OPTICAL FLOW ESTIMATION
FOR 3D RECONSTRUCTION

Ocean Systems Lab
School of EPS
Heriot-Watt University
Riccarton, Edinburgh
EH144AS United-Kingdom

Supervisors: Miguel Arredondo
Katia Lebart

ENST Bretagne
Technopole de Brest-Iroise
29280 Plouzane
France

Followed by: Jean-Marc Boucher
Acknowledgements

I would like to thank at first my supervisor Miguel Arredondo for his guidance and his enthusiastic dynamism.

I am also very grateful to Dr Katia Lebart for her pertinent advices and for the time she dedicated to me.

Finally I would like to thank Yvan Petillot and all the other people from the lab.
Contents

Introduction 6

1 Presentation of the university 9
   1.1 Heriot-Watt University .............................. 9
   1.1.1 A brief history ........................................ 9
   1.1.2 Research and knowledge .......................... 10
   1.2 Ocean Systems Laboratory ............................. 11
   1.2.1 Projects in the Ocean Systems Laboratory .... 11
   1.2.2 The AMASON project ................................. 15

2 Introduction to Optical Flow 16
   2.1 The Brightness Constancy Assumption .............. 16
   2.2 Constraining the optical flow ....................... 17
       2.2.1 The aperture problem ............................ 17
       2.2.2 The most popular constraints .................. 18
       2.2.3 Constraining the direction ..................... 19
   2.3 Conclusions ............................................. 20

3 Preliminary investigations 22
   3.1 The Negahdaripour algorithm ......................... 22
       3.1.1 Principle ............................................. 22
       3.1.2 Results ............................................... 23
   3.2 Using information from several images .............. 24
       3.2.1 Selection of the common region of the images 24
       3.2.2 Averaging ............................................. 26
       3.2.3 Results ............................................... 26
4 The regularization
  4.1 The aim of the regularization
     4.1.1 Theory
     4.1.2 The choice of the regularizer
  4.2 Practical considerations to solve the problem
     4.2.1 Reaching a solvable system
     4.2.2 The multi-scale process
     4.2.3 Key parameters and their influence
  4.3 Experimental results

Conclusion
# List of Figures

1. Trawling mark in the sea bed .......................... 7
1.1 Ancient and new logos of Heriot-Watt University. 10
1.2 Main entrance of Heriot-Watt University. .......... 10
1.3 The Ocean Systems Lab and the University .......... 12
1.4 A vehicle developed by the OSL, the Rauver (top) and experiments in the test tank of the laboratory (bottom). .. 14
1.5 AMASON logo .......................................... 15
2.1 The aperture problem ................................. 18
2.2 Comparison between the normal optical flow (c) and the direction-constrained optical flow (d) ................. 21
3.1 Trees ....................................................... 24
3.2 Optical Flow for the trees image: Normal Optical Flow (a) and Generalized Optical Flow (b) .................. 25
3.3 Selection of the common part between several images . 25
3.4 Selected common part on several images ............. 26
3.5 Two underwater images ................................ 27
3.6 Selected common region between 11 frames .......... 27
3.7 Direction-constrained optical flow (a) and depth estimation (b) 28
3.8 Averaged optical flow and depth estimation on 4 frames ... 29
3.9 Averaged optical flow and depth estimation on 11 frames .. 30
4.1 Trees ....................................................... 39
4.2 Optical Flow without regularization ................... 40
4.3 Resulting Optical Flow for the 3 proposed regularizations 41
4.4 Role of the weight of the regularization ................ 42
4.5 Influence of the actualization parameter ............... 43
4.6 Depth estimation of a trawling mark . . . . . . . . . . . . . . 44
4.7 3D representation of the trawling mark . . . . . . . . . . . . 45
Introduction

Human impacts on the environment are nowadays of prime importance. Scientists explore human impacts to study incidence on environment, especially in terms of fauna and flora survival. These considerations are also important for the sea bed. Indeed, human activities like fishing have great impact on the environment. For example, trawling fishing nets generally leave a mark on the floor, which destroys life activities. Trawling marks studies bring crucial information for marine scientists.

My traineeship has been held in this context at the Heriot-Watt University (Edinburgh, Scotland). I joined the Oceans Systems Laboratory team, specialized in underwater images processing, with Miguel Arredondo and Katia Lebart as supervisors.

This traineeship concludes my studies at the french engineering school Ecole Nationale Supérieure des Télécommunications de Bretagne (ENSTB) and is more precisely available for the DEA degree, in the domain of Signal and Image Processing.

The exact title of the traineeship was Optical Flow Estimation for 3D Reconstruction. The aim was to investigate robust methods of optical flow estimation to recover the depth of the sea floor. This work has been performed in great collaboration with Miguel Arredondo, who is doing a PhD thesis in the lab and is very familiar with the topic.

As shown on figure 1, trawling fishing boats generally leave marks on the sea bed.
These images are video images of the sea bed recorded by a camera attached to a moving vehicle. The aim is to estimate the depths on these images,
INTRODUCTION

Figure 1: Trawling mark in the sea bed

which contain important information for marine scientists. For non-zero translational motion of the vehicle, depth is highly related to optical flow. Optical flow is defined as \textit{the estimate of the projection in 2D of the real 3D motion}. As points closer to the camera move faster than points further away, the magnitude of the optical flow gives an estimate of the depth of the sea floor.

The traineeship was composed of two main tasks:

- In a first stage, my work focused on the understanding of the subject. This part lead to the implementation of preliminary algorithms in order to improve results obtained by Miguel Arredondo.

- The second stage focused on the implementation of an algorithm which intends to estimate optical flow smoothing homogeneous regions and preserving discontinuities of the image. This part lead to a promising depth estimation of the scene.

This work required technical knowledge, for example, in image processing and MATLAB implementation, as well as personal competences like English practise and team working.

The first chapter of the report presents the context of the traineeship, the second chapter proposes a technical introduction to the optical flow. The
third chapter details preliminary investigations to arrive to the main algorithm detailed in the final chapter.
Chapter 1

Presentation of the university

1.1 Heriot-Watt University

1.1.1 A brief history

Established in 1966, Heriot-Watt University now occupies a 154 ha (380 acre) campus at Riccarton, in SW Edinburgh. With 14,800 students (9000 of whom are engaged on distance-learning programmes), almost 1000 staff and an income exceeding £60 million, it is the city’s second University.

Founded by Leonard Horner (1785 - 1864) as the School of Arts of Edinburgh in 1821, specifically for the education of the working classes, it was one of the first technical colleges. Funds were raised for new premises in the memory of James Watt (1736 - 1819) and the Watt Institution came into being. As early as 1869, women were allowed to join classes. New premises were built in Chambers Street (Edinburgh) designed by David Rhind (1872), and following merger with George Heriot’s Hospital (1885) it was this building which remained the headquarters of the Heriot-Watt College until 1992. The College became Heriot-Watt University on the recommendation of the Royal Commission on Higher Education chaired by Lord Robbins (1966).

The Riccarton estate, which was formerly owned by the Gibson-Craig family, was purchased by the Midlothian Council (1967) and gifted to the University in 1969. A phased move from the centre of the city to this new Campus
1.1.2 Research and knowledge

Heriot-Watt has a flourishing research community with an international reputation for scientific and technological excellence, including the commercialisation of academic research. Major research initiatives are concentrated in niche areas including photonics and new optics-based technologies, business
logistics, technical textiles, advanced robotics, nano-science and technology, proteomics and nutrition Microsystems, biomimetics, virtual reality and engineering design. Heriot-Watt University is divided into fifteen departments regrouped in schools, including the Department of Electrical, Electronic & Computer Engineering (EECE) of the School of Engineering and Physical Sciences (EPS) which the Ocean System Laboratory in part of (figure 1.3).

1.2 Ocean Systems Laboratory

The Ocean Systems Laboratory (OSL) is a leading international centre for research, development and exploitation of acoustic, robotic and video-imaging systems, with particular emphasis on 1D, 2D and 3D signal processing, control, architecture and modelling. Under the direction of Prof. D.M. Lane, its staff counts about 40 people including 8 academic researchers, 13 research associates and 12 PhD students.

The laboratory works internationally with industry, government and research organisations and aims at developing novel technology for sub-sea, defence and other industry applications. Some of the OSL’s technology is currently being commercialised through a spin out partner, SeeByte Ltd.

1.2.1 Projects in the Ocean Systems Laboratory

Computer vision, sonar and image processing are some tools that the laboratory has been actively developing to further automate deep water operations. Throughout the past decade the OSL has been involved in many European projects addressing several aspects of marine technology: Sonar Data Processing, Computer Vision, Unmanned Underwater Vehicles (UUVs), Subsea Robotics, and Through Water Communications.

Sonar and image processing and Computer Vision are the major activities of research of the OSL. Sonars are used to detect, track or classify objects, to provide reactive path planning or to perform Concurrent Mapping and Localisation (CML). New tools are being developed in the context of the project AMASON (Advanced MAppling with SONar and video).

In the field of vision, the OSL has developed algorithms for reconstructing 3D shapes from 2D image sequences and for underwater robot position
Figure 1.3: The Ocean Systems Lab and the University
control from real-time processing of video data (visual servoing). But the laboratory is also interested in teleconferencing with the project VIRTUE (VIrtual Team User Environment) which aim is to develop a telepresence system.

The laboratory has developed over the years a small fleet of Unmanned Underwater Vehicles (UUVs). They have been developed to be Remotely Operated Vehicles (ROVs), but the main motivation behind their development is to test Autonomous Underwater Vehicle (AUV) technology. These vehicles were used in projects like ALIVE (Autonomous Light Intervention VEHICLE), ARAMIS (Advanced ROV package for Automatic Mobile Investigation of Sediments) or currently in RAUVER (Remote/Autonomous Underwater Vehicle for Experimentation and Research).
Figure 1.4: A vehicle developed by the OSL, the Rauver (top) and experimentations in the test tank of the laboratory (bottom).
These UUVs are also used to develop underwater robots. The OSL has pioneered the use of flexible robots for sub-sea applications. AMADEUS (Advanced MANipulation for DEep Underwater Sampling) had been a major project undertaken in the laboratory in the area of Subsea Robotics. The OSL was also involved in research on Through Water Communication using acoustic, electromagnetic and optical devices. They recently developed ForthBase, a test facility for long-term research into acoustic communications and environmental monitoring within a shallow water channel.

1.2.2 The AMASON project

AMASON is one of the European projects in the laboratory. The internship took place in the context of this project. Following is a brief presentation of its purposes.

![AMASON logo](image)

Figure 1.5: AMASON logo

Marine scientists need more affordable, better coverage and greater capability remote survey and data processing systems for monitoring and enforcement of sustainable ecosystem policies. AMASON (Advanced MAppling with SONar and Video) will support this by researching, implementing and evaluating an inexpensive, modular, reconfigurable multi-sonar and video sensor system, with advanced data processing algorithms implemented within a geographic information system (GIS). The plug-and-play system will be readily deployable from ROV’s, AUV’s and towfish of opportunity.

The traineeship has been held in this context on video data and 3D reconstruction of trawling mark.
Chapter 2

Introduction to Optical Flow

2.1 The Brightness Constancy Assumption

Recovering the relief of the sea bed can help to automate the assessment of the environmental impact of trawling, currently performed manually by marine scientists. One way to obtain a rough estimate of the relief is by estimating a disparity map between images using Optical Flow methods. Let us consider a unique camera, attached to a moving vehicle, recording the static sea bed on a video: the points in the foreground move faster than the ones in the background. The optical flow is defined as the estimate of the projection in 2D of the real 3D motion of the scene. Therefore, the magnitude of the optical flow field gives an estimate of the depth map of the scene. Thus, recovering a good 3D map of the scene requires on a robust optical flow estimation [3]. We consider for the rest of the document that the camera records images at the speed of 25 images/sec.

Several methods are developed in the literature to compute the optical flow [5]:

1. Differential techniques: The optical flow is computed from spatio-temporal derivatives of the image brightness.

2. Region-Based Matching: The optical flow is considered as the difference between image regions and the best corresponding ones on the next image.
3. **Frequency-Based Methods**: The optical Flow is estimated by using frequency methods like Fourier transform.

We investigate in this document a **differential technique** [15] based on the **Brightness Constancy Assumption**. This method assumes that the brightness of a scene remains constant over time [26]:

\[
\frac{dE}{dt} = 0
\]  \hspace{1cm} (2.1)

where \( E \) denotes the brightness of the image.

This equation can be written as [26]:

\[
E(x, y, t) = E(x + dx, y + dy, t + dt)
\]  \hspace{1cm} (2.2)

Using the Taylor series expansion of this term, we achieve the well-known form of the **Optical Flow Equation** [26]:

\[
E_x u + E_y v + E_t = 0
\]  \hspace{1cm} (2.3)

where \( u = dx/dt \) and \( v = dy/dt \) are the components of the optical flow and \( E_x, E_y \) and \( E_t \) are the partial derivatives of \( E \) with respect to \( x, y \) and \( t \).

Computing the optical flow consists in solving equation (2.3). It is the starting point for many algorithms described in the literature ([2], [9], [10], [12], [14], [20], [21], [22], [25], [26], [27], [30]).

We detail in the next section the limits of this equation and investigate some ways to solve the problem.

### 2.2 Constraining the optical flow

#### 2.2.1 The aperture problem

Computing the optical flow consists in solving equation (2.3). This equation is valid for each pixel. Therefore, it is a single equation for two unknowns: \( u \) and \( v \). The optical flow estimation is thus an under-constrained problem. Only one degree of freedom can be solved by this equation: it is called the
aperture problem [26].

If this problem is easily understandable from a mathematical point of vue, it has an easy physical interpretation as well [26]: *only the components of the optical flow in the direction of the spatial image gradient can be determined.* It means that only the motion of the camera perpendicular to the image boundaries is visible. Obviously this motion may not be the real one. Therefore, in the general case, it is impossible to recover the real motion of the camera (i.e. the optical flow). This interpretation is illustrated at figure 2.1. The two upper images show the data recorded by two images and the flow deducted from them; the two lower images show the real motion and the real flow.

![Visible Flow](image1)

![Real Flow](image2)

Figure 2.1: The aperture problem

Thus, in order to solve the optical flow (i.e. to move the problem to a well-or over-constrained one), it is necessary to find some additional constraints. Many constraints are found in the literature. The next section describes the most used constraints.

2.2.2 The most popular constraints

2.2.2.1 The Local Spatial Smoothness

It is the most common constraint used to solve equation (2.3). This constraint assumes that the optical flow is constant in a neighbourhood $\Omega_{ij}$ around the considered pixel $(i, j)$. Thus, we obtain a set of equations to
compute the optical flow for the considered pixel \((i, j)\) [26]:

\[ \forall (m, n) \in \Omega_{ij} \quad E_{x,m} u + E_{y,m} v + E_{t,m} = 0 \]

It is possible to gather the equations under matrix form and to use an algorithm such as Least-Squares to solve them.

### 2.2.2.2 Second-order derivations

Another classical constraint used to over-constrain the problem is the use of the *second-order derivation* of equation (2.3). The idea of this constraint is to use the brightness constancy equation but at higher levels of derivation. By using spatial derivations, we can write the following system for each pixel [30]:

\[
\begin{align*}
E_{xx} u + E_{yx} v + E_{tx} &= 0 \\
E_{xy} u + E_{yy} v + E_{ty} &= 0
\end{align*}
\]

Two equations for two unknowns is a solvable system. It is possible to use time derivation as well, which leads also to a system of two equations for two unknowns [12]:

\[
\begin{align*}
E_x u + E_y v + E_t &= 0 \\
E_{xt} u + E_{yt} v &= 0
\end{align*}
\]

This system contains 2 equations for 2 unknowns.

However, using a constraint with second-order derivatives requires to estimate discrete second-order derivatives, which are very sensitive to noise and introduce errors [9]. It follows the system may be ill-conditioned and not give any result.

The limitations of these methods is that they do not use any additional information on the images other than the constancy brightness equation. Thus, it is actually just one equation combined through different ways. The next section presents a constraint using extra information from the image.

### 2.2.3 Constraining the direction

One of the possible additional information is to use the direction of the camera. For pure translational motion, the camera follows a linear direction
which can be used to constrain the direction of the optical flow [4]. Thus, the problem to solve becomes:

\[
\min_{u,v} \quad (E_x u + E_y v + E_t)^2 \quad \text{subject to} \quad v = mu
\]

(2.4)

where \( m \) denotes the slope of the velocity vector and can be estimated in several ways. For example, it can be computed as \( m = \hat{u}/\hat{v} \) where \( \hat{u} \) and \( \hat{v} \) are the median of the velocity of an initial optical flow estimate. Another solution is to use a set of tracked points [4].

The efficiency of this method is shown on figure 2.2. The two images are equalized underwater images (a) and (b). The motion of the camera is purely translational, parallel to the vertical axis of the image. The motion direction is estimated through a tracker. The two lower images (c) and (d) compare the results obtained from the a traditional method to estimate optical flow with the results obtained with the constraint on the direction. It appears that the direction is very well estimated and it follows that the optical flow is smoother and more closer to the real motion.

### 2.3 Conclusions

This chapter illustrated the limits of the Brightness Constancy Equation and explained how the problem to solve is non-obvious. Additional constraints to equation (2.3) bring assumptions on images and therefore noise in the resulting optical flow. These constraints often smooth the image which is prejudicial on discontinuities, necessary for depth estimation. The rest of the document addresses this problem. The next chapter investigates two methods to reduce the noise when computing the optical flow, while chapter 4 presents a regularization method that smoothes the optical flow on homogeneous regions but preserves discontinuities.
Figure 2.2: Comparison between the normal optical flow (c) and the direction-constrained optical flow (d)
Chapter 3

Preliminary investigations

This chapter investigates two algorithms to improve the computation of optical flow from the Brightness Constancy Equation (2.3). The first one generalizes this equation, whereas the second one uses an averaging process in order to reduce the noise.

3.1 The Negahdaripour algorithm

This section presents the algorithm described by Negahdaripour and Yu [20]. The key-idea is to improve the Brightness Constancy Assumption used to build the Optical Flow Equation (2.3).

3.1.1 Principle

Negahdaripour and Yu proposed an approach to improve the Brightness Constancy Assumption. Contrary to equation (2.3), which assumes that the brightness remains constant over time, they propose an equation which takes into consideration the variation of the brightness through a linear transformation. It introduces a multiplier factor \( M(x, y, t) \) (taking into consideration a change of illumination or reflectance) and an offset term \( C(x, y, t) \) (taking into consideration changes like shadow or saturation). Equation (2.2) is rewritten as:
CHAPTER 3. PRELIMINARY INVESTIGATIONS

\[ E(x + dx, y + dy, t + dt) = M(x, y, t)E(x, y, t) + C(x, y, t) \quad (3.1) \]

Assuming that \( M \) is on the form \( M = 1 + \delta m \), the derivative of \( M \) is written:

\[ \frac{dM}{dt} = m_t = \lim_{t \to 0} \frac{\delta m}{\delta t} \]

On the same basis, if \( C \) is on the form \( C = \delta c \), we have:

\[ \frac{dC}{dt} = \alpha = \lim_{t \to 0} \frac{\delta c}{\delta t} \]

And finally, equation (2.3) becomes the Generalized Optical Flow Equation:

\[ E_x u + E_y v + E_t = Em_t + c_t \quad (3.2) \]

We use the Least-Square algorithm to solve this equation. Considering the optical flow constant on a small neighbourhood \( \Omega \) around the considered pixel, the least-square solution leads to equation:

\[
\begin{align*}
\int \int_{\Omega} \begin{pmatrix}
E_x^2 & E_x E_y & -E_x E & -E_x \\
E_x E_y & E_y^2 & -E_y E & -E_y \\
-E_x E & -E_y E & E_x^2 & E \\
-E_x & -E_y & E & 1
\end{pmatrix}
\begin{pmatrix}
u \\
v \\
m_t \\
c_t
\end{pmatrix}
\, dx
dy &= \int \int_{\Omega} \begin{pmatrix}
-E_x E_t \\
-E_y E_t \\
EE_t \\
E_t
\end{pmatrix}
\, dx
dy
\end{align*}
\]

(3.3)

This equation provides a solution for all points for which the 4x4 matrix (left term) is not singular.

3.1.2 Results

This algorithm was tested on the well-known images of the trees (figure 3.1); the motion on these images is a pure translational motion, parallel to the x-axis.
These images present interesting depths and motion and are very useful for tests. They were used for other tests as well.

Figure 3.2 presents the results. We compare the Optical Flow field computed from the Optical Flow Equation (2.3) and the field computed from the Generalized Optical Flow Equation (3.2).

We see on these images that the Generalized Optical Flow improves the results, especially in term of regularity. Actually, homogeneous regions are computed more regularly with the generalized algorithm, due to the presence of the multiplier factor.

3.2 Using information from several images

In order to reduce the noise on the Optical Flow computed on each frame of underwater images, it is interesting to compute an Averaged Optical Flow from results obtained on several frames.

3.2.1 Selection of the common region of the images

The aim is to average the optical flow from several frames, in order to reduce the random noise and emphasize homogeneous regions. It is obvious that this averaging can only be computed on the region common to all images (i.e.
Figure 3.2: Optical Flow for the trees image: Normal Optical Flow (a) and Generalized Optical Flow (b)

on the region which represents the same 3-D scene on all images). Thus, note that the more images we use, the smaller common region we have, as shown on figure 3.3. A compromise must be found between the number of frames to average and the size of the region on which is the optical flow is improved.

Figure 3.3: Selection of the common part between several images

Technically, assuming only that the distance of the camera to the scene stays constant, we compute the displacement between all two consecutive images
(e.g. using a set of tracked points). The next step consists in defining a reference frame (e.g. the first image) and computing the displacement of all images with respect to the reference frame. Therefore, knowing now the displacement of all images, it is possible to select the two relative furthest images on the direction $x$, and the two on the direction $y$. From these selected images, we extract first the size of the region, second the region itself on all images, as shown on figure 3.4.

![Selected common part on several images](image)

**Figure 3.4: Selected common part on several images**

### 3.2.2 Averaging

After selecting the common region on the frames, we compute the mean of the optical flow corresponding to this region. It should reduce the random noise which has a mean equal to 0. We investigate two approaches to perform that:

- Averaging by taking the *mean*
- Averaging by taking the *median*

### 3.2.3 Results

Test images are 11 frames of underwater images such as in figure 3.5
For practical reasons, these images have already been equalized in order to improve the contrast.

The first step is to select the common region between all images. Figure 3.6 shows the selected region for the first and the last images of the video sequence.
After having selected the common region, we compute the optical flow on these images. We use the direction-constrained optical flow described in chapter 2. The result on one frame is shown on figure 3.7.

![Optical Flow](image1.png) ![Depth estimation](image2.png)

Figure 3.7: Direction-constrained optical flow (a) and depth estimation (b)

The second image is the depth estimation of the scene. As explained in the introduction, it is the magnitude of the optical flow. Of course this depth is a relative estimation and needs a calibration of the camera to obtain a complete 3D reconstruction with real scale. On this depth map, brighter points are supposed to be closer to the camera.

We now average the optical flow on the common region to four frames. Results for the mean and the median are displayed on figure 3.8.
Figure 3.8: Averaged optical flow and depth estimation on 4 frames
Figure 3.9 shows the averaging on 11 frames.

![Mean optical flow](image1)

![Mean depth estimation](image2)

![Median optical flow](image3)

![Median depth estimation](image4)

Figure 3.9: Averaged optical flow and depth estimation on 11 frames

We see on these two sets of images the improvement of the averaging. It emphasizes the regions where the depth is higher (white regions) or deeper (black regions). Using the mean or median methods give similar results.
Chapter 4

The regularization

4.1 The aim of the regularization

We have seen in chapter 2 the need to find some additional constraints to equation (2.3). We described some additional constraints commonly found in the literature. However, these constraints often assume a priori knowledge on the image and bring noise in depth estimation. Moreover, they have the major drawback of smoothing the optical flow on all the image without discriminating depth discontinuities. We investigate in this chapter a method which smooths the flow in homogeneous regions while preserving depth discontinuities. These techniques are referred to as regularization [28].

4.1.1 Theory

The starting equation is still equation (2.3). The regularization consists in adding a term to this equation which takes into consideration the derivatives of the image and its optical flow. This term is called a regularizer and depends on the partial derivatives: \( V(\nabla E, \nabla u, \nabla v) \), where \( \nabla \) denotes the operator:

\[
\nabla = \left( \begin{array}{c} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \end{array} \right)
\]

\( E \) is the image brightness and \( (u, v) \) are the components of the optical flow.

By adding this term to the Optical Flow Equation (2.3), we have a unique equation for two unknowns: the problem is still under-constrained. Thus, it
is necessary to apply the *Spatial Smoothness Constraint* described in chapter 2 to the new equation to obtain the following equation to minimize [28]:

\[
\int \int_\Omega \left( (E_x u + E_y v + E_t)^2 + \alpha V(\nabla E, \nabla u, \nabla v) \right) dx dy \tag{4.1}
\]

where \( \Omega \) denotes the window in which the optical flow is considered constant and \( \alpha \) is the *regularization parameter* which allows to give more or less importance to the regularization term relative to the *Optical Flow Equation*.

### 4.1.2 The choice of the regularizer

One of the first regularizers was proposed by Horn and Schunck [13]:

\[
V(\nabla E, \nabla u, \nabla v) = |\nabla u|^2 + |\nabla v|^2 \tag{4.2}
\]

If we note \( s^2 = |\nabla u|^2 + |\nabla v|^2 \), then:

\[
V(s^2) = s^2 \quad \Rightarrow \quad \frac{dV(s^2)}{ds^2} = V'(s^2) = 1
\]

We can see that the first derivative of the regularizer does not depend on the discontinuities (i.e. the flow is also smoothed on the image boundaries) [29]. This will be more obvious in the next section, where the derivative of \( V \) is used to find the solution of the optical flow.

To correct this problem, several regularizers have been proposed. Two approaches are considered:

#### 4.1.2.1 Image-driven Smoothness

The smoothness term (i.e. the regularizer) depends on the derivatives of the image brightness. Hence, taking into consideration the image boundaries [29], Nagel proposed an image-driven regularizer [19] (improving the solution of Horn and Schunck) that reduces the term of regularization when it corresponds to large \( |\nabla E| \); thus the smoothness is reduced at image discontinuities:

\[
V(\nabla E, \nabla u, \nabla v) = \nabla u^T D(\nabla E) \nabla u + \nabla v^T D(\nabla E) \nabla v \tag{4.3}
\]

where

\[
D(\nabla E) = \frac{1}{|\nabla E|^2 + 2\lambda^2} (\nabla E^\perp \nabla E^\perp^T + \lambda^2 I) \tag{4.4}
\]
4.1.2.2 Flow-driven Smoothness

The regularizer depends on the derivatives of the flow. It takes into consideration the flow boundaries [29]. A flow-driven regularizer was proposed by Schnörr [23]:

\[ V(s^2) = \varepsilon s^2 + (1 - \varepsilon) \lambda^2 \sqrt{1 + \frac{s^2}{\lambda^2}} \]  

(4.5)

with \(0 < \varepsilon \ll 1\) and \(\lambda > 0\). The parameter \(\lambda\) can be seen as a contrast parameter [29]: if \(\lambda\) is small compared to the flow variation \(s^2\) (e.g. \(\lambda \ll s^2\)), then \(V(s^2) \approx \lambda s\). In the other hand, if \(\lambda \gg s^2\), then \(V(s^2) \approx \text{constant}\).

The choice of the smoothness term depends on the requirements of the application. Indeed, the flow discontinuities may not correspond to the image ones (for example a moving object with several textures). Therefore, when computing the optical flow, a flow-driven smoothness seems better, since an image-driven one may lead to an oversegmented flow.

So far, we have two types of equations:

- The flow-driven regularization (eq. 4.5) of Schnörr, where the regularizer is a function of \(s^2 = |\nabla u|^2 + |\nabla v|^2\):

  \[ V(\nabla E, \nabla u, \nabla v) = V(s^2) \]

  The similarity of the regularization (eq. 4.2) proposed by Horn and Schunck (it is a function of \(s^2 = |\nabla u|^2 + |\nabla v|^2\) as well) leads us to assimilate it to the regularization of Schnörr for the rest of the chapter, and especially for the computation of the solution.

- The image-driven regularization (eq. 4.3) of Nagel, where the regularizer is:

  \[ V(\nabla E, \nabla u, \nabla v) = \nabla u^T D(\nabla E) \nabla u + \nabla v^T D(\nabla E) \nabla v \]

  and is not a function of \(s^2 = |\nabla u|^2 + |\nabla v|^2\).

The next section investigates the approaches to solve the problem for the two different approaches.
4.2 Practical considerations to solve the problem

In this section, we lay down the theory to obtain the practical solvable equations for the two types of regularizers described in the previous section. We then introduce a multi-scale method and highlight some key parameters of the resulting algorithm.

4.2.1 Reaching a solvable system

4.2.1.1 The flow-driven regularization

Equation (4.1) is the starting equation, where:

\[ V(\nabla E, \nabla u, \nabla v) = V(s^2) \]

It is necessary to solve it by finding the minima. Using the Euler-Lagrange equations, we minimize this kind of function:

\[ \int \int_{\Omega} G(x, y, u, v, u_x, u_y, v_x, v_y) \, dx \, dy \]

which leads to the system:

\[
\begin{align*}
\frac{\partial G}{\partial u} - \frac{\partial}{\partial x} \left( \frac{\partial G}{\partial u_x} \right) - \frac{\partial}{\partial y} \left( \frac{\partial G}{\partial u_y} \right) &= 0 \\
\frac{\partial G}{\partial v} - \frac{\partial}{\partial x} \left( \frac{\partial G}{\partial v_x} \right) - \frac{\partial}{\partial y} \left( \frac{\partial G}{\partial v_y} \right) &= 0
\end{align*}
\]  \hspace{1cm} (4.6)

where \( u_x, u_y, v_x \) and \( v_y \) denote the partial derivatives of the optical flow components \( (u, v) \) with respect to \( x \) and \( y \).

Applying this system to our case, we obtain:

\[
\begin{align*}
div(V'(s^2) \nabla u) - \frac{1}{q} E_x (E_x u + E_y v + E_t) &= 0 \\
div(V'(s^2) \nabla v) - \frac{1}{q} E_y (E_x u + E_y v + E_t) &= 0
\end{align*}
\]  \hspace{1cm} (4.7)

where \( div \) denotes the operator:

\[ div \left( \begin{array}{c} a \\ b \end{array} \right) = \frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} \]

Obviously the first term of the equations corresponds to the regularization, whereas the second one corresponds to the Brightness Constancy Equation.
The key idea to solve the system is to consider the equations as the steady-state response of the optical flow equation. More precisely, by analogy with chemical area, the system can be seen as the diffusion-reaction system where the equations of the system (4.7) are the steady state ($\theta \to \infty$) of the following system [28]:

$$\begin{align*}
\frac{\partial u}{\partial t} &= \text{div}(V'(s^2) \nabla u) - \frac{1}{\alpha} E_x(E_x u + E_y v + E_t) \\
\frac{\partial v}{\partial t} &= \text{div}(V'(s^2) \nabla v) - \frac{1}{\alpha} E_y(E_x u + E_y v + E_t)
\end{align*}$$

(4.8)

Of course, the time denoted by $\theta$, the artificial time, is different from the real time $t$.

At this step, an iterative process is set up in order to discretize the system. Thus, at the iteration $k$ we consider we have estimated the optical flow $u_{ij}^k$ and $v_{ij}^k$ for the pixel $(i,j)$. By discretization, the system (4.8) is equivalent to compute:

$$\begin{align*}
u_{ij}^{k+1} - u_{ij}^k &= A_{ij}^k - \frac{1}{\alpha} E_{x,ij}(u_{ij}^k + E_{y,ij}v_{ij}^k + E_{t,ij}) \\
v_{ij}^{k+1} - v_{ij}^k &= B_{ij}^k - \frac{1}{\alpha} E_{y,ij}(u_{ij}^k + E_{y,ij}v_{ij}^k + E_{t,ij})
\end{align*}$$

(4.9)

where $\tau$ is the step size in the direction of $\theta$ and $A_{ij}^k$ and $B_{ij}^k$ represent the discretization of the divergence term (i.e. the regularization term).

The regularizer is approximated by the terms [29]:

$$\begin{align*}
A_{ij}^k &= \sum_{mn \in N(i,j)} \frac{V^k_{mn}(s_{mn}^2) + V^k_{ij}(s_{ij}^2)}{2} (u_{mn}^k - u_{ij}^k) \\
B_{ij}^k &= \sum_{mn \in N(i,j)} \frac{V^k_{mn}(s_{mn}^2) + V^k_{ij}(s_{ij}^2)}{2} (v_{mn}^k - v_{ij}^k)
\end{align*}$$

where $N(i,j)$ denotes the neighborhood of pixel $(i,j)$ of typically 4 or 8 pixels. This is the typical approximation of the divergence.

Finally, from (4.9), the components of the optical flow $(u,v)$ can be updated using the expressions:

$$\begin{align*}
u_{ij}^{k+1} &= u_{ij}^k + \tau A_{ij}^k - \frac{(\tau/\alpha)E_{x,ij}(E_{y,ij}v_{ij}^k + E_{t,ij})}{1 + (\tau/\alpha)E_{x,ij}^2} \\
v_{ij}^{k+1} &= v_{ij}^k + \tau B_{ij}^k - \frac{(\tau/\alpha)E_{y,ij}(E_{x,ij}u_{ij}^k + E_{t,ij})}{1 + (\tau/\alpha)E_{y,ij}^2}
\end{align*}$$

(4.10)
4.2.1.2 The image-driven regularization

By analogy with the flow-driven regularization, the system (4.7) is written as:

\[
\begin{align*}
\text{div}(D(\nabla E)\nabla u) - \frac{1}{\alpha} E_x (E_x u + E_y v + E_t) &= 0 \\
\text{div}(D(\nabla E)\nabla v) - \frac{1}{\alpha} E_y (E_x u + E_y v + E_t) &= 0
\end{align*}
\]

(4.11)

Therefore, the diffusion reaction system becomes:

\[
\begin{align*}
\frac{\partial u}{\partial t} &= \text{div}(D(\nabla E)\nabla u) - \frac{1}{\alpha} E_x (E_x u + E_y v + E_t) \\
\frac{\partial v}{\partial t} &= \text{div}(D(\nabla E)\nabla v) - \frac{1}{\alpha} E_y (E_x u + E_y v + E_t)
\end{align*}
\]

(4.12)

The discretized system to solve for the components of the optical flow (eq. 4.10) is the same, except that the terms \( A_{ij}^k \) and \( B_{ij}^k \) are estimated in a different way.

In this case, for each pixel \((i, j)\), the approximation depends on the term \( D(\nabla E) \) from equation (4.4), which has the form:

\[
D(\nabla E)_{ij} = \begin{pmatrix} d_{ij} & f_{ij} \\ f_{ij} & e_{ij} \end{pmatrix}
\]

Thus, the regularizer is approximated by expression [1]:

\[
A_{ij}^k = \frac{d_{i+1,j} + d_{i,j}}{2}(u_{i+1,j}^k - u_{ij}^k) + \frac{d_{i-1,j} + d_{i,j}}{2}(u_{i-1,j}^k - u_{ij}^k) + \frac{e_{i,j+1} + e_{i,j}}{2}(u_{i,j+1}^k - u_{ij}^k) + \frac{e_{i,j-1} + e_{i,j}}{2}(u_{i,j-1}^k - u_{ij}^k) + \frac{f_{i+1,j+1} + f_{ij}}{2} \frac{u_{i+1,j+1}^k - u_{ij}^k}{2} + \frac{f_{i-1,j-1} + f_{ij}}{2} \frac{u_{i-1,j-1}^k - u_{ij}^k}{2} - \frac{f_{i+1,j-1} + f_{ij}}{2} \frac{u_{i+1,j-1}^k - u_{ij}^k}{2} - \frac{f_{i-1,j+1} + f_{ij}}{2} \frac{u_{i-1,j+1}^k - u_{ij}^k}{2}
\]

and \( B_{ij}^k \) is the dual expression for \( v \).

4.2.2 The multi-scale process

The solution of equation (4.1) in this way is heavily dependent on the initialization of the values \( u_{ij}^0 \) and \( v_{ij}^0 \) at the beginning of the process. In order to achieve acceptable results, we set up a coarse-to-fine process to give good initializations for the algorithm [1]. Instead of applying the algorithm on the
original images \( E_1 \) and \( E_2 \), they are replaced by smoothed and size-reduced images \( E'_0 \) and \( E'_2 \) at scale \( s_0 \). We initialize for that scale the optical flow to zero. The next step is to re-compute the algorithm with images \( E'_1 \) and \( E'_2 \) at scale \( s_1 \) and initialize the components of the optical flow with the results of the previous iteration \( s_0 \). We execute this process at different scales \( s_i \) until we reach the original size of images:

\[
s_0 < s_1 < s_2 < \ldots < s_i < \ldots < s_{n-1} < s_n = 1
\]

The results of each scale gives initializations for the next scale and last results are the final results.

### 4.2.3 Key parameters and their influence

Several parameters are of prime importance in the regularization algorithm. We detail the importance of each of these parameters in order to understand the role of each of them and to use the algorithm efficiently.

#### 4.2.3.1 The weight of the regularization

The parameter \( \alpha \) in equation (4.1) allows to give more or less importance to the regularization term relative to the optical flow term equation. If \( \alpha \approx 0 \), it is equivalent to use the brightness constancy equation without any regularization. On the images we see later, a typical empiric value of \( \alpha \) is 10.

#### 4.2.3.2 The number of scales

The number of scales is important in the multi-scale approach of the algorithm. The more scales are used, the more precise the results are. However, a compromise between precision and computational cost is necessary. For example, for the 233x256 pixels image we use for tests, 12 scales with the first image as 1/10 of the original size is a good compromise.

#### 4.2.3.3 The step of actualization

This parameter corresponds to the speed of actualization relative to the artificial time \( \theta \). In the algorithm, we solve equations of the type:

\[
\frac{u^{k+1} - u^k}{\tau} = f(u^k, v^k)
\]
Thus, high values of $\tau$ lead to a quicker and less precise actualization of $u$ and $v$ but with a relative low computational cost, whereas low values of $\tau$ lead to a slower and more precise actualization with a higher computational cost. Again for our case, a good compromise between these is for example $\tau = 1/10$. Note that the solution may not converge for high values of $\tau$.

4.2.3.4 The stop criterion

For each scale, the system (4.8) is computed until it converges to a solution precise enough to give a good initialization for the next scale. As the actualization gradually converges to the ideal solution, we can compute the actualization on the optical flow components $(u_{ij}, v_{ij})$ at pixel $(i, j)$ between iterations $k$ and $k + 1$ as:

$$(u_{ij}^{k+1} - u_{ij}^k)^2 + (v_{ij}^{k+1} - v_{ij}^k)^2$$

The algorithm stops when the mean of the squared difference of the improvement on the whole image is below a given threshold $\gamma$:

$$\frac{1}{N_i N_j} \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} (u_{ij}^{k+1} - u_{ij}^k)^2 + (v_{ij}^{k+1} - v_{ij}^k)^2 \leq \gamma$$

where $(N_i, N_j)$ is the size of the image at the current scale.

Again, a compromise must be found between precision and computational cost. A first optimization can be done by considering that the last scale gives the final result and thus could be considered the more important. So we shall impose a stronger stop criterion than the other scales. For our case, a typical value is 0.1; and for the last scale is 0.01.

4.3 Experimental results

We present in this section some results obtained with the regularization on the well-known images of the trees (figure 3.1) and show the results we obtain on some underwater images (figure 1). Let us remember that the motion on the trees images is a pure translational motion, parallel to the x-axis:
In order to observe the improvement brought by the regularization, we compute the optical flow without any regularization term (corresponding to $\alpha = 0$). The resulting optical flow is on figure 4.2.
CHAPTER 4. THE REGULARIZATION

Figure 4.2: Optical Flow without regularization

In order to compare the influences of the regularizers, we display on figure 4.3 the resulting optical flow for each type of regularizer.
Figure 4.3: Resulting Optical Flow for the 3 proposed regularizations
We see on these images the improvement of the regularization, especially in terms of convergence: there are less outliers than on figure 4.2. Moreover, the flow is more homogeneous (for example in the background) and the discontinuities are clearer. However, it is important to note that on these images, the three regularization types presented in this document have similar results.

Logically, as shown on figure 4.4, these results depend on the weight of the regularization term in equation (4.1). We have considered for these tests the regularization proposed by Nagel.

![Figure 4.4: Role of the weight of the regularization](image)

The influence of the parameter $\tau$ in the algorithm is illustrated on figure 4.5. We consider the regularization proposed by Nagel.
Figure 4.5: Influence of the actualization parameter
As expected, we see in these results that the optical flow is more regularized for smaller values of $\tau$. As explained in the introduction, the depth estimation is the magnitude of the optical flow. Of course this depth is a relative estimation and needs a calibration of the camera to obtain a complete 3D reconstruction with real scale. We see on these images that the depth is more emphasized for smaller values of $\tau$. Note the loss of magnitude on the optical flow for small values of $\tau$.

To conclude this part, let us now observe the results on underwater images (figure 1). If we consider the regularization proposed by Nagel and if we set the parameters of the algorithm to $\alpha = 10$ and $\tau = 1/10$, we obtain the optical flow shown on figure 4.6. Note that figure 1 and other images on which we work are equalized. The second image is the magnitude of the estimated optical flow (i.e. the depth map where brighter points are supposed to be closer to the camera).

![Optical Flow with regularization](image1.png)  ![Depth estimation](image2.png)

Figure 4.6: Depth estimation of a trawling mark

The 3D representation of this image is presented at the figure 4.7.
Figure 4.7: 3D representation of the trawling mark
Conclusion

Results on regularization are promising and give a good idea of the 3D representation (figure 4.7) of the underwater image given in introduction (figure 1). These results must still be improved by investigating other regularizers than those proposed for equation (4.1) and other ways to solve this equation.

One possible investigation is to couple an efficient constraint (the principle is described in chapter 2), which over-constrains the problem, with a regularizer, which preserves discontinuities of the image. For example, by coupling the constraint on the direction (chapter 2) with regularization, equation (4.1) becomes:

$$\min_{u,v} \int_\Omega \left( (E_x u + E_y v + E_t)^2 + \alpha V(\nabla E, \nabla u, \nabla v) \right) dxdy \text{ subject to } v = mu$$

By replacing the constraint in the first term, we have:

$$\min_{u} \int_\Omega \left( (E_x u + E_y mu + E_t)^2 + \alpha V(\nabla E, \nabla u) \right) dxdy$$

which leads to a unique Euler-Lagrange equation instead of the system (4.6):

$$\frac{\partial G}{\partial u} - \frac{\partial}{\partial x} \left( \frac{\partial G}{\partial u_x} \right) - \frac{\partial}{\partial y} \left( \frac{\partial G}{\partial u_y} \right) = 0$$

where

$$G(x, y, u, u_x, u_y) = \left( (E_x + E_y m) u + E_t \right)^2 + \alpha V(\nabla E, \nabla u)$$

It seems important to combine constraints and regularization in order to take advantages of the both.
Moreover, Miguel Arredondo is investigating other constraints [4] like the *differential epipolar constraint*.

Finally, the traineeship was very useful to emphasize the role of regularization, especially on discontinuities of 3D reconstruction of underwater images and will be included as part of his future work by Miguel Arredondo for his PhD thesis.

To conclude, this traineeship was personally and technically enlightening. From a **technical point of view**, it helped me to investigate the domains of *image processing and computer vision*, especially the 3D reconstruction, which was a new domain for me. This experience gave me the opportunity to work in a research team in a university and to discover the researcher work. From a **personal point of view**, my work was a part of a team work included in the european project AMASON and was in great collaboration with Miguel Arredondo. Furthermore, this experience abroad and in a foreign language will be for sure very useful for my future.

In this context and after this enriching experience, I decided to continue in the domain of research in *Signal and Image Processing* by starting a PhD next year in France on *vocal tract modeling*.
Bibliography


