)$$

and that an equivalence class $[f] = \{ g \in L_2(\mathbb{R}^2) \mid f \sim g \}$ of representant $f$ is exactly given by those images that have the same features as image $f$. Notice to this end that

$$(\kappa_i, f)_A = (\kappa_i, g)_A \quad \text{for all} \quad i = 1, \ldots, N \iff f - g \in V^\perp. \quad (A.4)$$

Next we show that the unique element $g$ within $[f]$ that minimizes the energy $E[g] = \|g\|_A^2$ is given by the $A$-orthogonal projection of $f$ on $V$, $P_V f$:

$$\min_{g \in [f]} \|g\|_A^2 = \min_{g \in [f]} \|g - P_V f + P_V f\|_A^2 = \min_{g \in [f]} \|g - P_V f\|_A^2 + \|P_V f\|_A^2 \quad (A.5)$$

and this equals $\|P_V f\|_A^2$ only in the case $g = P_V f$. Notice with respect to the last equality (equation (A.5)) is Pythagoras theorem, which can be applied since $(g - P_V f) = (g - P_V g) \in V^\perp$ and $P_V f \in V$. 


Appendix B

Parallel Mathematica Source Code

Only a small part of the source code of the optic flow estimation algorithm is presented. This part of the source code implements the parallel computation of the Gramm matrix (recall equation (3.21)), which is the most time consuming part of the algorithm. We proceed with the documentation of the presented module.

B.1 Documentation of the makeParallelGramm Module

The Mathematica module \texttt{makeParallelGramm} returns the Gramm matrix and evaluates different inner products using the previously generated sampled kernel $\phi_\gamma$. \texttt{GaussianDerivativeAt} is evaluated at each ($x_{\text{size}}/2 + 1$) point difference because of the origin that is shifted to $(0, 0)$. In Mathematica the origin is by default located at $(1, 1)$.

Mathematica module \texttt{SubF} calculates all derivatives at a single position in scale space. A request for calculation is submitted to remote kernels using the module \texttt{Queue} and the needed data are delivered to the different kernels using \texttt{ExportEnvironment}. The \texttt{Wait} module collects all answers produced by the remote kernels. Note that efficiency is increased when higher order derivatives are taken into account. Exploiting the symmetry of the Gramm matrix does not pay off since this will load the local CPU with extra reordering work. Local evaluation does benefit from an implementation that exploits the symmetry.

\texttt{Outer[Plus, points, \{1, -1, -1\}\#&\@points, 1]} constructs all "differences", $(t_i + t_j, x_i - x_j)$. In Section B.2 the program code belonging to this documentation is given.
B.2 Source Code of the makeParallelGramm Module

```plaintext
makeParallelGramm[Rkernel_, Rorders_, Rpoints_, (Rxsize_, Rysize_)] :=
Module[{@tensor, gramm, timer, kernel, orders, points, xsize, ysize, pids},
  Clear[sub@];
  kernel = Rkernel;
  orders = Rorders;
  points = Rpoints;
  xsize = Rxsize;
  ysize = Rysize;

  sub@[{st_, dx_, dy_}] := Partition[
    Apply[GaussianDerivativeAt, Flatten[
      Map[{{xsize/2 + 1 + dx, st, #[[1]]}, {ysize/2 + 1 + dy, st, #[[2]]}} &,
        Outer[Plus, orders, orders, 1], (2)],
        1]][kernel],
    Length[orders]
  ] // N
  ExportEnvironment[sub@, orders, kernel, points, xsize, ysize];

  timer = AbsoluteTime[];
  pids =
    Map[
      Composition[Queue, sub@], #] &,
    Outer[Plus, points, {1, -1, -1} # &/@points, 1]
  ];
  @tensor = Wait[pids];
  Print["Remote evaluation took ", AbsoluteTime[] - timer, " seconds."];

  gramm = Chop[Flatten[
    Map[
      Flatten,
      Transpose[@tensor, (2, 4, 1, 3)],
      (2)
    ],
    1
  ]];
  gramm
]
```
Bibliography


