**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**
PROGRAMA DE MAESTRÍA Y DOCTORADO EN INGENIERÍA
INGENIERÍA ELÉCTRICA – PROCESAMIENTO DIGITAL DE SEÑALES

# ESTIMACIÓN DE FLUJO ÓPTICO POR MEDIO DE LA TRANSFORMADA DE HERMITE

TESIS
QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN INGENIERÍA

PRESENTA:
ERNESTO MOYA ALBOR

TUTOR PRINCIPAL
DR. BORIS ESCALANTE RAMÍREZ, FACULTAD DE INGENIERÍA
COMITÉ TUTOR
DR. FRANCISCO GARCÍA UGALDE, FACULTAD DE INGENIERÍA
DR. JOSÉ LUIS MARROQUÍN ZALETA, CENTRO DE INVESTIGACIÓN EN MATEMÁTICAS

MÉXICO, D. F. MAYO 2013

**JURADO ASIGNADO:**

Presidente: DR. FRANCISCO GARCÍA UGALDE

Secretario: DR. PABLO ROBERTO PEREZ ALCÁZAR

Vocal: DR. BORIS ESCALANTE RAMÍREZ

1er. Suplente: DR. ENRIQUE PABLO VALLEJO VENEGAS

2do. Suplente: DR. JOSÉ LUIS MARROQUÍN ZALETA

Lugar donde se realizó la tesis: FACULTAD DE INGENIERÍA, C.U.

**TUTOR DE TESIS:**

DR. BORIS ESCALANTE RAMÍREZ

-------------------------------------------------
**FIRMA**

*Dedicated to ...*

Thesis advisor
**Dr. Boris Escalante Ramírez**

Author
**M. en I. Ernesto Moya Albor**

### Optical Flow Estimation using the Hermite Transform

# Abstract

This thesis describes a new method to estimate the heart's motion in computer tomography images with the inclusion of a bio–inspired image representation model. Our proposal is based on the polynomial decomposition of each of the images using the steered Hermite transform as a representation of the local characteristics of images from an perceptual approach within a multiresolution scheme.

The Hermite transform is a model that incorporates some of the more important properties of the first stages of the human visual system, such as the overlapping Gaussian receptive fields, the Gaussian derivative model of early vision and the multiresolution analysis.

We propose an approach for optical flow estimation that incorporates image structure information extracted from the steered Hermite coefficients, that is later used as local motion constraints in a differential estimation method that involves several of the constraints seen in the current differential methods, which allows obtaining accurate flows.

Considering the importance of understanding the movement of certain structures such as left ventricular and myocardial wall for better medical diagnosis, our main goal is to find an estimation method useful to assist diagnosis tasks in computer tomography images.

# Acknowledgments

This work was supported by UNAM grants IN113611 and IX100610. E. Moya-Albor thanks to Universidad Nacional Autónoma de México for the granted doctoral scholarship.

# Table of Contents

# Chapter 1

# Introduction

> The fool who thinks he is wise is just a fool. The fool who knows he is a fool is wise indeed.
>
> ---
>
> Sidhartha Gautama (563–483 b.c).
> Dhammapada.

Congestive heart failure has increased worldwide due to both left ventricular and right ventricular failure. In the United States, about 5 million patients suffer from this problem and about 500,000 patients develop this condition each year. Previously it was due to the left ventricle (LV) pumping blood inefficiently (systolic ventricular failure). More recently, there has been emphasis on diastolic ventricular failure, where the systolic function appears to be normal, but diastolic ventricular function is impaired, being the cause of 50% of congestive heart failure in these patients. In order to develop better treatments for congestive heart failure, it is necessary to first understand the basic physiology and movement of both normal and abnormal ventricular relaxation and contraction [18].

Understanding the movement of certain structures such as left ventricle and myocardial wall is fundamental for better medical diagnosis. This thesis describes a new method to estimate the heart's motion in computed tomography (CT) images. We propose a differential approach to optical flow estimation using the steered Hermite transform, which is a tool that performs a decomposition of the images into visual patterns that are relevant to the human vision system (HSV), such as directional edges, textures, etc. The Hermite transform is an image representation model that mimics some of the more important properties of early vision such as local processing and the Gaussian derivative models of receptive fields [85, 52, 51].

Although physicians are often interested in the boundaries of structures, our method can also estimate motion in homogeneous areas that are also of interest for medical diagnosis.

The Hermite transform uses a Gaussian window to extract local information from a image. This information is expanded in terms of a family of orthogonal polynomials. There are reasons why a Gaussian window is used. From a perceptual standpoint the Gaussian window is a good model of the overlapping receptive fields found in physiological experiments [67, 44]. According to the scale-space theory, the Gaussian window minimizes the uncertainty product in the spatial and frequency domain [83]. It is also mathematically tractable and there is an approximation of the Gaussian function in the discrete case. The polynomials used to approximate the local information are determined by the analysis window. For the case of a Gaussian window these polynomials involve the Gaussian derivative operators, found in the psychophysical modeling of the HSV [50, 12] which provides *natural operators* in agreement with the theory scale-space [43]. In [85, 86, 87] it is showed that these functions model the measured receptive field data more accurately than the Gabor functions do, with the additional advantage of being orthogonal at the same location of analysis. Like the receptive fields, both Gabor functions and Gaussian derivatives are spatially local and consist of alternating excitatory and inhibitory regions within a decaying envelope.

A multiresolution decomposition using the Hermite transform can be obtained through a pyramidal scheme [24, 73], where the image is decomposed into a number of band-pass or low-pass subimages, which are then subsampled in proportion to their spatial resolution offering a way of relating image structures between different scales. Like [78] where the retinal image is represented through a wavelet-like transform with difference of two Gaussian functions (DoG) as basis filters. We use the Hermite transform for the multiscale decomposition to emulate the behavior of retinal ganglion cell receptive fields that can be described using the DoG at different scales [66]. A multiresolution approach allows using small windows to better detect fine details and large windows to analyze low resolution objects, which is one of the main properties of the HVS.

A rotated version of the Hermite transform provides a very efficient representation of oriented patterns which enables an adaptation to local orientation content at each window position over the image, indicating the direction of one-dimensional pattern. The steered Hermite coefficients are obtained projecting the cartesian Hermite coefficients onto one-dimensional coefficients on an axis that makes an angle $\theta$ with the $x$ axis, where the angle $\theta$ represents the direction of maximum energy. The Hermite filters form a steerable basis because they are products of polynomials with a radially symmetric Gaussian window. Filters of increasing order analyze successively higher radial frequencies and filters of the same order and different (directional) index distinguish between different orientations in the image, which is relevant from a psychophysical approach considering that the HSV calculates Gaussian derivatives at different scales and different directions.

Optical flow estimation methods calculate apparent velocities that can be associated with a variation of brightness patterns in a sequence of images. In general these methods consider that motion is due to the change in position of different cardiac structures and not due to the relative motion between the observer and the scene or to variations in scene illumination. Many of the differential optical flow approaches are based on the work of Horn and Schunck [38], and Lucas and Kanade [49], which incorporated certain constraints in order to handle the *Ill-posed Opening problem.*

We propose an approach for optical flow estimation that involves several of the constraints seen in the current differential methods, which allows obtaining an accurate optical flow [16, 17, 63]. We use local image constraints as mentioned in [38, 63] and assume that the flow is piecewise smooth as in [16, 15, 17]. Similarly to [11, 17], the proposed optimization function is robust to outliers and a multiresolution strategy was implemented to handle large displacements as mentioned in [11, 5, 53, 54].

Our contribution includes local restrictions using the steered Hermite transform as a representation of the local image characteristics from an perceptual approach. In [63] the effect of different local restrictions on the data term (intensity, gradient, Hessian, Laplacian) is described, but in our approach the steered Hermite coefficients allow including polynomial decomposition of the image and take these parameters as constraints that include intensity and higher order derivatives, which are useful to analyze the image in a similar way as is done by the HSV. The use of Gaussian derivatives allows incorporating image structure information from neighboring pixels that is robust to noise [49, 10]. This feature is incorporated in a global differential functional that allows obtaining dense flow fields [38, 57].

The nature of CT images compels the estimation of optical flow algorithms to be robust in noisy environments, this complicates the accurate calculation of the derivatives used in classical differential methods that usually uses a smoothing pre-filter. Using Gaussian derivatives the estimation is more robust to noise because the smoothing is implicitly included in the proposed functional applied at different scales of analysis.

In order to expose the patient to less radiation CT cardiac images are captured in a few discrete times of the cardiac cycle. Therefore there are large displacements of the structures between two consecutive images. Including a Gaussian window in the polynomial decomposition of the Hermite transform allows incorporating information from neighboring pixels that helps finding optical flow in areas that change dramatically from one image to another. Moreover the inclusion of a multiresolution analysis in the proposed functional also helps dealing with large displacements. In this sense, motion discontinuities are also dealt with using filters similar to those found in HSV.

In summary, our proposal includes a multiresolution HSV-inspired perspective to optical flow estimation in CT images. The steered Hermite transform extracts the local image characteristics, that are later used as motion constraints in a differential estimation method, which allows obtaining accurate flows.

This thesis is structured as follows. In Chapter 2 we present the theory and the mathematical foundations of the cartesian Hermite transform, which allows to consider the Hermite transform as a good model for describing the images. In Section 2.2 we introduce the rotated version of the Hermite transform, which allows locally adapting the analysis to representative visual patterns. Related work and elements of the state of the art of optical flow estimation are shown in Chapter 3. A brief introduction to the cardiac anatomy is shown in Chapter 2. Chapter 4 presents our method differential optical flow estimation using the steered Hermite coefficients as the main tool of analysis, which includes elements that allow obtaining a robust optical flow in a perceptual framework. The results obtained are reported in Chapter 5 and conclusions are given in Chapter 6. Appendixes A and B give an introduction to CT images and its application to cardiac images. Appendix C provides a proof for a Hermite transform property that allows us to simplify the solution of the proposed functional. Finally, some simplifications of the numerical solution are shown in the Appendixes D and E.

# Chapter 2

# The Hermite transform as a model of image representation

> The seeker after truth should be humbler than the dust.
> The world crushes the dust under its feet, but the seeker after truth should so humble himself that even the dust could crush him.
> Only then, and not till then, will he have a glimpse of truth
>
> Mahatma Gandhi (1869–1948).

The Hermite transform [52, 51] is a special case of polynomial transform, it can be considered as an image description model. In order to calculate the Hermite transform, the original image $L(x, y)$ (where $(x, y)$ are the coordinates of the pixels) is located at various positions multiplying $L(x, y)$ by the window function $v^2(x - x_0, y - y_0)$ at positions $(x_0, y_0)$ that conform the sampling lattice $S$:

$$L_v(x - x_0, y - y_0) = L(x, y)v^2(x - x_0, y - y_0) \tag{2.1}$$

By replicating the window function over the sampling lattice, we can define a weight function different from zero for all $(x, y)$:

$$V(x, y) = \sum_{(x_0,y_0)\in S} v^2(x - x_0, y - y_0) \tag{2.2}$$

Therefore the original image is represented within the window by:

$$L(x, y) = \frac{1}{V(x, y)} \sum_{(x_0,y_0)\in S} L(x, y)v^2(x - x_0, y - y_0) \tag{2.3}$$

The local information for each analysis window $L_v(x - x_0, y - y_0)$ is expanded in terms of a family of orthogonal polynomials $G_{m,n-m}(x, y)$ of $m$ degree in $x$ and $(n - m)$ in $y$:

$$L_{m,n-m}(x_0, y_0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[L(x, y)v^2(x - x_0, y - y_0)\right] G_{m,n-m}(x - x_0, y - y_0)\, dxdy \tag{2.4}$$

The polynomials $G_{m,n-m}(x, y)$ used to approximate the information within the window are determined by the analysis window and satisfy the orthogonality condition:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} v^2(x, y)G_{m,n-m}(x, y)G_{l,k-l}(x, y)\, dxdy = C_{nk}\delta_{nk}\delta_{ml} \tag{2.5}$$

for $n, k = 0, \ldots, \infty$, $m = 0, \ldots, n$ and $l = 0, \ldots, k$; where $\delta_{nk}$ denotes the Kronecker function ($\delta_{nk} = 1$ for $n = k$ and $\delta_{nk} = 0$ for $n \neq k$). The polynomials are orthonormal after a normalization by the factor $C_{nk}$ or when $C_{nk} = 1$.

For example, for a decomposition up to order $N = 2$ where $N$ is the polynomial degree, we have that $n = 0, 1, \ldots, N$ and $m = 0, \ldots, n$ are the indexes of the polynomials used in the decomposition as defined in Table 2.1.

## 2.1 Cartesian Hermite transform

We mentioned above that the polynomials $G_{m,n-m}(x, y)$ are orthogonal with respect to the window function. From a perceptual standpoint and according to the

| $\boldsymbol{n}$ | $\boldsymbol{m = 0, 1, \ldots, n}$ | $\boldsymbol{G_{m,n-m}(x, y)}$ |
|---|---|---|
| $n = 0$ | $m = 0$ | $G_{0,0-0}(x, y) = G_{0,0}(x, y)$ |
| $n = 1$ | $m = 0$ | $G_{0,1-0}(x, y) = G_{0,1}(x, y)$ |
|  | $m = 1$ | $G_{1,1-1}(x, y) = G_{1,0}(x, y)$ |
| $n = 2$ | $m = 0$ | $G_{0,2-0}(x, y) = G_{0,2}(x, y)$ |
|  | $m = 1$ | $G_{1,2-1}(x, y) = G_{1,1}(x, y)$ |
|  | $m = 2$ | $G_{2,2-2}(x, y) = G_{2,0}(x, y)$ |

Table 2.1: Indexes of the polynomials used in a decomposition of order $N = 2$ ($n = 0, 1, \ldots, N$ and $m = 0, \ldots, n$).

scale-space theory, our option would be a Gaussian window (Fig. 2.1):

$$v(x, y) = \frac{1}{\sigma\sqrt{\pi}} \exp^{\left(-\frac{(x^2+y^2)}{2\sigma^2}\right)} \tag{2.6}$$

where the normalization factor defines a unitary energy for $v^2(x, y)$.

The reasons why a Gaussian window is of interest are numerous. First, the theory related to the Gaussian function is mathematically tractable. In addition, there is discrete approximation of the Gaussian function, which is useful for implementation purposes. Secondly, adjacent Gaussian windows separated by twice the standard deviation $\sigma$ are a good model of overlapping receptive fields found in physiological experiments [67, 44]. Third, a signal decomposition into orthogonal polynomials with respect to the Gaussian window involves Gaussian derivative operators (Fig. 2.2), which are found in the psychophysical modeling of the HSV [50, 12, 85, 86, 87]. In the human eye the retina (Fig. 2.3) is made up of several layers of different kinds of cells. The sensitive rods and cones are at the back of the retina, facing away from the light. To reach them the light first has to travel through blood vessels, nerve fibers and then several layers of retinal nerve cells. In contrast, in the octopus eye its rods and cones face forwards, and blood supply and nerves come from behind. For Richard Gregory [35] the ”design flaw” in the human eye is less serious than it seems. He notes that the blood vessels and nerve fibers visible with an ophthalmoscope skirt around the fovea centralis where visual acuity is crucial, so they interfere very little with precise vision. It seems that we have a greater need for high acuity vision than

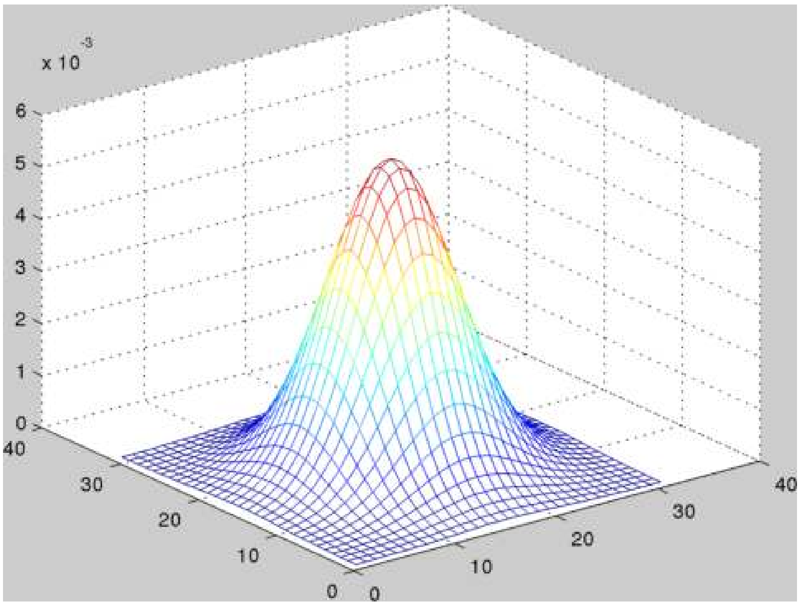


Figure 2.1: Gaussian window $v(x, y)$.

the octopus does and accordingly our eyes have a richer supply of nutrients than the octopus does. The retinas of our eyes gets nutrition both from the blood vessels in front and from the choroid layer directly behind the rods and cones. The blood vessels in front supplement the supply from behind. The retina responds to changes of light and not directly to light, this indicates that the eye works like a spatial differentiator of the scene. The ganglion cells (Fig. 2.4) are known to be spot detectors that have receptive fields that look like $[-1 \quad 2 \quad 1]$ or $[1 \quad -2 \quad 1]$ so that they sense the local second derivative (Fig. 2.5) [65]. The ganglion cells are involved in color vision, they compare signals from many different cones. Finally, the Gaussian window minimizes the product of uncertainty in the spatial and frequency domains [83].

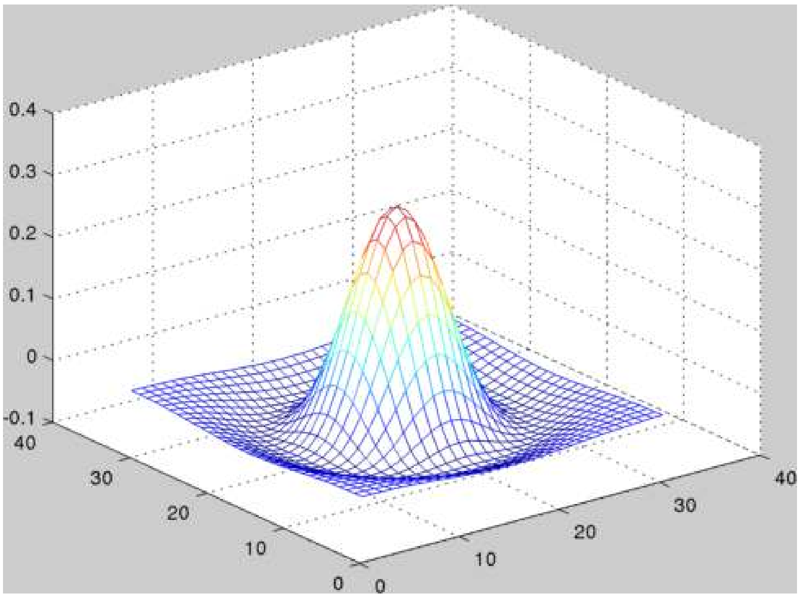


Figure 2.2: Second derivative of a 2D Gaussian.

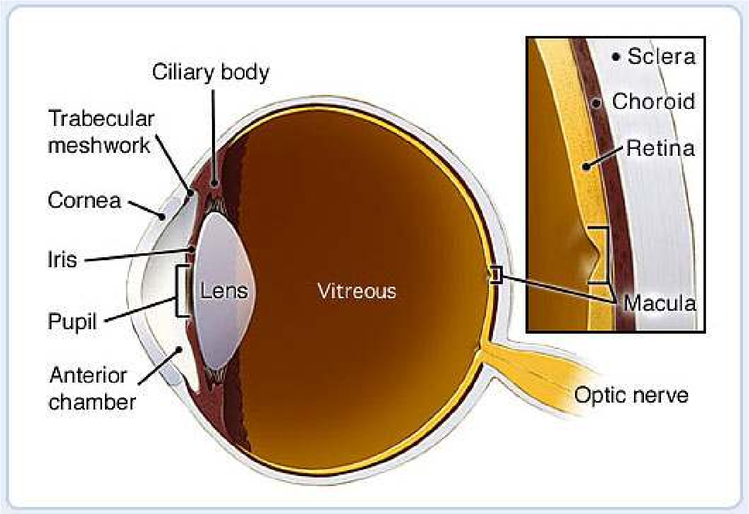


Figure 2.3: Anatomy of the human eye [4].

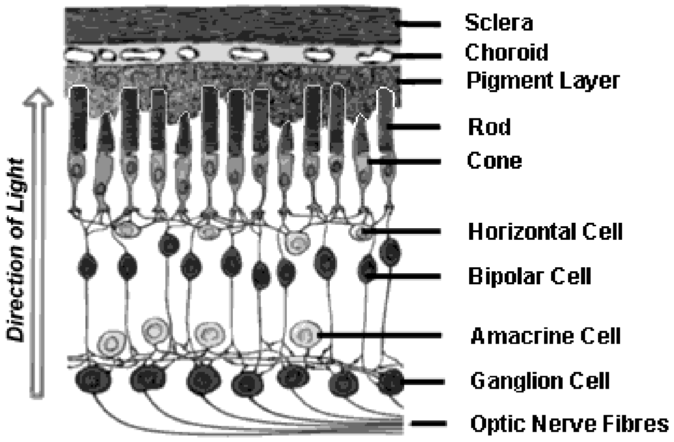


Figure 2.4: A cross section of the retina. The light has to pass through the ganglion–cell and bipolar–cell layers before it gets to the rods and cones [22].

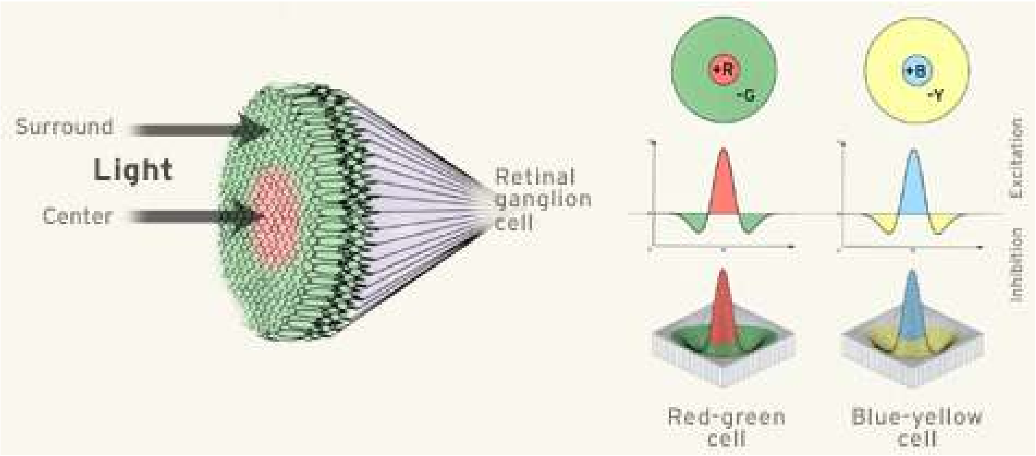


Figure 2.5: Receptive fields of the ganglion cells [23].

With a Gaussian window function, the associated orthogonal polynomials are the Hermite polynomials [74]:

$$G_{m,n-m}(x,y) = \frac{1}{\sqrt{2^n m!(n-m)!}} H_m\left(\frac{x}{\sigma}\right) H_{n-m}\left(\frac{y}{\sigma}\right) \tag{2.7}$$

where $H_n(x)$ denotes the *nth* Hermite polynomial given by Rodrigues' formula [1]:

$$H_n(x) = (-1)^n \exp^{x^2} \frac{d^n}{dx^n} \exp^{-x^2} \tag{2.8}$$

Substituting $\mathcal{G}_{\sigma^2}(x)$ (with variance $\sigma^2$) by $\exp^{-\frac{x^2}{\sigma^2}}$ in Eq. (2.8), we obtain the generalized Hermite polynomials with respect to the Gaussian function $\mathcal{G}_{\sigma^2}(x)$:

$$H_n\left(\frac{x}{\sigma}\right) = (-1)^n \mathcal{G}_{\sigma^2}^{-1}(x) \frac{d^n}{dx^n} \mathcal{G}_{\sigma^2}(x) \tag{2.9}$$

Thus, from Eq. (2.4) the polynomial coefficients $L_{m,n-m}(x,y)$ are calculated by a convolution of the original image $L(x,y)$ with the filter function $D_{m,n-m}(x,y)$, followed by a subsampling $(T)$ at position $(x_0, y_0)$ of the sampling lattice $S$:

$$L_{m,n-m}(x_0,y_0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} L(x,y) D_{m,n-m}(x-x_0, y-y_0)\, dxdy \tag{2.10}$$

$$n = 0, 1, \ldots, \infty$$

$$m = 0, 1, \cdots, n$$

where $D_{m,n-m}(x,y) = G_{m,n-m}(-x,-y)v^2(-x,-y)$.

$L_{m,n-m}$ are the Hermite coefficients and $(n-m)$ and $m$ denote the analysis order in $x$ and $y$ direction respectively. In Fig. 2.6 we show the analysis process of the Hermite transform.

In Fig. 2.7 we show a typical graphical distribution of the Hermite coefficients in function of the order of the polynomials in the $x$ and $y$ directions.

The filter function $D_{m,n-m}(x,y)$ are separable because the Gaussian window is rotationally symmetric, therefore:

$$D_{m,n-m}(x,y) = D_m(x) D_{n-m}(y) \tag{2.11}$$

The Hermite filters can be computed by:

$$D_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}} \frac{1}{\sigma\sqrt{\pi}} H_n\left(\frac{x}{\sigma}\right) \exp^{-\frac{x^2}{\sigma^2}} \tag{2.12}$$

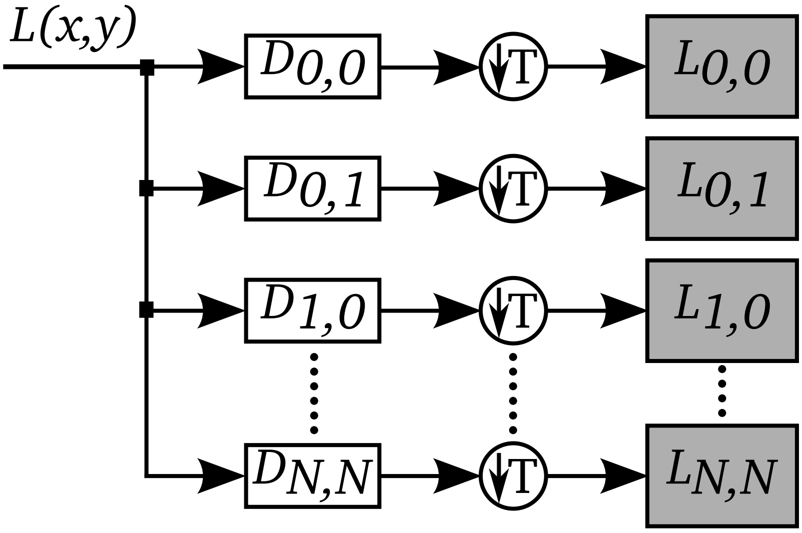
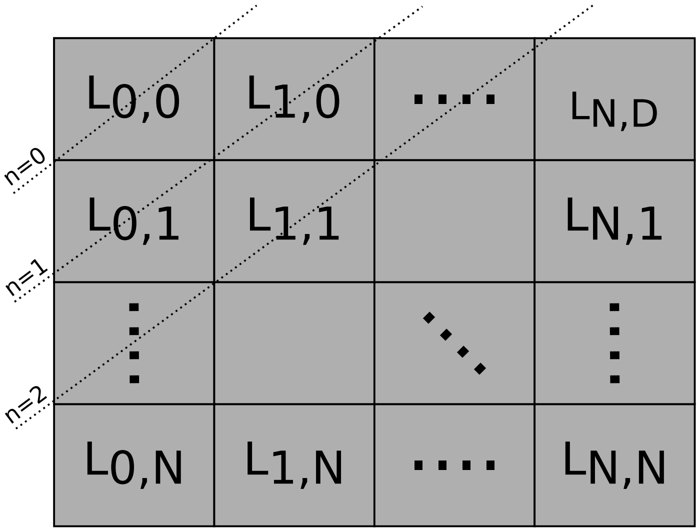


Figure 2.6: Analysis process of the Hermite transform.



| | | | |
|---|---|---|---|
| $L_{0,0}$ | $L_{1,0}$ | .... | $L_{N,D}$ |
| $L_{0,1}$ | $L_{1,1}$ | | $L_{N,1}$ |
| $\vdots$ | | $\ddots$ | $\vdots$ |
| $L_{0,N}$ | $L_{1,N}$ | .... | $L_{N,N}$ |

Figure 2.7: Typical graphical distribution of the Hermite coefficients.

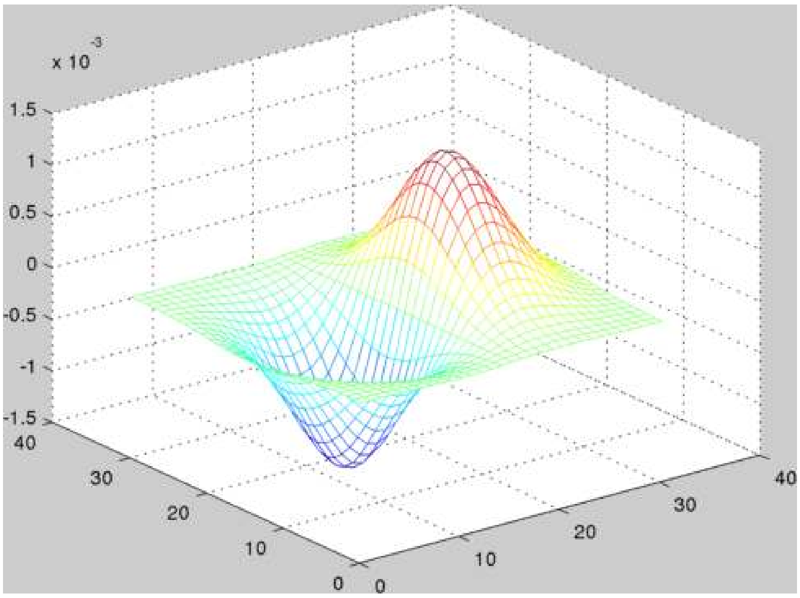


Figure 2.8: First derivative of a 2D Gaussian.

The analysis functions of the Hermite transform are similar to Gaussian derivatives (Fig. 2.8), which, as argued before, are good models of some of the important retinal and cortical cells of the HSV. They model filter operations in human vision with the same accuracy as the Gabor filters, with the advantage that they accomplish this task with fewer parameters [52, 85, 87]. Fig. 2.9 shows the Hermite filters $D_n(x)$ for $N = 4$ $(n = 0, 1, 2, 3, 4)$.

Fig. 2.10 shows the *House* test image and in Fig. 2.11 we obtain the cartesian Hermite coefficients for $N = 3$ $(n = 0, 1, 2, 3)$.

In order to recover the original image an interpolation of the transform coefficients with the synthesis filters $P_{m,n-m}(x, y)$ is performed followed by an oversampling and adding all the elements:

$$\hat{L}(x,y) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} \sum_{(x_0,y_0)\in S} L_{m,n-m}(x_0, y_0) P_{m,n-m}(x - x_0, y - y_0) \tag{2.13}$$

where

$$P_{m,n-m}(x,y) = \frac{G_{m,n-m}(x,y)v^2(x,y)}{V(x,y)} \tag{2.14}$$

for $n = 0, \ldots, \infty$ and $m = 0, \ldots, n$. In Fig. 2.12 we show the synthesis process of the

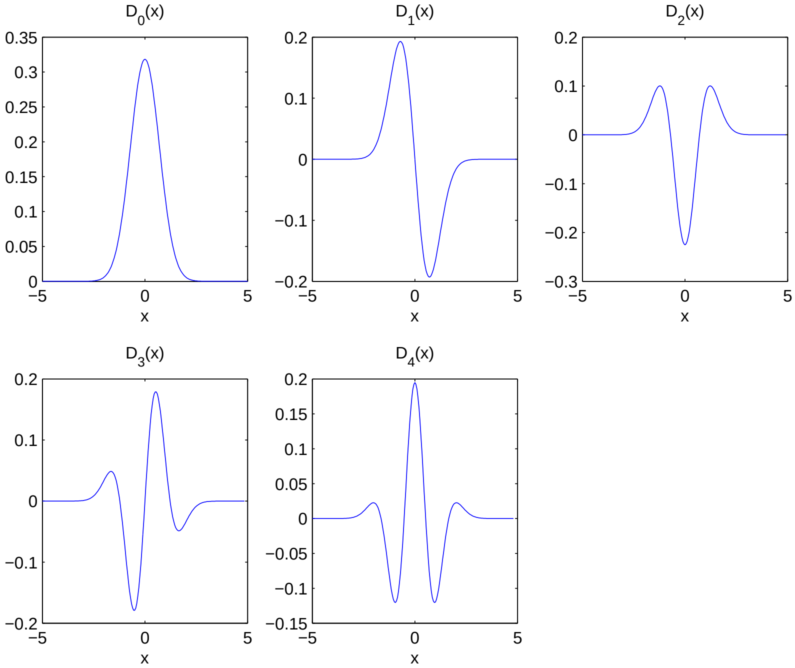
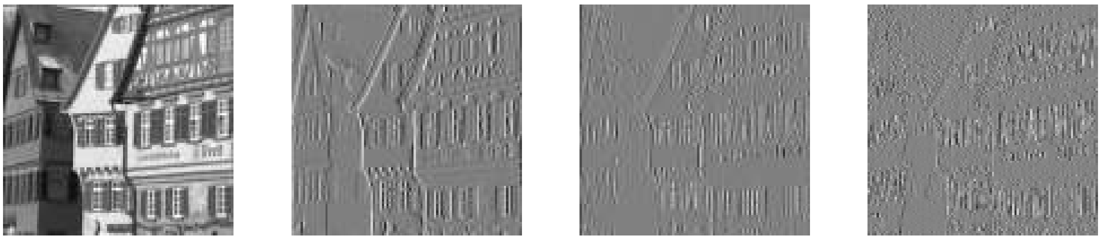


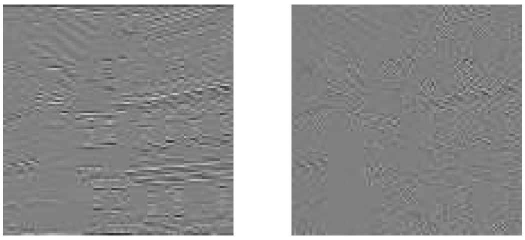
Figure 2.9: The Hermite filters $D_n(x)$ for $N = 4$ $(n = 0, 1, 2, 3, 4)$.

Figure 2.10: *House* test image [37].



(a) $L_{0,0}(x,y)$ (b) $L_{1,0}(x,y)$ (c) $L_{2,0}(x,y)$ (d) $L_{3,0}(x,y)$

(e) $L_{0,1}(x,y)$ (f) $L_{1,1}(x,y)$ (g) $L_{2,1}(x,y)$



(h) $L_{0,2}(x,y)$ (i) $L_{1,2}(x,y)$

(j) $L_{0,3}(x,y)$

Figure 2.11: Cartesian Hermite coefficients of the *House* test image for $N = 3$ ($n = 0, 1, 2, 3$). (a) $L_{0,0}(x,y)$ (represents the DC Hermite coefficient). (b) $L_{1,0}(x,y)$. (c) $L_{2,0}(x,y)$. (d) $L_{3,0}(x,y)$. (e) $L_{0,1}(x,y)$. (f) $L_{1,1}(x,y)$. (g) $L_{2,1}(x,y)$. (h) $L_{0,2}(x,y)$. (i) $L_{1,2}(x,y)$. (j) $L_{0,3}(x,y)$.

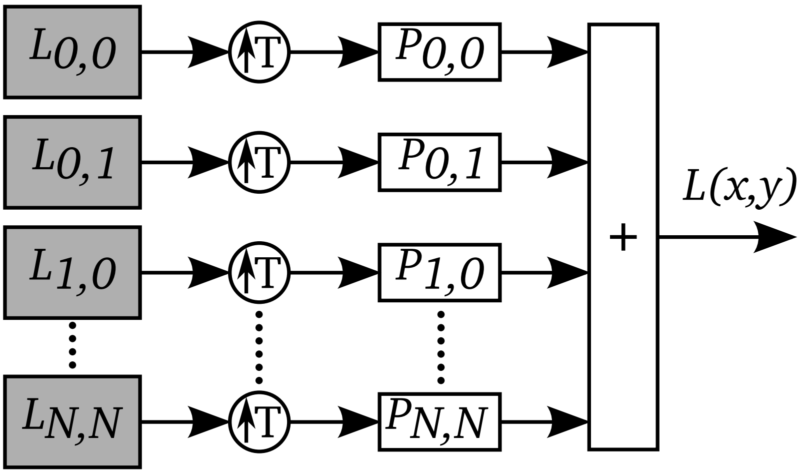


Figure 2.12: Synthesis process of the Hermite transform.

Hermite transform.

The free parameters of the Hermite transform are the maximum derivative order $N$, the subsampling factor and the scale $\sigma$ which must be related to the spatial scale of the image structures to be analyzed. Small windows are better to detect fine details and large windows allow analyzing low resolution objects. The Hermite transform has a multiresolution extension that allows analyzing objects at different scales [24, 73].

## 2.2 Steered Hermite transform

Oriented filters are a class of filters that are rotated copies of each filter, constructed as a linear combination of a set of basis filters [31]. The orientation feature of the Hermite filters explains why they are products of polynomials with a radially symmetric window function (Gaussian function). The $N + 1$ Hermite filters of order $n$ form a steerable basis for each individual filter of order $n$. Because of this characteristic, Hermite filters at each position in the image are adapted to local content [77]. The resulting filters can be interpreted as directional derivatives of a Gaussian function.

For orientation analysis purposes, it is convenient to work with a rotated version of the Hermite transform. The polynomial coefficients can be computed through a convolution of the image with the filter functions $D_m(x)D_{n-m}(y)$. They are separable in space and polar domains, and its Fourier transform can be expressed in polar coordinates. If $\omega_x = \omega cos(\theta)$ and $\omega_y = \omega sin(\theta)$, then

$$d_m(\omega_x)d_{n-m}(\omega_y) = g_{m,n-m}(\theta)d_n(\omega) \tag{2.15}$$

where $d_n(\omega)$ is the Fourier transform of each filter function, which expresses radial frequency selectivity of the $nth$ derivative of the Gaussian but with a radial coordinate $r$ for $x$:

$$d_n(\omega) = \frac{1}{\sqrt{2^n n!}}(-j\omega\sigma)^n \exp^{\left(-\frac{(\omega\sigma)^2}{4}\right)} \tag{2.16}$$

and $g_{m,n-m}(\theta)$ expresses the directional selectivity of the filter.

$$g_{m,n-m}(\theta) = \sqrt{\binom{n}{m}}\left(\cos^m(\theta)\right)\left(\sin^{n-m}(\theta)\right) \tag{2.17}$$

Filters of increasing order $n$ analyze successively higher radial frequencies and filters of the same order $n$ and different (directional) index $m$ distinguish between different orientations in the image [52]. Note that the radial frequency selectivity $d_n(\omega)$ is the same for all $N+1$ filters of the order $n$ and that these filters differ only in their orientation selectivity. In terms of orientation frequency functions, this property of the Hermite filters can be expressed by:

$$g_{m,n-m}(\theta-\theta_0) = \sum_{k=0}^{n}\left(l_{k,n-k,\theta_0}(x,y)\right)\left(g_{k,n-k}(\theta)\right) \tag{2.18}$$

where $l_{k,n-k,\theta_0}(x,y)$ are the steered Hermite coefficients. The Hermite filter rotation at each position over the image is an adaptation to local orientation content.

In order to obtain the steered Hermite coefficients, the Hermite coefficients are rotated toward the estimated local orientation, according to a criterion of maximum oriented energy at each window position. For local 1D patterns, the steered Hermite transform provides a very efficient representation. This representation consists of a parameter $\theta$ that indicates the orientation of the pattern and a small number of coefficients that represent the profile of the pattern perpendicular to its orientation. For a 1D pattern with orientation $\theta$, the following relation holds:

$$l_{m,n-m,\theta}(x,y) = \begin{cases} \sum_{k=0}^{n} \Big(L_{k,n-k}(x,y)\Big)\Big(g_{k,n-k}(\theta)\Big), & m = 0 \\ 0, & m > 0 \end{cases} \tag{2.19}$$

For such pattern, steering over $\theta$ results in a compaction of energy into the coefficients $l_{0,n,\theta}(x,y) = l_{n,\theta}(x,y)$, while all other coefficients are set to zero.

Using Hermite coefficients, the energy content can be expressed according to Parseval Theorem as:

$$E_N = \sum_{n=0}^{N} \sum_{m=0}^{n} \left[L_{m,n-m}(x,y)\right]^2 \tag{2.20}$$

up to order $N$.

The steered Hermite transform offers a way to describe 1D patterns on the basis of their orientation and profile. We can differentiate 1D energy terms and 2D energy terms. That is, for each local signal:

$$E_N^{1D} = \sum_{n=1}^{N} \left[l_{0,n,\theta}(x,y)\right]^2 \tag{2.21}$$

$$E_N^{2D} = \sum_{n=0}^{N} \sum_{m=0}^{n} \left[l_{m,n-m,\theta}(x,y)\right]^2 \tag{2.22}$$

where the energy content of Eq. (2.20) can be expressed as:

$$E_N = \left[L_{0,0}(x,y)\right]^2 + E_N^{1D} + E_N^{2D} \tag{2.23}$$

In Fig. 2.14 we steer the cartesian Hermite coefficients of Fig. 2.11 according to maximum energy direction, the angle $\theta$ (Fig. 2.13) was estimated using the phase of the gradient, which is a good indicator of the direction of the edges, for this we use the coefficients $L_{01}$ and $L_{10}$:

$$\theta(x,y) = \arctan \frac{L_{0,1}(x,y)}{L_{1,0}(x,y)} \tag{2.24}$$

where $L_{01}$ and $L_{10}$ are a good approach to optimal edge detectors in the horizontal and vertical directions respectively.

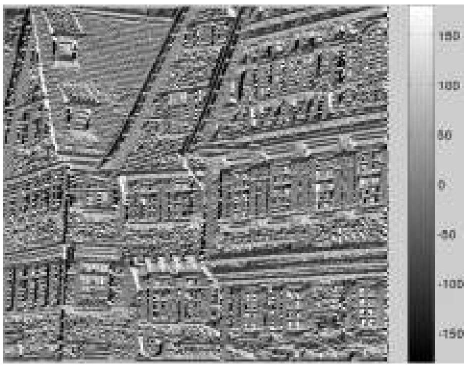
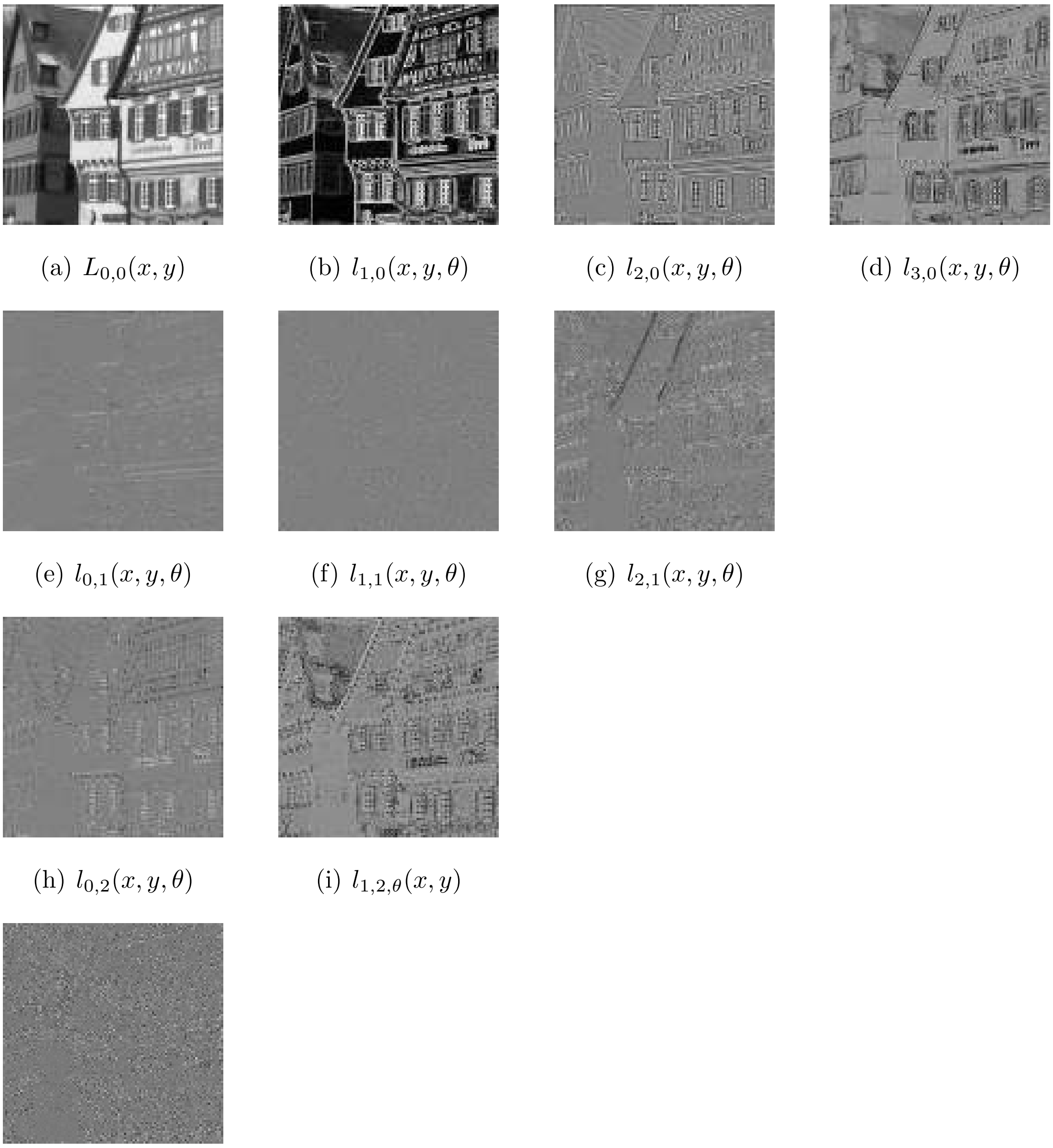


Figure 2.13: Angle $\theta$ of maximum energy ( $^\circ$).

It is noticeable that the energy is concentrated in only three coefficients (Fig. 2.14b-d), which represent the orientation of the different structures of the image.



(a) $L_{0,0}(x,y)$ (b) $l_{1,0}(x,y,\theta)$ (c) $l_{2,0}(x,y,\theta)$ (d) $l_{3,0}(x,y,\theta)$

(e) $l_{0,1}(x,y,\theta)$ (f) $l_{1,1}(x,y,\theta)$ (g) $l_{2,1}(x,y,\theta)$

(h) $l_{0,2}(x,y,\theta)$ (i) $l_{1,2,\theta}(x,y)$

(j) $l_{0,3,\theta}(x,y)$

Figure 2.14: Steered Hermite coefficients of the *House* test image for $N = 3$ ($n = 0,1,2,3$). (a) $l_{0,0}(x,y)$ (represents the DC Hermite coefficient). (b) $l_{1,0,\theta}(x,y)$. (c) $l_{2,0,\theta}(x,y)$. (d) $l_{3,0,\theta}(x,y)$. (e) $l_{0,1,\theta}(x,y)$. (f) $l_{1,1,\theta}(x,y)$. (g) $l_{2,1,\theta}(x,y)$. (h) $l_{0,2,\theta}(x,y)$. (i) $l_{1,2,\theta}(x,y)$. (j) $l_{0,3,\theta}(x,y)$.

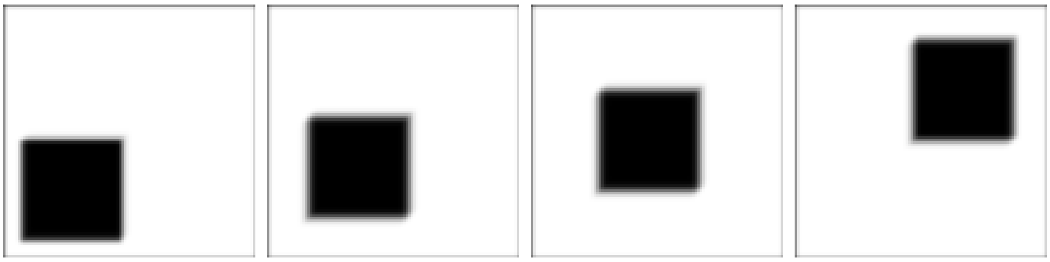
# Chapter 3

# Optical flow estimation

> The one who has conquered himself
> is a far greater hero than he who
> has defeated a thousand times a
> thousand men.
>
> Sidhartha Gautama (563–483 b.c).
> Dhammapada.

Optical flow estimation is still one of the key problems in computer vision. Estimation of the displacement field between two images, applies in those situations requiring the correspondence between the pixels of an image to another. Problems of this type are not only restricted to motion estimation, they are also present in a similar fashion in 3D reconstruction and image registration. Optical flow provides not only a basis for motion–based image sequence interpretation but also for a second generation of video algorithms and applications including, for instance, experimental dynamics, meteorology, medical imaging, video compression, segmentation, object and activity detection, key frame extraction and interpolation in time to mention a few.

When we watch a movie, we see a sequence of images in which objects appear in a number of positions. Although each frame represents a frozen instant of time, the movie gives us a convincing impression of motion. Somehow the visual system



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

Figure 3.1: Apparent motion in a sequence of images.

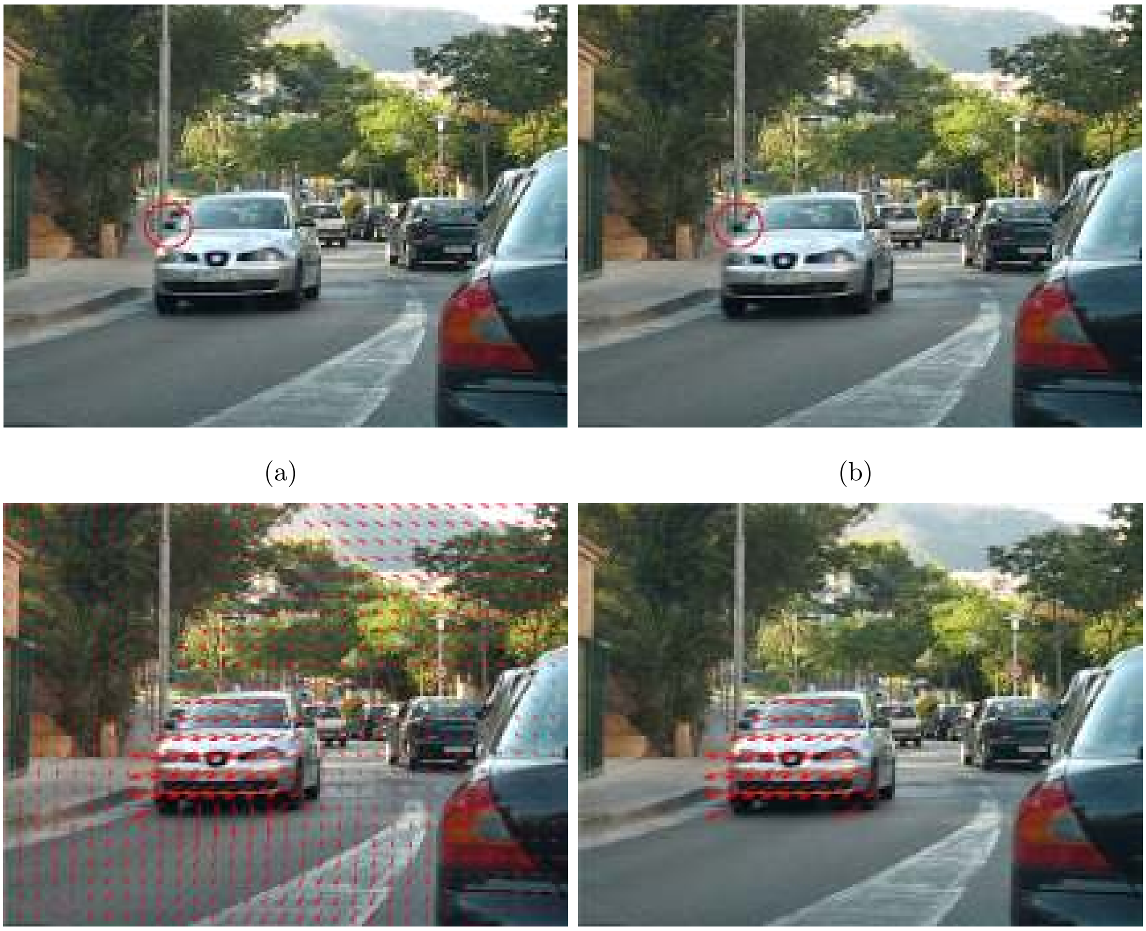
interprets the succession of still images as a continuously moving scene [2] (Fig. 3.1).

The optical flow can be defined as 2D distribution of apparent velocities that can be associated with a variation of brightness patterns in a sequence of images [33, 34]. It can be represented by a vector field induced by the motion of objects (or camera), which encodes the displacement of each pixel in the image. For example in Fig. 3.2 we show the optical flow of *Cameramotion* sequence [47], Fig. 3.2(c) shows the vector field between images Fig. 3.2(a) and Fig. 3.2(b). In Fig. 3.2(d) we show that the optical flow can be used to segment the car in the sequence by grouping pixels with similar displacements.

To detect the optical flow in two images taken at different times we observe the change from one image to another, due to:

- The motion of an object in the scene.
- The relative movement between the observer and the scene.
- Variations in scene illumination.

One of the first applications of optical flow estimation was made by [60] where the discontinuities in the optical flow helped segmenting images into regions that correspond to different objects. Others attempts to segmentation used differences between successive image frames [41, 40, 42, 45, 56].



(a) (b)

(c) (d)

Figure 3.2: Optical flow representation in a sequence of images. (a) Frame 43 of the *Cameramotion* sequence. (b) Frame 44 of the *Cameramotion* sequence. (c) Optical flow. (d) Segmentation of the car using displacement vectors.

## 3.1 Methods of optical flow estimation

There are different ways to group the methods of optical flow estimation. In 1994 Barron et al. classified the existing techniques into four groups [7]:

- Differential methods
- Region–based matching methods
- Energy–based methods
- Phased–based methods

In [9] they grouped energy–based and phased–based methods into frequency–based methods and added two new methods called multiple motion and temporal refinement methods. The boundaries between groups are not always clear because there are methods that incorporate more than one technique or their characteristics can not be placed in a single class.

### 3.1.1 Differential methods

Differential techniques compute image velocity from spatiotemporal derivatives of image intensities or filtered versions of the image [7].

The first approaches use first order derivatives and are based on image translation [27, 38, 57]. Assuming intensity conservation the *Optical Flow Constraint* yields the orientation and normal speed of spatial contours of constant intensity but the vertical and horizontal components are constrained by one linear equation.

Second order differential methods [58, 76], use second order derivatives (the Hessian of the intensity function) to constrain 2–d velocity.

Global and local first- and second-order methods based on *Optical Flow Constraint* can be used to compute optical flow. Global methods use *Optical Flow Constraint* and an additional global constraint, usually a smoothness regularization term, to compute dense optical flows over large image regions. Local methods use normal velocity information in local neighborhoods to perform a least squares minimization to find the best fit for the vertical and horizontal components of displacement [9].

Differential approaches calculate optic flow as the minimizer of an energy functional, which consists of a data term and a smoothness term.

The data term in the energy functional involves optic flow constraints such as the assumption that corresponding pixels in different frames should reveal the same gray value. The smoothness term usually requires that the optic flow field should vary smoothly in space [38]. Such a term may be modified in an image–driven way in order to suppress smoothing at or across image boundaries [57, 3]. As an alternative, flow–driven modifications have been proposed which reduce smoothing across flow discontinuities [11, 69, 80].

From both a theoretical and practical point of view, it can be attractive to use energy functionals that are convex. They have a unique minimum, and this global minimum can be found in a stable way by using standard techniques from convex optimization, for instance gradient descent methods. Having a unique minimum allows to use globally convergent algorithms, where every arbitrary flow initialization leads to the same solution: the global minimum of the functional.

Minimizing continuous energy functionals leads in a natural way to partial differential equations (PDEs): applying gradient descent, for instance, yields a system of coupled diffusion–reaction equations for the two flow components. The fastly emerging use of PDE–based image restoration methods has motivated many researchers to apply similar ideas to estimate optic flow [3, 69, 80].

### 3.1.2 Region–based matching methods

Authors such as [39] proposed region–based matching techniques arguing that the exact numerical differentiation of the differential methods may be impractical because of the noise, this by having a small number of frames or by the smoothed images in the acquisition process.

These methods assume a conservation of the local intensity distribution, they work with two successive images finding the best match between the pixel being analyzed (1st image) and pixels candidates (2nd image). They either to maximize a similarity measure, such as the normalized cross–correlation or minimize a distance measure, such as the sum–of–squared differences (SSD).

For large displacements, the region–based matching methods become sensitive to false matches, due to increase of search spaces. In order to reduce the amount of computations and the potential for mismatches, one may use coarse motions estimates to direct the matching process [9].

In 1992, Ogata and Sato [61] propose the correlation computation with coarse estimates of motion obtained from velocity–tuned Gabor filters. These estimates can then be used to restrict the sizes of search areas and thereby reduce the number of computations usually necessary to obtain disparities.

Anandan's method [5] is based on a Laplacian pyramid and a coarse–to–fine SSD–based matching strategy. The Laplacian pyramid allows for the estimation of large inter frame disparities and helps to enhance image structures, such as edges.

### 3.1.3 Frequency–based methods

Frequency based methods rely on the use of velocity–tuned filters. These techniques use orientation sensitive filters in the Fourier domain of time-varying images. Among advantages brought by these methods, it is found that motion-sensitive mechanisms operating on spatiotemporally oriented energy in Fourier space can estimate motion in image signals for which matching approaches would fail [9].

Adelson and Bergen in [2] propose a class of computational schemes that exploits the fact that detecting image motion is equivalent to extracting spatiotemporal orientation. Gabor filtering[1] is presented as a technique for extracting spatiotemporal energy. The filters oriented in space-time and tuned in spatial frequency are constructed in quadrature (two filters that are 90 deg out of phase) avoiding phase sensitivity that exists at the output of each filter separately.

Watson and Ahumada in [79] defined an orientation-selective mechanism that agrees with psychophysical measurements of human motion sensing. Their mechanism uses a combination of 2D spatial Gabor functions and 1D temporal filters tuned to several orientations for the estimation of local image velocity, noting that the elements

---

[1] A Gabor filter is a Gaussian function multiplied by a sine or cosine wave.

of motion sense appear tuned by two spatial frequencies and frequency spectrum of the motion of an image.

### 3.1.4 Phased–based methods

These methods are referred to as phased–based because velocity is defined in terms of the phase behaviour of band–pass filter outputs [7]. The use of phase was first proposed by [29, 30]. Their approach defines velocity components in terms of the instantaneous motion of level phase contours in the output of band-pass velocity–tuned filters. Band–pass filters are used to decompose the input signal according to scale, speed and orientation.

The method developed by Heeger [36] presents a computational model for the estimation of image velocity which uses quadrature pairs of spatiotemporal Gabor filters. They profit from energy contained in the visual stimulus, considering a textured image. A family of Gabor-energy filters, tuned to the same spatial frequency band but to different spatial orientations, is defined. Using 12 Gabor filters at each spatial scale, the computations of image velocity is formulated as a least squares fit of the filter energies to a plane in frequency space. First, a Gaussian Pyramid is built and then, each pyramid level is band–pass filtered. Gabor filters are then applied to the scale–specific channels, from which velocities are measured [7].

### 3.1.5 Evaluation of the optical flow estimation methods

There are many approaches to determine the optical flow from a sequence of images. Differential methods and phase-based methods are the techniques with better performance according to [7, 32] where a detailed review of computational methodologies for motion analysis is presented. Their study emphasizes the measurement accuracy and concludes that the most accurate methods are the differential proposals where methods using global smooth constraint appear to produce visually attractive flow fields.

## 3.2 Related work on differential methods

Many of the methods of differential optical flow estimation are based on the work of [38] and [49], which incorporated certain restrictions in order to handle the *Ill–posed Opening Problem*[2].

Until a few years ago, it was considered that differential methods for optical flow estimation were either unreliable or slow. In recent years, novel differential approaches have emerged that have suggested some additional restrictions to overcome these problems [15, 14, 17, 63]:

- Local image restrictions.
- Spatial coherence.
- Combining local and global methods.
- Statistical robust optimization functions.
- Multiresolution strategy.

### 3.2.1 Notation

Let $L(x, y, t)$ be an image sequence, where $(x, y)$ represents the location within a rectangular image domain $\Omega$, and $t \in [0, \tau]$ denotes time. Let $u$ and $v$ be the displacement of a pixel at position $(x, y)$ between an image at time $t$ and another image at time $(t + 1)$ in the directions $x$ and $y$ respectively.

We define the following relations:

$$W := (u, v, 1)^\top; \qquad |\nabla W|^2 := |\nabla u|^2 + |\nabla v|^2; \qquad |\nabla_3 W|^2 := |\nabla_3 u|^2 + |\nabla_3 v|^2;$$

$$\nabla u := (u_x, u_y)^\top; \qquad \nabla_3 u := (u_x, u_y, u_t)^\top; \qquad \Delta u := (u_{xx}, u_{yy})^\top;$$

$$X := (x, y, t)^\top; \qquad L_* := \frac{\partial L}{\partial *}; \qquad \nabla_3 L := (L_x, L_y, L_t)^\top$$

---

[2] *Ill–posed problem* in the sense of Hadamard, data only calculated the normal component of optical flow. A problem is called *Well–posed* if it has a unique solution that depends in a continuous (and therefore predictable) way on the input data.

### 3.2.2 Local image restrictions

Horn and Schunck [38] assumes that the intensities of the pixels of the objects remain constant: *Constant Intensity Constraint*:

$$L(x+u, y+v, t+1) - L(x, y, t) = 0 \tag{3.1}$$

Considering linear displacements, we can expand Eq. (3.1) by Taylor's series obtaining the *Optical Flow Constraint* equation:

$$W^{\top}\left(\nabla_3 L\right) = 0 \tag{3.2}$$

In order to minimize Eq. (3.2), [49] considers the flow constant within a neighborhood $\rho$ (Gaussian function $K_\rho$ of standard deviation $\rho$), and determine the two constants $u$ and $v$ at a point $(x, y, t)$ using a weighted least squares approach:

$$E_{LK}(W) = W^{\top} J_\rho(\nabla_3 L) W \tag{3.3}$$

where

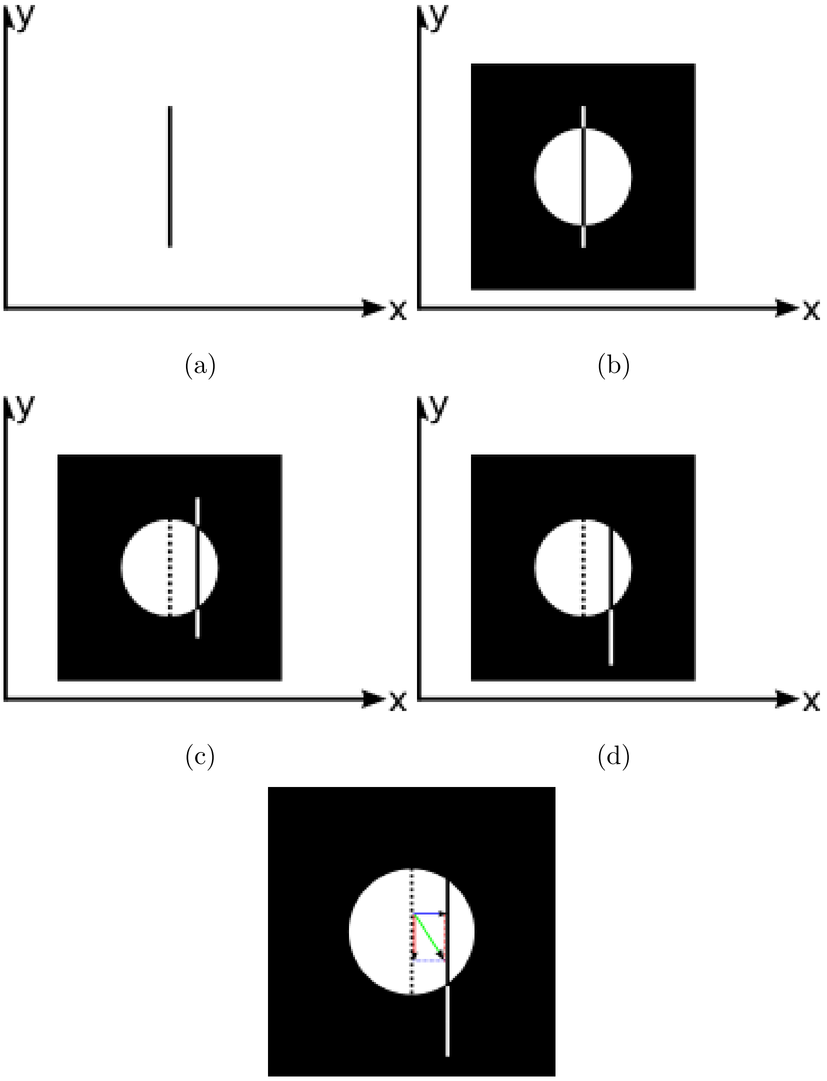
$$J_\rho\left(\nabla_3 L\right) := K_\rho \circledast \left(\nabla_3 L \nabla_3 L^{\top}\right) \tag{3.4}$$

and $\circledast$ represents the convolution operator.

The intensity does not always remain constant from one image to another, therefore an independent intensity change measure is required. In [57, 59] a *Constant Gradient Constraint* (generally used in local methods to handle the aperture problem) is proposed:

$$\nabla L(x+u, y+v, t+1) - \nabla L(x, y, t) = 0 \tag{3.5}$$

The disadvantages of considering only local restrictions are observed in homogeneous areas where movement is observed locally and only the normal component of the movement can be estimated (aperture problem). In Fig. 3.3 we show an example of the aperture problem, in Fig. 3.3(a) we show a straight line in a $xy$ plane, in Fig. 3.3(b) the viewing field is locally restricted using a circular window. If the line moves to the right we can easily see in Fig. 3.3(c) the horizontal displacement through the window, but if the movement is downward and to the right (3.3(d)) only the horizontal displacement is observed (blue line in Fig. 3.3(e)) ignoring the vertical displacement (red line in Fig. 3.3(e)).



(a) (b)

(c) (d)

(e)

Figure 3.3: Aperture problem. (a) A straight line in a $xy$ plane. (b) Viewing field restricted locally by a circular window. (c) Horizontal displacement detected through the window. (d) Only the horizontal displacement is detected through the window ignoring the vertical displacement. (e) Actual displacement in (d).

### 3.2.3 Spatial coherence

Eq. (3.2) is not sufficient to determine the two unknown functions $u$ ans $v$ uniquely. In order to recover a unique flow field, an additional restriction is therefore required. Regularization based optic flow methods additionally assume that the optical flow field should be smooth (or at least piecewise smooth). The basic idea is to recover the optic flow as a minimizer of some energy functional of type [82]

$$E(W) = \int_{\Omega} \Big( W^{\top} \left( \nabla_3 L \nabla_3 L^{\top} \right) W + \alpha U(\nabla L, \nabla u, \nabla v) \Big) dX \tag{3.6}$$

The first term in Eq. (3.6) is a data term requiring that the *Optical Flow Constraint* equation be fulfilled, while the second term penalizes deviations from (piecewise) smoothness. The smoothness term $U(\nabla L, \nabla u, \nabla v)$ is called *regularizer*, and the positive smoothness weight $\alpha$ is the regularization parameter. One would expect that the specific choice of the regularizer has a strong influence on the result.

It is a classic result from the calculus of variations [21] that – under mild regularity conditions– a minimizer $W$ of some energy functional[3] $\mathcal{F}$ [82]

$$E(W) := \int_{\Omega} \mathcal{F}\Big( x, y, u, v, \nabla u, \nabla v \Big) dx\, dy \tag{3.7}$$

satisfies necessarily the so–called Euler–Lagrange equations:

$$\begin{aligned} \partial_x \left( \frac{\partial \mathcal{F}}{\partial u_x} \right) + \partial_y \left( \frac{\partial \mathcal{F}}{\partial u_y} \right) - \frac{\partial \mathcal{F}}{\partial u} &= 0, \\ \partial_x \left( \frac{\partial \mathcal{F}}{\partial v_x} \right) + \partial_y \left( \frac{\partial \mathcal{F}}{\partial v_y} \right) - \frac{\partial \mathcal{F}}{\partial v} &= 0 \end{aligned} \tag{3.8}$$

with homogeneous Neumann boundary conditions:

$$\begin{aligned} \partial_n u &= 0 \ \text{ on } \partial\Omega, \\ \partial_n v &= 0 \ \text{ on } \partial\Omega \end{aligned}$$

---

[3] The object of the calculus of variations is to find extrema of functionals rather than extrema of functions of a finite number of independent variables. By a "functional" we mean a quantity or function which depends on the entire course of one or more functions rather than on a number of discrete variables. The domain of a functional is a set or "space" of admissible functions rather than a region of a coordinate space [21].

where $n$ is a vector normal to the image boundary $\partial\Omega$.

Eqs. (3.8) can be regarded as the steady state ($\mathfrak{t} \to \infty$) of the diffusion–reaction system:

$$\partial_{\mathfrak{t}} u = \partial_x \left( \frac{\partial \mathcal{F}}{\partial u_x} \right) + \partial_y \left( \frac{\partial \mathcal{F}}{\partial u_y} \right) - \frac{\partial \mathcal{F}}{\partial u},$$
$$\partial_{\mathfrak{t}} v = \partial_x \left( \frac{\partial \mathcal{F}}{\partial v_x} \right) + \partial_y \left( \frac{\partial \mathcal{F}}{\partial v_y} \right) - \frac{\partial \mathcal{F}}{\partial v} \qquad (3.9)$$

where $\mathfrak{t}$ denotes an artificial evolution parameter which differs from the time $t$ of the image sequence.

The diffusion–reaction system from Eq. (3.9) can be seen as an analogy to the equation of the anisotropic diffusion equation of Perona and Malik [64].

### Homogeneous regularization

In 1981 Horn and Schunck [38] proposed the regularizer:

$$U_{HS}(\nabla L, \nabla u, \nabla v) := |\nabla u|^2 + |\nabla v|^2 \qquad (3.10)$$

where it is assumed that the apparent speed of the intensity pattern in the image varies smoothly, that is, neighboring points of the objects have similar velocities: *Smoothness Constraint*, leading to the following energy functional:

$$E_{HS}(W) = \int_{\Omega} \Big( W^{\top} \left( \nabla_3 L \nabla_3 L^{\top} \right) W + \alpha |\nabla W|^2 \Big) \, dX \qquad (3.11)$$

Minimizing Eq. (3.11) leads to the PDEs:

$$\Delta u - \frac{1}{\alpha} L_x \Big( W^{\top} \left( \nabla_3 L \right) \Big) = 0,$$
$$\Delta v - \frac{1}{\alpha} L_y \Big( W^{\top} \left( \nabla_3 L \right) \Big) = 0 \qquad (3.12)$$

where $\Delta$ denotes the Laplace operator.

Eqs. (3.12) lead to the corresponding diffusion–reaction equations given by:

$$\partial_{\mathfrak{t}} u = \Delta u - \frac{1}{\alpha} L_x \Big( W^{\top} \left( \nabla_3 L \right) \Big),$$
$$\partial_{\mathfrak{t}} v = \Delta v - \frac{1}{\alpha} L_y \Big( W^{\top} \left( \nabla_3 L \right) \Big) \qquad (3.13)$$

The underlying diffusion process in the Horn and Schunck approach is the linear diffusion equation:

$$\begin{aligned} \partial_t u &= \Delta u = div\Big(g \nabla u\Big), \\ \partial_t v &= \Delta v = div\Big(g \nabla v\Big) \end{aligned} \tag{3.14}$$

with $g := 1$. *div* denoting the divergence operator:

$$div\left(\overrightarrow{F}\right) := \nabla \cdot \overrightarrow{F} = \frac{\partial F_{(x)}}{\partial x} + \frac{\partial F_{(y)}}{\partial y}$$

where $\overrightarrow{F}$ is a vector field[4].

The *Uniform Smoothness Constraint* used in [38] causes an over–smoothing at borders of objects, creating blurry optic flow fields due to the diffusivity $g$ equal to 1 everywhere in Eq. (3.14), smoothing in a completely homogeneous way.

### Isotropic image–driven regularization

A simple way to prevent smoothing at motion boundaries, which in most cases are image boundaries, is introducing a weight function into the Horn and Schunck regularizers that becomes small at image edges [3]. This modification yields the regularizer:

$$U_{II}(\nabla L, \nabla u, \nabla v) := g\left(|\nabla L|^2\right)\left(|\nabla u|^2 + |\nabla v|^2\right) \tag{3.15}$$

where $g$ is a decreasing positive function.

Using the linearized data term of Eq. (3.6) and the regularizer from Eq. (3.15) the energy functional to minimize is:

$$E_{II}(W) = \int_{\Omega} \Big(W^{\top}\left(\nabla_3 L \nabla_3 L^{\top}\right) W + \alpha g\left(|\nabla L|^2\right)\left(|\nabla u|^2 + |\nabla v|^2\right)\Big) dX \tag{3.16}$$

where the corresponding diffusion–reaction equations are given by:

$$\begin{aligned} \partial_t u &= div\Big(g\left(|\nabla L|^2\right)\nabla u\Big) - \frac{1}{\alpha} L_x \Big(W^{\top}\left(\nabla_3 L\right)\Big), \\ \partial_t v &= div\Big(g\left(|\nabla L|^2\right)\nabla v\Big) - \frac{1}{\alpha} L_y \Big(W^{\top}\left(\nabla_3 L\right)\Big) \end{aligned} \tag{3.17}$$

---

[4] Let $C$ be a subset on $\mathbb{R}^2$; a vector field on $\mathbb{R}^2$ is a function $\overrightarrow{F}$ that assigns to each point $(x, y)$ in $C$ a two dimensional vector $\overrightarrow{F}(x, y) = F_{(x)}(x, y)\overrightarrow{i} + F_{(y)}(x, y)\overrightarrow{j}$.

The underlying diffusion process is:

$$\begin{aligned} \partial_t u &= div\Big(g\left(|\nabla L|^2\right)\nabla u\Big), \\ \partial_t v &= div\Big(g\left(|\nabla L|^2\right)\nabla v\Big) \end{aligned} \tag{3.18}$$

The scalar–valued diffusivity $g$ depends on the image gradient, this method can therefore be classified as inhomogeneous, isotropic and image–driven. Isotropic refers to the fact that a scalar–valued diffusivity guarantees a direction–independent smoothing behaviour, while inhomogeneous means that this behaviour may be space–dependent. The diffusion process is linear, since the diffusivity does not depend on the flow [82]. When $g\left(|\nabla L|^2\right) = 1$ the regularizer of the Eq. (3.15) becomes equal to the homogeneous regularizer of the Eq. (3.10).

### Anisotropic image–driven regularization

An anisotropic modification of the Horn and Schunck functional due to [57] reduces smoothing across image boundaries, while encouraging smoothing along image boundaries. This is achieved by considering the regularizer:

$$U_{AI}(\nabla L, \nabla u, \nabla v) := \nabla u^\top \mathcal{D}\left(\nabla L\right)\nabla u + \nabla v^\top \mathcal{D}\left(\nabla L\right)\nabla v \tag{3.19}$$

where $\mathcal{D}\left(\nabla L\right)$ is a regularized projection matrix perpendicular to $\nabla L$:

$$\mathcal{D}(\nabla L) := \frac{1}{|\nabla L|^2 + 2\lambda^2}\Big(\nabla L^\perp \nabla L^{\perp\top} + \lambda^2 I\Big) \tag{3.20}$$

and $I$ denotes the unit matrix.

Using the linearized data term of Eq. (3.6) and the regularizer from Eq. (3.19) the energy functional to minimize is:

$$\begin{aligned} &E_{AI}(W) = \\ &\qquad \int_\Omega \Bigg( W^\top \left(\nabla_3 L \nabla_3 L^\top\right) W \\ &\qquad + \alpha\Big(\nabla u^\top \mathcal{D}\left(\nabla L\right)\nabla u + \nabla v^\top \mathcal{D}\left(\nabla L\right)\nabla v\Big)\Bigg) dX \end{aligned} \tag{3.21}$$

This methods leads to the diffusion–reaction equations:

$$\partial_t u = div\Big(\mathcal{D}\left(\nabla L\right)\nabla u\Big) - \frac{1}{\alpha} L_x\Big(W^\top\left(\nabla_3 L\right)\Big),$$
$$\partial_t v = div\Big(\mathcal{D}\left(\nabla L\right)\nabla v\Big) - \frac{1}{\alpha} L_y\Big(W^\top\left(\nabla_3 L\right)\Big) \tag{3.22}$$

The underlying diffusion process is:

$$\partial_t u = div\Big(\mathcal{D}\left(\nabla L\right)\nabla u\Big),$$
$$\partial_t v = div\Big(\mathcal{D}\left(\nabla L\right)\nabla v\Big) \tag{3.23}$$

The usage of a diffusion tensor $\mathcal{D}\left(\nabla L\right)$ instead of a scalar–valued diffusivity allows a direction–dependent smoothing behaviour (anisotropic). The diffusion tensor depends on the image $L$ but not on the unknown flow. It is an image–driven process that is linear in its diffusion part [82].

The eigenvectors of $\mathcal{D}$ are $e_1 := \nabla L, e_2 := \nabla L^\top$, and the corresponding eigenvalues are given by:

$$\mu_1(|\nabla L|) = \frac{\mu^2}{|\nabla L|^2 + 2\mu^2},$$
$$\mu_2(|\nabla L|) = \frac{|\nabla L|^2 + \mu^2}{|\nabla L|^2 + 2\mu^2} \tag{3.24}$$

Inside the objects, $|\nabla L| \to 0$ with $\mu_1 \to \frac{1}{2}$ and $\mu_2 \to \frac{1}{2}$. At ideal edges $|\nabla L| \to \infty$ with $\mu_1 \to 0$ and $\mu_2 \to 1$. Thus, this regularizer presents an isotropic behaviour within regions, and at image boundaries the process smoothes anisotropically along the edges.

Homogeneous and isotropic image–driven regularizers are special cases of Eq. (3.19), where $\mathcal{D}(\nabla L) := I$ and $\mathcal{D}(\nabla L) := g\left(|\nabla L|^2\right) I$ are chosen.

### Isotropic flow–driven regularization

Image–driven regularization methods reduce the smoothing on the images edges, but may create over–segmentations of the flow field for strongly textured objects (there are much more image boundaries than motion boundaries). In order to reduce smoothing at motion boundaries, one may consider using a purely flow–driven regularizer. In [68, 80], the authors considered regularizers of type [82]:

$$U_{IF}(\nabla L, \nabla u, \nabla v) := \Psi\Big(|\nabla u|^2 + |\nabla v|^2\Big) \tag{3.25}$$

where $\Psi\Big(s^2\Big)$ is a differentiable and increasing smooth function that is convex in $s$.

Using the linearized data term of Eq. (3.6) and the regularizer from Eq. (3.25) the energy functional to minimize is:

$$\begin{aligned} &E_{IF}(W) = \\ &\quad \int_\Omega \left( W^\top \left(\nabla_3 L \nabla_3 L^\top\right) W + \alpha \Psi\Big(|\nabla u|^2 + |\nabla v|^2\Big)\right) dX \end{aligned} \tag{3.26}$$

A regularizer of type of Eq.(3.25) leads to the diffusion–reaction system:

$$\begin{aligned} \partial_t u &= div\left(\Psi'\Big(|\nabla u|^2 + |\nabla v|^2\Big)\nabla u\right) - \frac{1}{\alpha} L_x\Big(W^\top \left(\nabla_3 L\right)\Big), \\ \partial_t v &= div\left(\Psi'\Big(|\nabla u|^2 + |\nabla v|^2\Big)\nabla v\right) - \frac{1}{\alpha} L_y\Big(W^\top \left(\nabla_3 L\right)\Big) \end{aligned} \tag{3.27}$$

The underlying diffusion process is:

$$\begin{aligned} \partial_t u &= div\left(\Psi'\Big(|\nabla u|^2 + |\nabla v|^2\Big)\nabla u\right), \\ \partial_t v &= div\left(\Psi'\Big(|\nabla u|^2 + |\nabla v|^2\Big)\nabla v\right) \end{aligned} \tag{3.28}$$

where $\Psi'$ denotes the derivative of $\Psi$ with respect to its argument. The scalar–valued diffusivity $\Psi'\Big(|\nabla u|^2 + |\nabla v|^2\Big)$ shows that this model is isotropic and flow–driven. In general, the diffusion process is nonlinear.

For the specific choice $\Psi\Big(s^2\Big) := s^2$ homogeneous regularization with diffusivity $\Psi'\Big(s^2\Big) = 1$ is recovered.

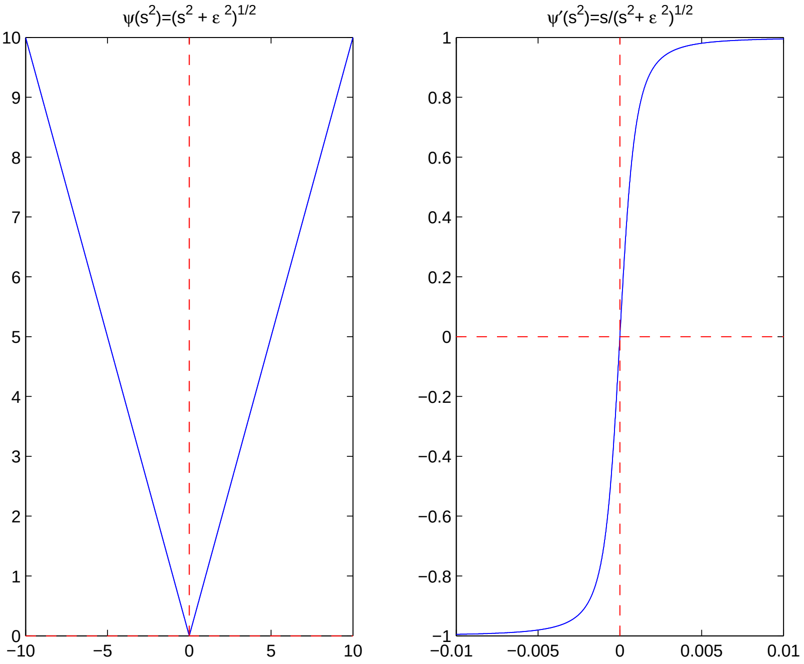


Figure 3.4: Modified $\ell^1$ norm: $\Psi\left(s^2\right) = \sqrt{\left(s^2 + \epsilon^2\right)}$, $\epsilon = 0.001$.

The modified $\ell^1$ norm (Fig. 3.4) presented in [11, 17] is robust in the presence of flow discontinuities:

$$\Psi\left(s^2\right) = \sqrt{\left(s^2 + \epsilon^2\right)} \tag{3.29}$$

where $\epsilon$ ensures the differentiability of $\Psi\left(s^2\right)$ in $s = 0$, so that it is chosen reasonably small, for example $\epsilon \approx 0.001$ [63].

### Anisotropic flow-driven regularization

In previous sections we have seen isotropic and anisotropic image–driven regularizers as well as isotropic flow–driven ones. This section discusses the anisotropic flow–driven regularizer.

In the context of nonlinear diffusion filtering, anisotropic models with a diffusion tensor instead of a scalar–valued diffusivity offer advantages for images with noisy edges or interrupted structures [81].

In the anisotropic flow–driven regularizer one could expect a smoother behaviour along flow discontinuities and less fluctuations than for isotropic flow–driven regularizers. In [82], Weickert and Schnörr propose a novel class of regularizer:

$$U_{AF}(\nabla L, \nabla u, \nabla v) := tr\Psi\Big(\nabla u \nabla u^\top + \nabla v \nabla v^\top\Big) \tag{3.30}$$

where $tr$ is the trace operator[5]..

The scalar–valued function $\Psi(s)$ is extended to a matrix–valued function $\Psi(J)$. Let $J$ denote some symmetric $\mathcal{N} \times \mathcal{N}$ matrix with orthonormal eigenvectors $e_1, \ldots, e_{\mathcal{N}}$ and corresponding eigenvalues $\mu_1, \ldots, \mu_{\mathcal{N}}$:

$$\Psi(J) = \sum_i \Psi(\mu_i)\, e_i e_i^\top \tag{3.31}$$

For Eq.(3.30) we define:

$$J := \nabla u \nabla u^\top + \nabla v \nabla v^\top \tag{3.32}$$

as a symmetric $\mathcal{N} \times \mathcal{N}$ matrix with two orthonormal eigenvectors $e_1, \ldots, e_2$ and their corresponding eigenvalues $\mu_1, \mu_2$. These eigenvalues specify the contrast of the vector–value image $(u, v)$ in the directions $e_1$ and $e_2$ respectively [82].

A regularizer of type of Eq.(3.30) leads to the corresponding descent diffusion–reaction system given by:

$$\begin{aligned}
\partial_t u &= div\left(\Psi'\Big(\nabla u \nabla u^\top + \nabla v \nabla v^\top\Big)\nabla u\right) - \frac{1}{\alpha} L_x\Big(W^\top(\nabla_3 L)\Big),\\
\partial_t v &= div\left(\Psi'\Big(\nabla u \nabla u^\top + \nabla v \nabla v^\top\Big)\nabla v\right) - \frac{1}{\alpha} L_y\Big(W^\top(\nabla_3 L)\Big)
\end{aligned} \tag{3.33}$$

where $\Psi'$ is a matrix–valued function as in Eq. (3.31).

The underlying diffusion process is:

$$\begin{aligned}
\partial_t u &= div\left(\Psi'\Big(\nabla u \nabla u^\top + \nabla v \nabla v^\top\Big)\nabla u\right),\\
\partial_t v &= div\left(\Psi'\Big(\nabla u \nabla u^\top + \nabla v \nabla v^\top\Big)\nabla v\right)
\end{aligned} \tag{3.34}$$

---

[5] For a an $\mathcal{N} \times \mathcal{N}$ square matrix $A$: $tr(A) := \sum_{i=1}^{\mathcal{N}} a_{ii}$

Notably, the isotropic flow–driven regularization from Eq. (3.25) is recovered by just exchanging the trace operator and the penalty function $\Psi$ in Eq. (3.30) [82]:

$$U_{IF}(\nabla L, \nabla u, \nabla v) := \Psi\Big(tr\left(\nabla u \nabla u^{\top} + \nabla v \nabla v^{\top}\right)\Big) \tag{3.35}$$

An advantage of the anisotropic flow–driven regularization in noisy images is that the diffusion tensor differentiates the noise of the important structures of the image.

**Occlusion of objects**

The most popular regularizers (isotropic and/or anisotropic smoothness operators) have some drawbacks, when there are occlusions in the input images these methods cannot correctly handle the flow estimation for the occluded region. An approach taking into account the occlusions was proposed by [84] which uses bilateral filtering [75] in space, intensity and flow.

### 3.2.4 Combining local and global methods

The smoothness term of Eq. (3.6) can get information from neighbors in regions where the intensity gradient is zero, resulting in dense flow fields and making unnecessary interpolation stages commonly found in local differential methods. In this context, [16, 15, 17] proposed a combination of local and global estimation, combining the robustness against noise of the local differential approaches [49, 10] and dense flow fields of global differential methods [38, 57].

By rewriting the Eq. (3.11) to take the form of Eq. (3.3) we have the following expression:

$$E_{HS}(W) = \int_{\Omega}\Big(W^{\top} J_0\left(\nabla_3 L\right) W + \alpha|\nabla W|^2\Big)\, dX \tag{3.36}$$

where $J_0\left(\nabla_3 L\right) := 1 \circledast \Big(\nabla_3 L \nabla_3 L^{\top}\Big)$.

In [16, 15, 17] a Combined Local–Global (CLG) method combining the Eq. (3.3) and Eq. (3.36) is defined:

$$E_{CLG}(W) = \int_{\Omega}\Big(W^{\top} J_{\rho}\left(\nabla_3 L\right) W + \alpha|\nabla W|^2\Big)\, dX \tag{3.37}$$

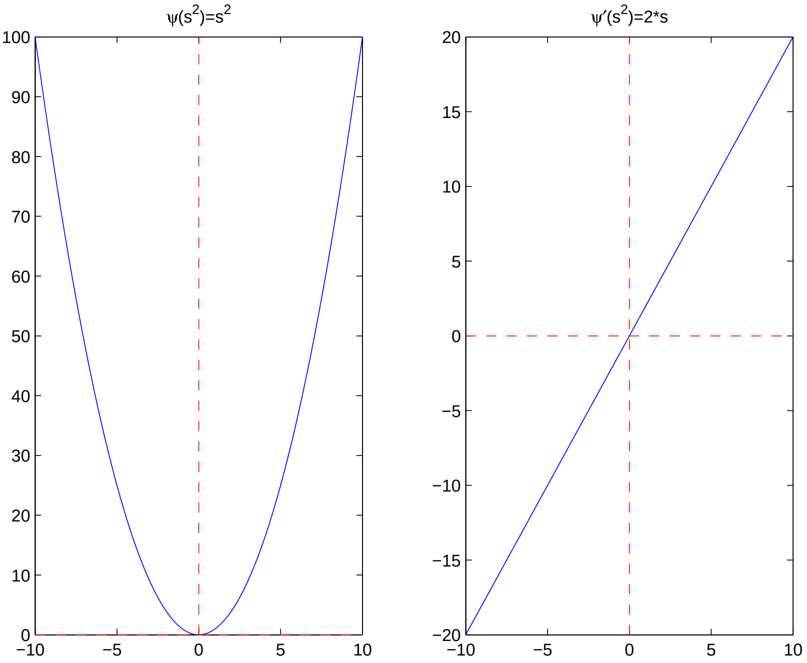


Figure 3.5: $\ell^2$ norm : $\Psi(s^2) = s^2$.

In order to improve robustness to noise, a space–time formulation, performing a convolution with a three–dimensional Gaussian function and considering soft flows in the temporal direction was made by [16]:

$$E_{CLG3}(W) = \int_{\Omega\times[0,\tau]} \left( W^\top J_\rho\,(\nabla_3 L)\, W + \alpha|\nabla_3 W|^2 \right) dX \tag{3.38}$$

### 3.2.5 Statistical robust optimization functions

Eq. (3.3) and Eq. (3.11) use quadratic optimization $\Psi(s^2) = s^2$ ($\ell^2$ norm) to assign considerable weight to outliers. This is shown in its associated influence function[6] ($\Psi' = \frac{\partial\Psi(s^2)}{\partial s}$), which increases linearly and without borders (Fig. 3.5). On the other hand $\ell^1$ norm has a constant value in its influence function but it is used to a lesser extent because of the complexity to minimize the functional.

---

[6] The influence function characterizes deviations of a particular measurement in the solution and is proportional to the derivative of the optimization function.

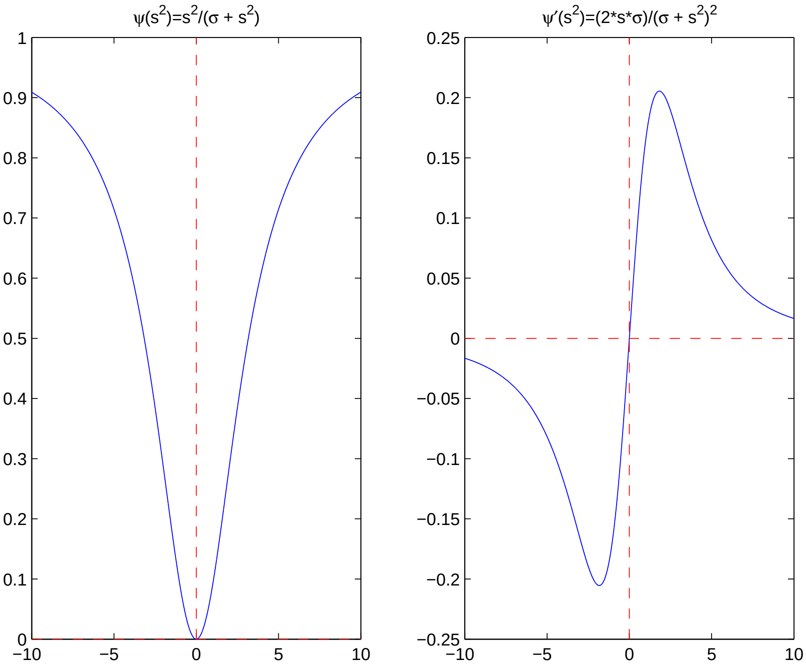


Figure 3.6: Geman–MacClure norm: $\Psi(s^2) = \frac{s^2}{\sigma + s^2}$, with scaling parameter $\sigma = 10$.

Optimization functions more lenient with respect to outliers and with influence functions tending to zero, should be considered, e.g, the Geman–MacClure (Fig. 3.6) and Lorentzian influence functions [11, 17].

A non lineal approach to Eq. (3.38) is defined by [17]:

$$E_{CLG3-N}(W) = \int_{\Omega\times[0,\tau]} \left( \Psi_1\left(W^\top J_\rho(\nabla_3 L) W\right) + \alpha\Psi_2\left(|\nabla_3 W|^2\right)\right) dX \tag{3.39}$$

with Charbonnier optimization function [20]:

$$\Psi_i(s^2) = 2\beta_i^2 \sqrt{\left(1 + \frac{s^2}{\beta_i^2}\right)} \quad i = 1,2 \tag{3.40}$$

where $\beta$ is a scaling parameter.

### 3.2.6 Multiresolution strategy

The equation of *Optical Flow Constraint* is valid when the displacements are relatively small. In order to handle large displacements multiresolution strategies are used [11, 5, 53, 54]. In this context, [17] proposes a functional where the estimated flow in a coarse resolution is used to correct (*warping*) the original sequence before going to the next finer level. In this way, it creates a hierarchy of problems, where only small displacements are calculated at each stage.

The final displacement field is most accurate in comparison with the linearization of the *Optical Flow Constraint.* The non lineal CLG multiscale approach is given by:

$$\begin{aligned} &E^m_{CLG3-N}(\delta W^m) = \\ &\qquad \int_{\Omega\times[0,\tau]} \left( \Psi_1\Big(\delta W^{mT} J_\rho\Big(\nabla_3 L(X+W^m)\Big)\delta W^m\Big) \right. \\ &\qquad \left. + \alpha\Psi_2\Big(|\nabla_3(W^m+\delta W^m)|^2\Big)\right) dX \qquad (3.41) \end{aligned}$$

where $W^{m+1} = W^m + \delta W^m$ and $\delta W^m$ denotes the motion increment at resolution level $m$, with $m = 0$ being the coarsest level with initialization $w^0 = (0, 0, 0)$. The penalty function $\Psi_i$ is given by the Eq. (3.40).

In [63], Papenberg et al. proposes a functional that combines *Constant Intensity Constraint* [59], *Constant Gradient Constraint* [76], *Spatio–temporal Smoothness Constraint* and multiscale approach (*warping*):

$$\begin{aligned} &E_{Warp}(u, v) = \\ &\qquad \int_{\Omega\times[0,\tau]} \Psi\Big(|L(X+W) - L(X)|^2 + \\ &\qquad \gamma|\nabla L(X+W) - \nabla L(X)|^2\Big) dX \\ &\qquad + \alpha \int_{\Omega\times[0,\tau]} \Psi\Big(|\nabla_3 u|^2 + |\nabla_3 v|^2\Big) dX \qquad (3.42) \end{aligned}$$

where $\Psi\left(s^2\right) = \sqrt{(s^2+\epsilon^2)}$ which yields the total variation regularize and $\gamma$ is a weight between the *Constant Intensity Constraint* and the *Constant Gradient Constraint.*

# Chapter 4

# Optical flow estimation using the steered Hermite transform

> You do not really understand
> something unless you can explain it
> to your grandmother.
>
> Albert Einstein (1879–1955).

There have been previous approaches to motion estimation based on the Hermite transform. Liu et al. [48] propose a method that includes a spatio-temporal filtering using the Hermite transform and generalized motion models, such as the affine model, into a single spatial scale. This scheme is proposed the classical *Constant Intensity Constraint* of [38] considering small displacements. More recently, in [72] a directional energy was defined in terms of the 1D Hermite transform coefficients of local projections. Each projection was described by the Hermite transform in terms of a directional derivative analysis of the input at a given scale. The Hermite transform coefficients were then used to detect 1D or 2D spatio-temporal patterns within the 3D stack of images. In [25], a spatio-temporal energy based method to estimate motion in an image sequences was presented. A directional energy was defined in terms of the Radon projections of the Hermite transform. The Radon transform provides a suitable representation for image orientation analysis, while the Hermite transform describes image features locally in terms of Gaussian derivatives. A directional response derived from the directional energy is then used to estimate local motion as well as to compute a confidence matrix. This matrix is used to propagate the velocity information toward directions with high uncertainty.

Unlike the above methods of optical flow estimation using the Hermite transform [48, 72, 25], the present proposal poses a differential approach to estimation. It has been show that differential methods offer better optical flow estimation [7, 32].

In Section 3.1 we mentioned that the boundaries between groups of techniques of optical flow estimation are not always clear because there are methods that incorporate more than one technique. Our differential approach to optical flow estimation using the steered Hermite transform may be regarded as a frequency–based method that incorporates a biological model of human vision, in this sense the Hermite transform is a a model of image representation which performs a decomposition of the images into visual patterns that are relevant to the human vision system.

## 4.1 Model

Given two consecutives images $L(x, y, t)$ and $L(x+u, y+v, t+1)$ in a short period of time, our proposal is based on the polynomial decomposition of each of the images using the steered Hermite transform as a representation of the local characteristics of images from an perceptual approach within a multiresolution scheme. We include elements found in recents advances of differential methods which allow obtaining an accurate optical flow [16, 17, 63].

### 4.1.1 Constancy assumptions

In [63] the effect of different local image restrictions [38] on the data terms (intensity, gradient, Hessian, Laplacian) of the differential methods are described. Our contribution includes a polynomial decomposition of the image and takes as local restrictions the zero order coefficient of the Hermite transform which represents the intensity of the image and the directional higher order derivatives implicitly contained in the steered Hermite coefficients to analyze the image in a similar way as it is done in the HSV. The use of the derivatives of a Gaussian function allows incorporating information from neighboring pixels to determine the local image constraints, similar to local differential methods, as described in Eq. (3.3) which are robust to noise [49, 10]. This feature is included in a global differential functional in order to obtain dense flow fields as in [38, 57].

The local image restrictions of our proposal are:

- *Constant Intensity Constraint*

  The coefficient of order 0 of the Hermite transform $L_{0,0}$ ($L_0$ for simplicity) contains a smoothed version of the original image, so it can be used to define the *Constant Intensity Constraint*:

  $$L_0(x, y, t) = L_0(x + u, y + v, t + 1) \tag{4.1}$$

  Where the smoothing implicit in the DC coefficient allows eliminate any component of high frequency noise.

  Fig. 4.1 shows the coefficients $L_0$ from the image 42 and 43 of *Cameramotion* sequence, where there is a displacement $(u, v)$ of the pixels in the position $(x, y)$ between an image at time $t$ and another image at time $(t + 1)$.

- *Steered Hermite Coefficient Constraint*

  A global change of intensity of the images on two consecutive times violates the *Constant Intensity Constraint* but not the gradient of the images, which is only affected in the magnitude but not in its direction. Moreover, in [63] the *Constant Gradient Constraint* considers higher order derivatives for the formulation of constancy assumptions such as the constancy of the Hessian, the constancy of the Laplacian, the constancy of the norm of the Hessian and the constancy of the determinant of the Hessian. It is mentioned that derivatives of order larger than two can also be considered. Therefore in our proposal we include the steered Hermite coefficients $l_{n,\theta}$ up to order $N$ for dealing with various movements, such as translational and rotational motions. This allows

(a) (b)

Figure 4.1: Cartesian Hermite coefficients $L_{00}$ from the image 42 and 43 of *Cameramotion* sequence. (a) $L_{0,0}(x, y, t)$. (b) $L_{0,0}(x + u, y + v, t + 1)$.

including higher derivatives simply by changing the maximum order $N$ of the decomposition of the Hermite transform.

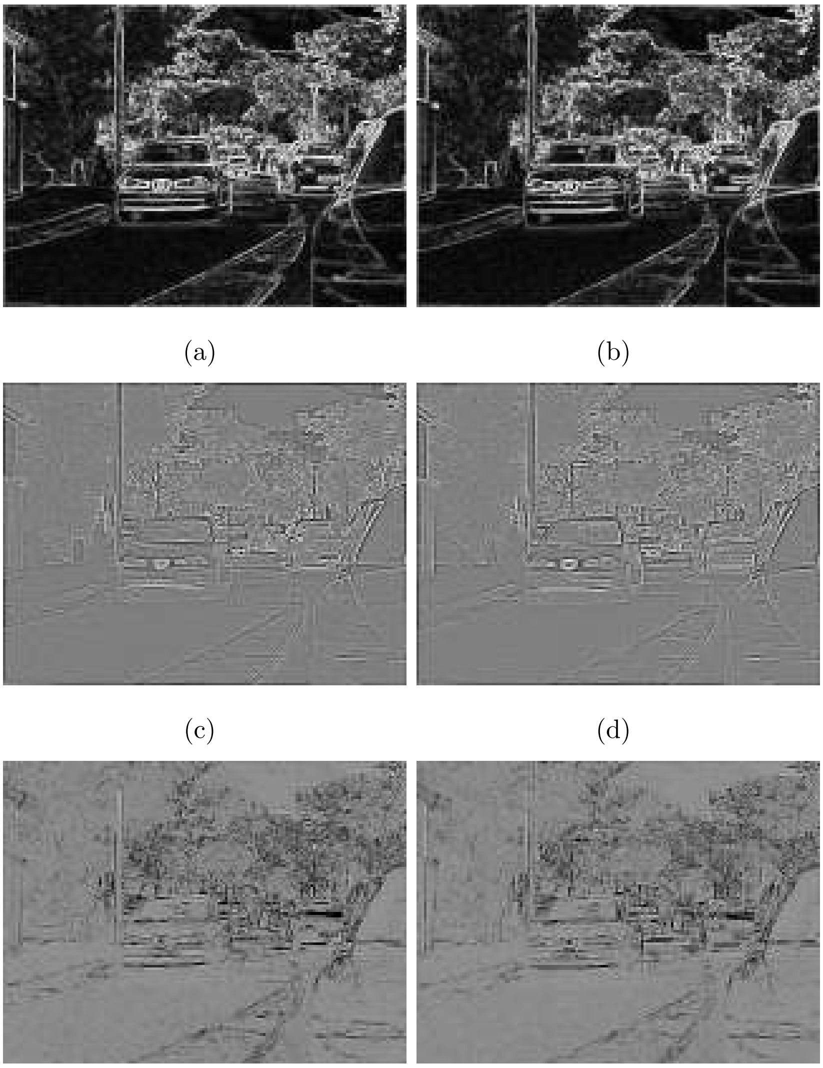
In this sense we define the *Steered Hermite Coefficient Constraint*:

$$l_{n,\theta}(x, y, t) = l_{n,\theta}(x + u, y + v, t + 1) \tag{4.2}$$

where $n = 1, 2, \cdots, N$ and $\theta$ is the angle of maximum energy for the position $(x, y)$, which is given by the Eq. (2.24).

Using the steered Hermite coefficients we obtain rotation invariance as demonstrated in [26]. This allows defining local image restrictions considering constancy in the polynomial decomposition of images and identifying perceptually relevant visual patterns that represent the most important characteristics of the image.

Fig. 4.2 shows the coefficients $l_{n,\theta}$ from the image 42 and 43 of *Cameramotion* sequence.



(a) (b)

(c) (d)

(e) (f)

Figure 4.2: Steered Hermite coefficients up to order $N = 3$ from the image 42 and 43 of *Cameramotion* sequence. (a) $l_{1,\theta}(x, y, t)$. (b) $l_{1,\theta}(x+u, y+v, t+1)$. (c) $l_{2,\theta}(x, y, t)$. (d) $l_{2,\theta}(x+u, y+v, t+1)$. (e) $l_{3,\theta}(x, y, t)$. (f) $l_{3,\theta}(x+u, y+v, t+1)$.

- *Smoothness Constraint*:

  We assume that the flow is piecewise smooth as in [16, 15, 17] allowing to have robustness to discontinuities of flow:

  $$min \int_{\Omega} \Psi\Big(|\nabla u|^2 + |\nabla v|^2\Big)\, dX \tag{4.3}$$

  where $\Psi(s^2)$ is a non-quadratic penalty function which enables capture also locally non-smooth motion, allowing outliers in the smoothness assumption.

### 4.1.2 Energy

We define a functional that penalizes deviations from model assumptions. All deviations from the assumption of *Constant Intensity Constraint* (Eq. (4.1)), *Steered Hermite Coefficient Constraint* (Eq. (4.2)) and *Smoothness Constraint* (Eq. (4.3)) are measured by the energy:

$$E(u,v) = E_{Data}(u,v) + \alpha E_{Smooth}(u,v) \tag{4.4}$$

where the regularization parameter $\alpha > 0$ is a smoothness weight.

The data term $E_{Data}(u,v)$ is given by:

$$\begin{aligned} E_{Data}(u,v) = &\\ &\int_{\Omega} \Psi\Big(|L_0(X+W) - L_0(X)|^2 + \\ &\gamma |l_{n,\theta}(X+W) - l_{n,\theta}(X)|^2\Big)\, dX \end{aligned} \tag{4.5}$$

where $\gamma$ is a weight between the *Constant Intensity Constraint* and the *Steered Hermite Coefficient Constraint.*

As optimization function we use the modified $\ell^1$ norm (Eq. (3.29)), which is robust to outliers:

$$\Psi(s^2) = \sqrt{(s^2 + \epsilon^2)} \tag{4.6}$$

where $\epsilon = 0.001$ ensure the differentiability of $\Psi\left(s^2\right)$ in $s = 0$.

We include the smoothness term of Eq. (4.3):

$$E_{Smooth}(u,v) = \int_{\Omega} \Psi\Big(|\nabla u|^2 + |\nabla v|^2\Big)\, dX \tag{4.7}$$

Therefore, our global energy functional to minimize that includes the data and smoothness term has the form:

$$\begin{aligned} E(u,v) = & \\ & \int_{\Omega} \Psi\Big( |L_0(X+W) - L_0(X)|^2 \\ & + \gamma |l_{n,\theta}(X+W) - l_{n,\theta}(X)|^2 \Big) \, dX + \\ & \alpha \int_{\Omega} \Psi\Big( |\nabla u|^2 + |\nabla v|^2 \Big) \, dX \end{aligned} \tag{4.8}$$

The proposed functional, includes the coefficient of order 0 ($n = 0$) and the steered Hermite coefficients of the Hermite transform as basic elements to define the local image restrictions through a polynomial decomposition of degree $N$ ($n = 1, 2, \cdots, N$).

## 4.2 Numerical solution

The corresponding Euler–Lagrange equations that satisfy the minimization of Eq. (4.8) in function of $W$ are

$$\begin{aligned}
&\Psi'\Big(|L_0(X+W)-L_0(X)|^2\\
&+\gamma|l_{n,\theta}(X+W)-l_{n,\theta}(X)|^2\Big)\cdot\\
&\Bigg[|L_0(X+W)-L_0(X)|L_{10}(X+W)\\
&+\gamma\Big(|l_{n,\theta}(X+W)-l_{n,\theta}(X)|l_{n,\theta_{(m)+1}}(X+W)\Big)\Bigg]-\\
&\alpha\ div\left(\Psi'\Big(|\nabla u|^2+|\nabla v|^2\Big)\nabla u\right)=0
\end{aligned} \tag{4.9}$$

$$\begin{aligned}
&\Psi'\Big(|L_0(X+W)-L_0(X)|^2\\
&+\gamma|l_{n,\theta}(X+W)-l_{n,\theta}(X)|^2\Big)\cdot\\
&\Bigg[|L_0(X+W)-L_0(X)|L_{01}(X+W)\\
&+\gamma\Big(|l_{n,\theta}(X+W)-l_{n,\theta}(X)|l_{n,\theta_{(n)+1}}(X+W)\Big)\Bigg]-\\
&\alpha\ div\left(\Psi'\Big(|\nabla u|^2+|\nabla v|^2\Big)\nabla v\right)=0
\end{aligned} \tag{4.10}$$

For a description of the simplification procedure see Appendix D.

### 4.2.1 Iterative approach

Defining a fixed-point iterative scheme, as in [63], to find a solution for $W$ in Eq. (4.9) and Eq. (4.10), leads to the following equations:

$$
\begin{aligned}
&\Psi'\Big(\left|L_0(X+W^{k+1})-L_0(X)\right|^2 \\
&+\gamma\left|l_{n,\theta}(X+W^{k+1})-l_{n,\theta}(X)\right|^2\Big)\cdot \\
&\Bigg[\left|L_0(X+W^{k+1})-L_0(X)\right|L_{10}(X+W^k) \\
&+\gamma\Big(\left|l_{n,\theta}(X+W^{k+1})-l_{n,\theta}(X)\right|l_{n,\theta_{(m)+1}}(X+W^k)\Big)\Bigg]- \\
&\alpha\ div\left(\Psi'\Big(\left|\nabla u^{k+1}\right|^2+\left|\nabla v^{k+1}\right|^2\Big)\nabla u^{k+1}\right)=0
\end{aligned}
\tag{4.11}
$$

$$
\begin{aligned}
&\Psi'\Big(\left|L_0(X+W^{k+1})-L_0(X)\right|^2 \\
&+\gamma\left|l_{n,\theta}(X+W^{k+1})-l_{n,\theta}(X)\right|^2\Big)\cdot \\
&\Bigg[\left|L_0(X+W^{k+1})-L_0(X)\right|L_{01}(X+W^k) \\
&+\gamma\Big(\left|l_{n,\theta}(X+W^{k+1})-l_{n,\theta}(X)\right|l_{n,\theta_{(n)+1}}(X+W^k)\Big)\Bigg]- \\
&\alpha\ div\left(\Psi'\Big(\left|\nabla u^{k+1}\right|^2+\left|\nabla v^{k+1}\right|^2\Big)\nabla v^{k+1}\right)=0
\end{aligned}
\tag{4.12}
$$

where $W^k=(u^k,v^k,1)^T$ are the unknown variables $u^k$ and $v^k$ in the outer iteration $k$, thus $W^{k+1}$ is the solution for Eq. (4.11) and Eq. (4.12).

Eqs. (4.11) and (4.12) are nonlinear in the terms of the form $f(X+W^{k+1})-f(X)$, therefore we use a 1st order Taylor expansion of such terms (see Appendix E). The resulting equation system reads:

$$
\begin{aligned}
&\Psi'\Big(\Big|L_0(X+W^k)-L_0(X)\\
&+du^kL_{10}(X+W^k)+dv^kL_{01}(X+W^k)\Big|^2\\
&+\gamma\Big|l_{n,\theta}(X+W^k)-l_{n,\theta}(X)\\
&+du^kl_{n,\theta_{(m)+1}}(X+W^k)+dv^kl_{n,\theta_{(n)+1}}(X+W^k)\Big|^2\Big)\cdot\\
&\Bigg[\Big|L_0(X+W^k)-L_0(X)\\
&+du^kL_{10}(X+W^k)+dv^kL_{01}(X+W^k)\Big|L_{10}(X+W^k)\\
&+\gamma\Big(\Big|l_{n,\theta}(X+W^k)-l_{n,\theta}(X)\\
&+du^kl_{n,\theta_{(m)+1}}(X+W^k)+dv^kl_{n,\theta_{(n)+1}}(X+W^k)\Big|\cdot\\
&l_{n,\theta_{(m)+1}}(X+W^k)\Big)\Bigg]-\\
&\alpha\ div\Bigg(\Psi'\Big(\big|\nabla(u^k+du^k)\big|^2+\big|\nabla(v^k+dv^k)\big|^2\Big)\cdot\\
&\nabla(u^k+du^k)\Bigg)=0
\end{aligned}
\tag{4.13}
$$

$$
\begin{aligned}
&\Psi'\Big(\Big|L_0(X+W^k)-L_0(X)\\
&+du^kL_{10}(X+W^k)+dv^kL_{01}(X+W^k)\Big|^2\\
&+\gamma\Big|l_{n,\theta}(X+W^k)-l_{n,\theta}(X)\\
&+du^kl_{n,\theta_{(m)+1}}(X+W^k)+dv^kl_{n,\theta_{(n)+1}}(X+W^k)\Big|^2\Big)\cdot\\
&\Bigg[\Big|L_0(X+W^k)-L_0(X)\\
&+du^kL_{10}(X+W^k)+dv^kL_{01}(X+W^k)\Big|L_{01}(X+W^k)\\
&+\gamma\Big(\Big|l_{n,\theta}(X+W^k)-l_{n,\theta}(X)\\
&+du^kl_{n,\theta_{(m)+1}}(X+W^k)+dv^kl_{n,\theta_{(n)+1}}(X+W^k)\Big|\cdot\\
&l_{n,\theta_{(n)+1}}(X+W^k)\Big)\Bigg]-\\
&\alpha\ div\Bigg(\Psi'\Big(\big|\nabla(u^k+du^k)\big|^2+\big|\nabla(v^k+dv^k)\big|^2\Big)\cdot\\
&\nabla(v^k+dv^k)\Bigg)=0
\end{aligned}
\qquad (4.14)
$$

In order to remove the nonlinearity in $\Psi'$ we apply a second fixed-point iteration, with index $l$ for the inner iteration:

$$
\begin{aligned}
&\Psi'\Big(\Big|L_0(X+W^k)-L_0(X)\\
&+du^{k,l}L_{10}(X+W^k)+dv^{k,l}L_{01}(X+W^k)\Big|^2\\
&+\gamma\Big|l_{n,\theta}(X+W^k)-l_{n,\theta}(X)\\
&+du^{k,l}l_{n,\theta_{(m)+1}}(X+W^k)+dv^{k,l}l_{n,\theta_{(n)+1}}(X+W^k)\Big|^2\Big)\cdot\\
&\Bigg[\Big|L_0(X+W^k)-L_0(X)\\
&+du^{k,l+1}L_{10}(X+W^k)+dv^{k,l+1}L_{01}(X+W^k)\Big|L_{10}(X+W^k)\\
&+\gamma\Big(\Big|l_{n,\theta}(X+W^k)-l_{n,\theta}(X)\\
&+du^{k,l+1}l_{n,\theta_{(m)+1}}(X+W^k)+dv^{k,l+1}l_{n,\theta_{(n)+1}}(X+W^k)\Big|\cdot\\
&l_{n,\theta_{(m)+1}}(X+W^k)\Big)\Bigg]-\\
&\alpha\ div\Bigg(\Psi'\Big(\big|\nabla(u^k+du^{k,l})\big|^2+\big|\nabla(v^k+dv^{k,l})\big|^2\Big)\cdot\\
&\nabla(u^k+du^{k,l+1})\Bigg)=0
\end{aligned}
\tag{4.15}
$$

$$
\begin{aligned}
&\Psi'\Big(\Big|L_0(X+W^k)-L_0(X)\\
&+du^{k,l}L_{10}(X+W^k)+dv^{k,l}L_{01}(X+W^k)\Big|^2\\
&+\gamma\Big|l_{n,\theta}(X+W^k)-l_{n,\theta}(X)\\
&+du^{k,l}l_{n,\theta_{(m)+1}}(X+W^k)+dv^{k,l}l_{n,\theta_{(n)+1}}(X+W^k)\Big|^2\Big)\cdot\\
&\Bigg[\Big|L_0(X+W^k)-L_0(X)\\
&+du^{k,l+1}L_{10}(X+W^k)+dv^{k,l+1}L_{01}(X+W^k)\Big|L_{01}(X+W^k)\\
&+\gamma\Big(\Big|l_{n,\theta}(X+W^k)-l_{n,\theta}(X)\\
&+du^{k,l+1}l_{n,\theta_{(m)+1}}(X+W^k)+dv^{k,l+1}l_{n,\theta_{(n)+1}}(X+W^k)\Big|\cdot\\
&l_{n,\theta_{(n)+1}}(X+W^k)\Big)\Bigg]-\\
&\alpha\ div\Bigg(\Psi'\Big(\big|\nabla(u^k+du^{k,l})\big|^2+\big|\nabla(v^k+dv^{k,l})\big|^2\Big)\cdot\\
&\nabla(v^k+dv^{k,l+1})\Bigg)=0
\end{aligned}
\tag{4.16}
$$

### 4.2.2 Multiscale strategy

In order to consider large displacements in our approach we need to include a multiscale strategy [11, 5, 53, 54] that does not contradict the linearization considered in Eqs. (4.13) and (4.14).

For the multiresolution strategy a Gaussian pyramid of the image is generated using a downsampling factor $\eta \in (0, 1)$, where $\eta$ remains constant for each stage. The coarse grid is obtained by scaling the images $L(x, y, t)$ and $L(x + u, y + v, t + 1)$ by a factor $\eta^i$ for $i = M-1, M-2, \cdots, 0$, where $M$ represents the number of decomposition levels.

Starting at the coarse level ($i = M - 1$) with $W^0 = (0, 0, 1)^T$, $du^{k,0} = 0$ and $dv^{k,0} = 0$ the inner iteration ($l$) allows obtaining the increment $du^k$ for the outer iteration ($k$), where $u^{k+1} = u^k + du^{k,l+1}$ is the solution for the linear system of the Eqs.(4.15) and (4.16) in the current level [63]. Then, the solution is interpolated and propagate to the next finer level ($i = i - 1$) where it is employed for the initialisation of the outer iteration.

# Chapter 5

# Results

> If my relativity theory is verified, Germany will proclaim me a German and France will call me a citizen of the world. But if my theory is proved false, France will emphasize that I am a German and Germany will say that I am a Jew.
>
> ---
>
> Albert Einstein (1879–1955).
> Speaking at the Sorbonne during the 1930s.

## 5.1 Test images

In order to test the frame estimation performance of our method we used sequences with ground truth as well as computed tomography (CT) images (see Appendix A). In the sequences with ground truth we calculated the angular error between the estimated optical flow and the ground truth. In the case of CT images, we ran objective and subjective tests. Objective tests consisted of reconstructing sequential images from their adjacent images and their motion vectors. We then measured the backward reconstruction error.

| **Parameter name** | **Symbol** | **Value** |
|---|---|---|
| Hermite Polynomial degree | $N$ | 3 |
| Downsampling factor | $\eta$ | 0.95 |
| Regularization parameter | $\alpha$ | 50 |
| Local constraint weight | $\gamma$ | 90 |
| Decomposition levels | $M$ | 40 |
| Outer iteration number | $k$ | 15 |
| Inner iteration number | $l$ | 50 |

Table 5.1: Optical flow parameters using steered Hermite transform approach.

For the optical flow estimation using the steered Hermite transform, we used the parameters of Table 5.1.

## 5.2 Algorithm validation

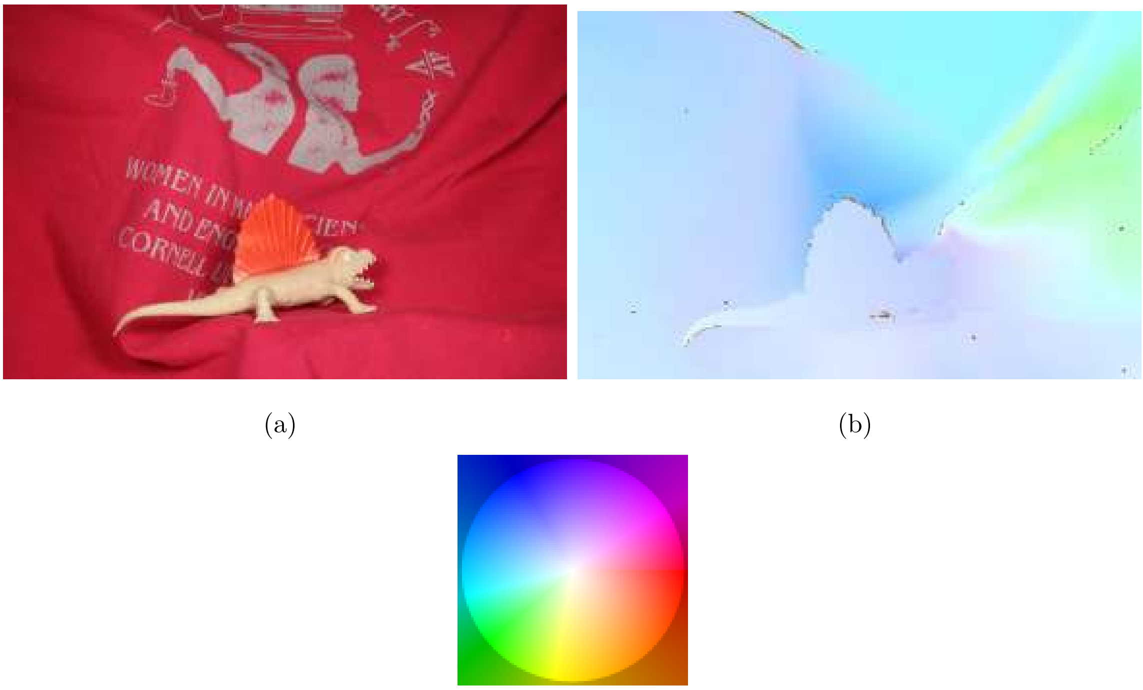
In order to validate our method we compared our implementation with the 2D algorithm of [63] using several sequences of the new set of benchmarks of [6]. They defined sequences with nonrigid motion where the *ground truth* flow was determined by tracking hidden fluorescent texture. This methodology pushing the limits of current technology, revealing where current algorithms fail, and evaluating the next generation of optical flow algorithms.

As a performance measure we calculate the angular error ($AE$) proposed in [6]. The angular error is defined as the angle in 3D space between the dot product of two flow normalized vectors $\overrightarrow{u} = (u_0, u_1)$ and $\overrightarrow{v} = (v_0, v_1)$:

$$AE = \arccos(\overrightarrow{u} \cdot \overrightarrow{v}) \tag{5.1}$$

where

$$\overrightarrow{u} \cdot \overrightarrow{v} = \hat{u}_0\hat{v}_0 + \hat{u}_1\hat{v}_1 \tag{5.2}$$



(a) (b)

(c)

Figure 5.1: Optical flow of the sequence *Dimetrodon*. (a) Frame 10 of the sequence. (b) *Ground truth*. (c) Reference color wheel.

Fig. 5.1(a) shows the frame 10 of the *Dimetrodon* sequence and its *ground truth* is displayed in Fig. 5.1(b). In Fig. 5.1(c) a reference color wheel is used to encode the direction of the flow vector, where each color represents the direction and its tone the magnitude of the vector.

In Fig. 5.2 we show the optical flow for the *Dimetrodon* sequence using the approach of [63] (Fig. 5.2(a)) and the steered Hermite approach (Fig. 5.2(c)). For the implementation of [63] we used the binaries of the author's website published in [13].

Figs. 5.2(b) and 5.2(d) show the average error for both approaches, namely $AE = 3.03^\circ$ for approach [63] and $AE = 2.7^\circ$ for our proposal. Noting that the error is smaller with our approach.

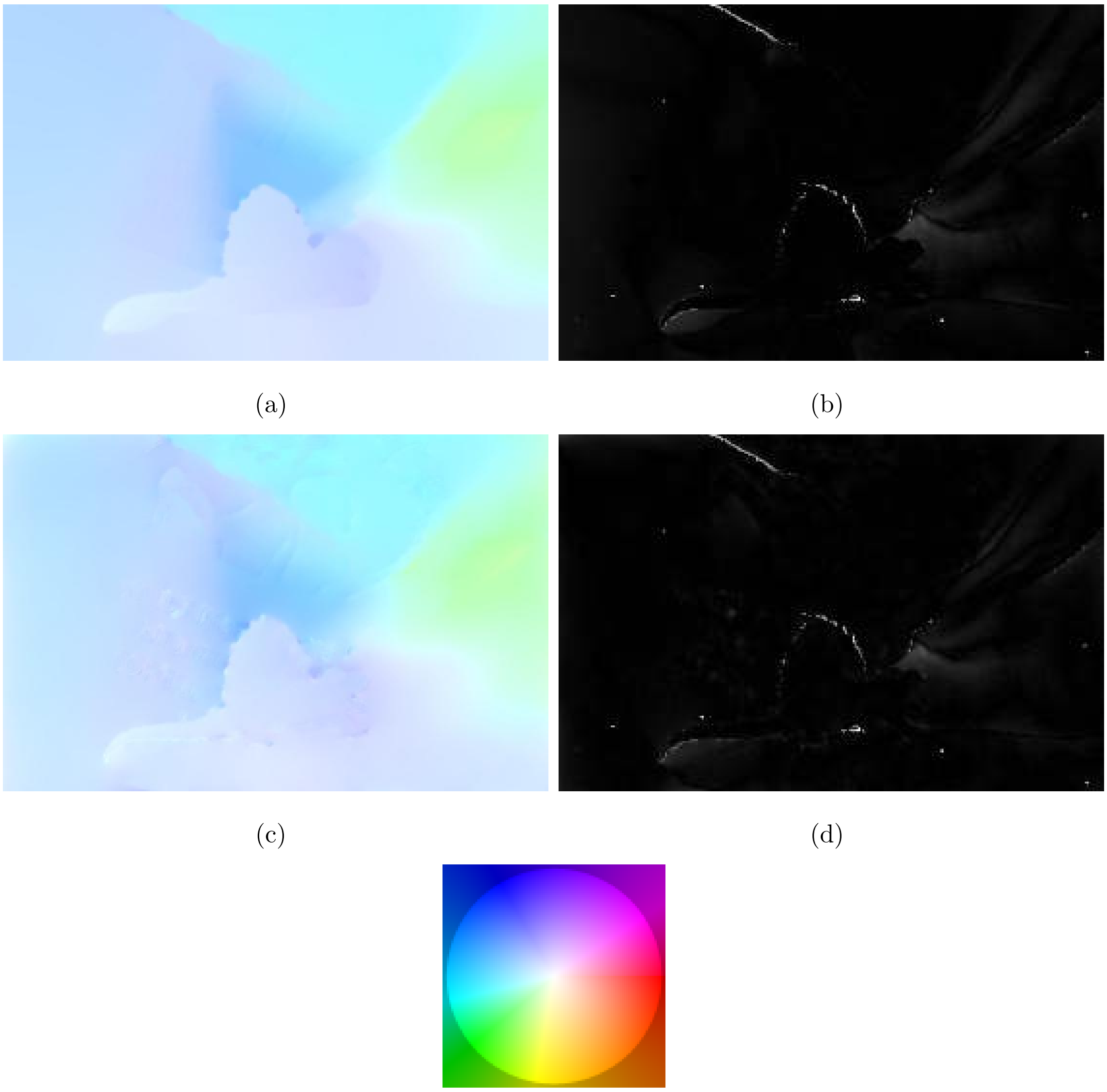
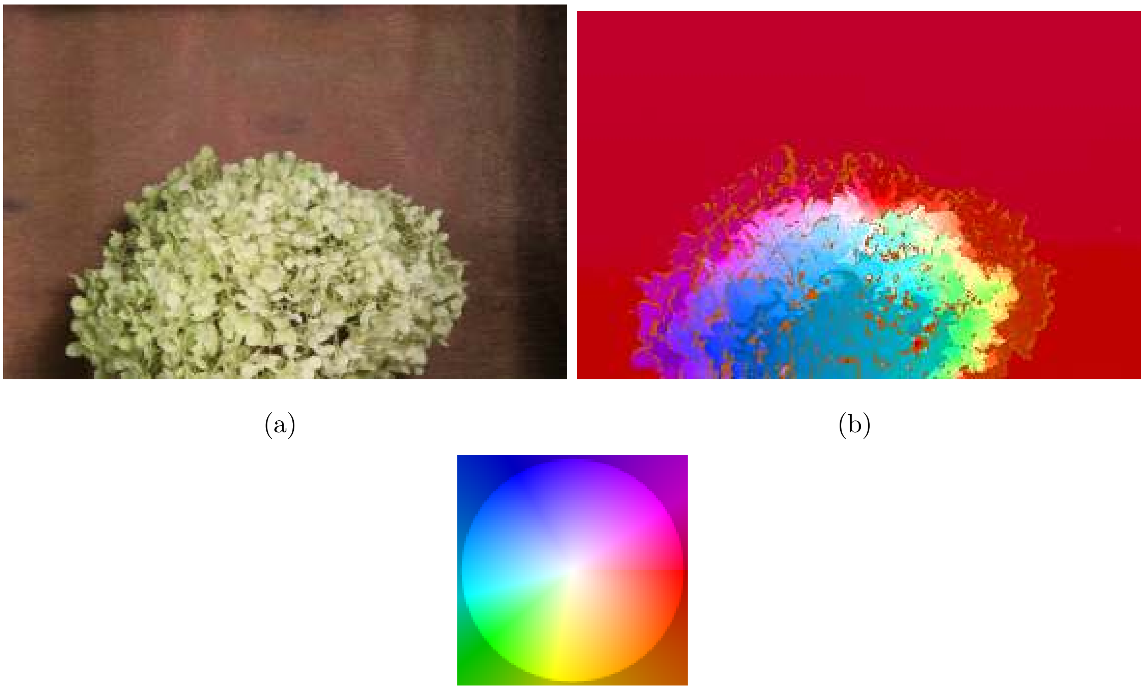

(a) (b) (c) (d) (e)

Figure 5.2: Optical flow fields of the *Dimetrodon* sequence. (a) Optical flow approach of [63]. (b) Absolute error between the *ground truth* and the optical flow approach of [63] $AE = 3.03\,^{\circ}$. (c) Optical flow using the steered Hermite coefficients. (d) Absolute error between the *ground truth* and the optical flow using steered Hermite coefficients $AE = 2.7\,^{\circ}$. (e) Reference color wheel.



(a) (b)

(c)

Figure 5.3: Optical flow of the sequence *Hydrangea*. (a) Frame 10 of the sequence. (b) *Ground truth*. (c) Reference color wheel.

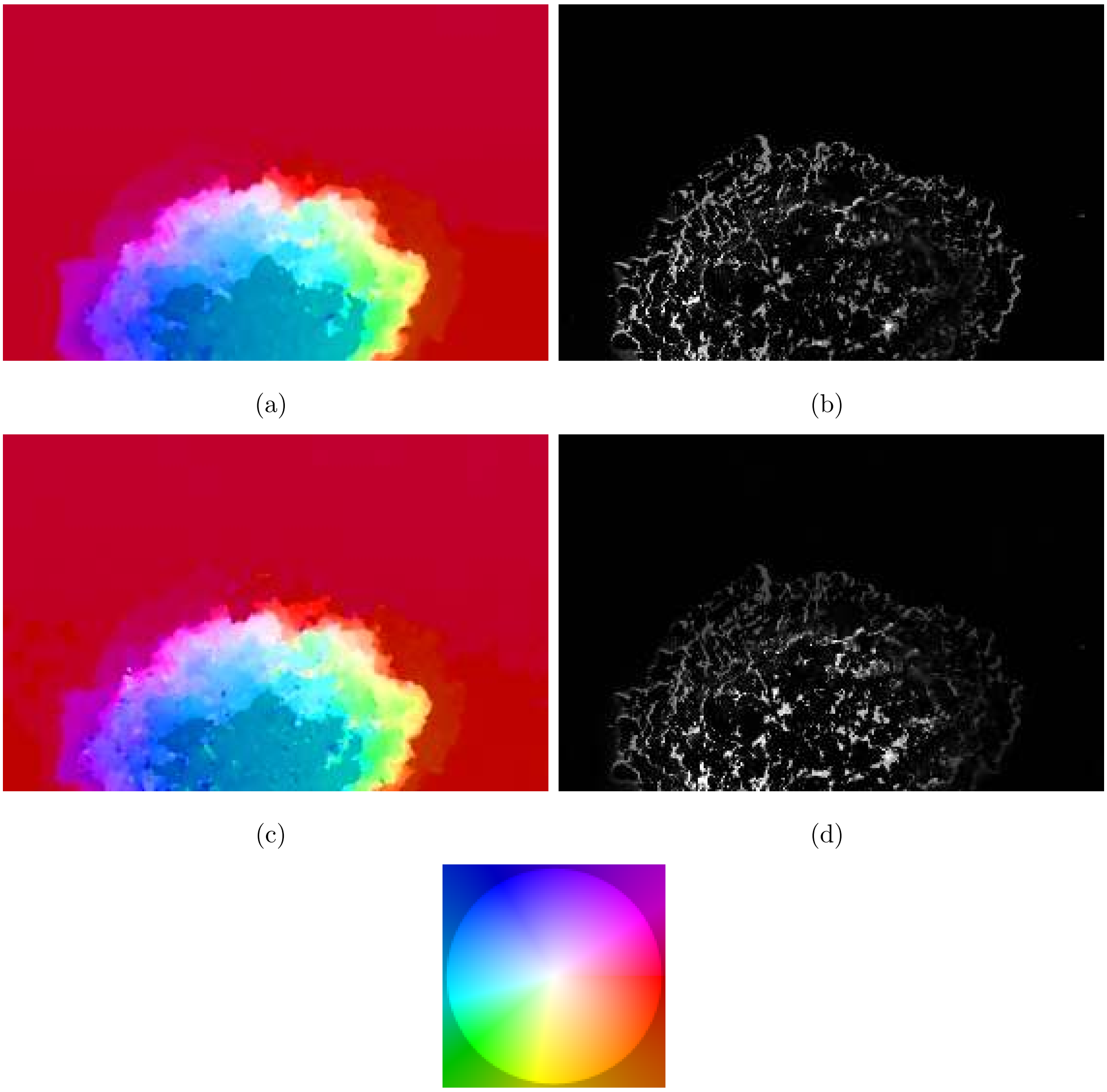
Fig. 5.3(a) shows the frame 10 of the *Hydrangea* sequence and its *ground truth* is displayed in Fig. 5.3(b). Fig. 5.3(c) shows its reference color wheel.

In Fig. 5.4 we show the optical flow for the *Hydrangea* sequence using the approach of [63] (Fig. 5.4(a)) and the steered Hermite approach (Fig. 5.4(c)).

Figs. 5.4(b) and 5.4(d) show the average error for both approaches. The approach [63] gets an error $AE = 6.94\,^\circ$ and using our proposal we get an error $AE = 6.88\,^\circ$

The frame 10 of the *Rubberwhale* sequence is shown in Fig. 5.5(a) whereas its *ground truth* is displayed in Fig. 5.5(b). Fig. 5.5(c) represents the color coding for the displacement.

The optical flow results are show in Fig. 5.6 for the approach of [63] (Fig. 5.6(a)) and using the steered Hermite approach (Fig. 5.6(c)). The corresponding errors are shown in Fig. 5.6(b) for [63] ($AE = 6.61\,^\circ$) and Fig. 5.6(d) for our approach ($AE = 7.69\,^\circ$).

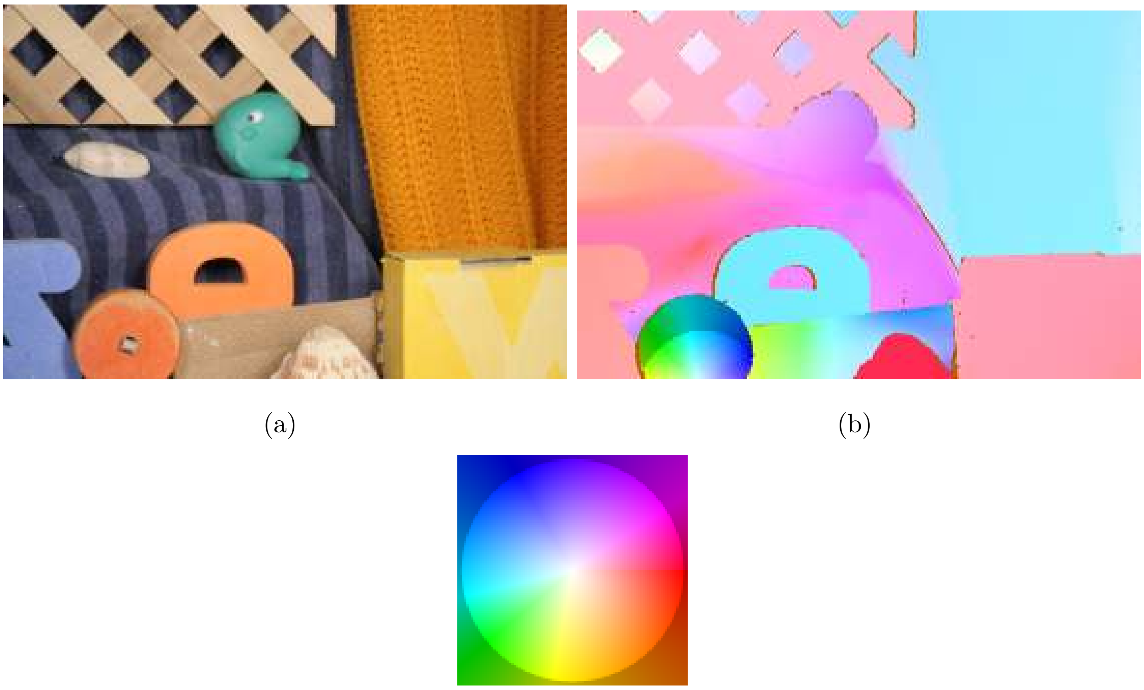


(a) (b)

(c) (d)

(e)

Figure 5.4: Optical flow fields of the *Hydrangea* sequence. (a) Optical flow approach of [63]. (b) Absolute error between the *ground truth* and the optical flow approach of [63] $AE = 6.94\,^{\circ}$. (c) Optical flow using the steered Hermite coefficients. (d) Absolute error between the *ground truth* and the optical flow using steered Hermite coefficients $AE = 6.88\,^{\circ}$. (e) Reference color wheel.

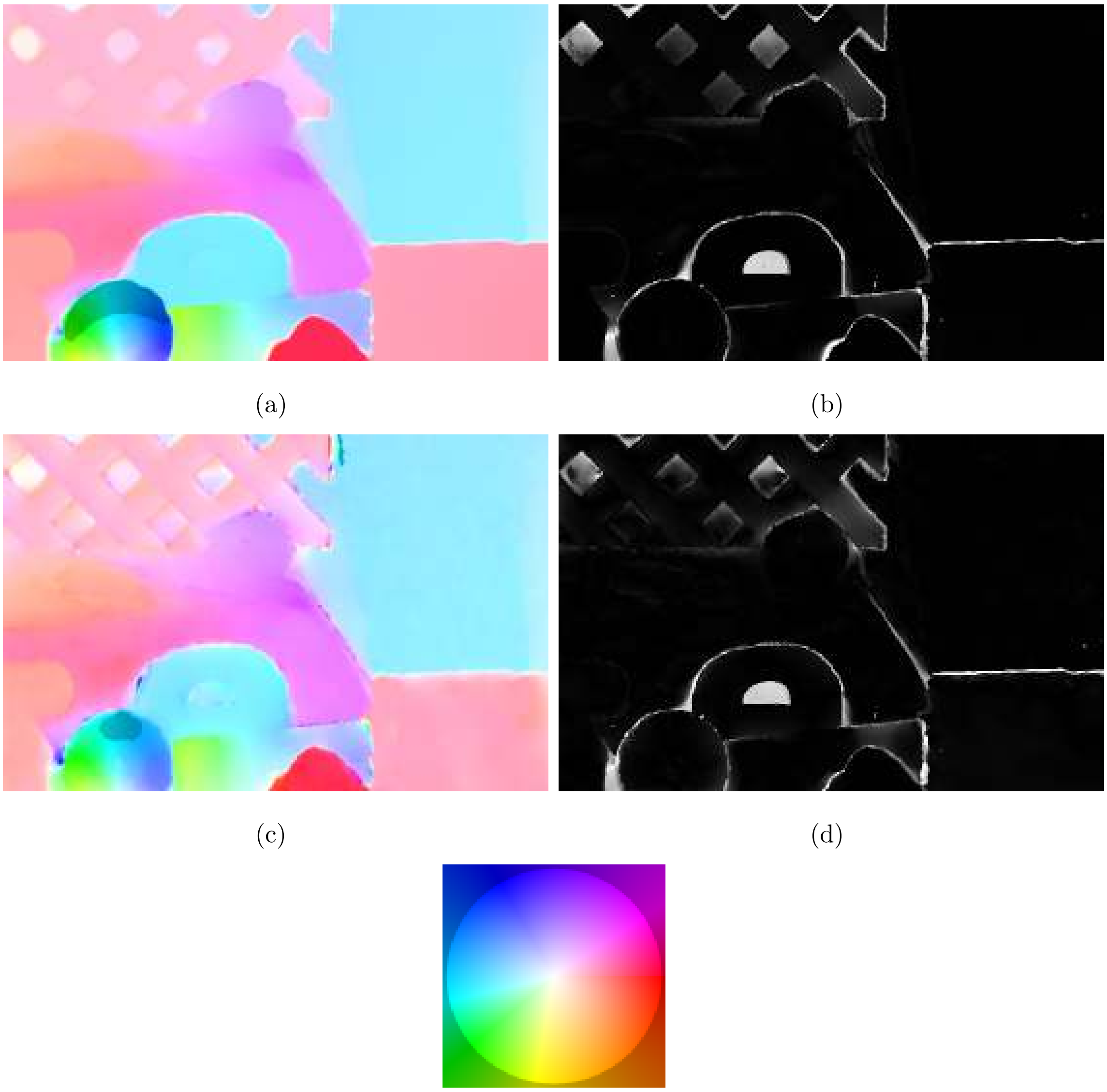


(a)

(b)

(c)

Figure 5.5: Optical flow of the sequence *Rubberwhale.* (a) Frame 10 of the sequence. (b) *Ground truth.* (c) Reference color wheel.



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

Figure 5.6: Optical flow fields of the *Rubberwhale* sequence. (a) Optical flow approach of [63]. (b) Absolute error between the *ground truth* and the optical flow approach of [63] $AE = 6.61\,^{\circ}$. (c) Optical flow using the steered Hermite coefficients. (d) Absolute error between the *ground truth* and the optical flow using steered Hermite coefficients $AE = 7.69\,^{\circ}$. (e) Reference color wheel.

| **Sequence** | **Angular error (°) [63]** | **Angular error (°) Hermite** |
|---|---|---|
| *Dimetrodon* | 3.03 | 2.70 |
| *Hydrangea* | 6.94 | 6.88 |
| *Rubberwhale* | 6.61 | 7.69 |

Table 5.2: Angular errors.

Table 5.2 shows the angular error from *Dimetrodon*, *Hydrangea* and *Rubberwhale* sequences, first applying the algorithm of [63] and then using the steered Hermite transform.

## 5.3 Optical flow estimation in cardiac CT images

Getting a picture of the heart has always been a technical challenge because the heart has a continuous movement. CT images of the heart in motion have been used in the diagnosis with the development of MSCT synchronized with the Electrocardiography (ECG) examination. One of the main advantages of the introduction of MSCT is the speed of acquisition of the images, thus shortening examination of the patient to reduce the amount of contrast required to increase and enhance vascular consistency. The use of MSCT of the heart synchronized with the patient's ECG allows reconstruction of transverse images of the entire cardiac volume during any phase of the cardiac cycle. Usually, the entire stack of images is only reconstructed during diastole to freeze the heart's motion. By performing a reconstruction of the volume during systole and diastole function it is possible to find basic parameters of cardiac function, such as left and right ventricular ejection fraction and the myocardial wall thickness [71].

The strongest cardiac movement is present during contraction of the atria and ventricles in systole, approximately between 0% and 30% of cardiac cycle. The short end–systolic rest phase is followed by a continuous filling phase of the ventricles during diastole that slows down towards mid- and end–diastole. This movement is seen in the displacement of the left ventricular wall, the aortic valve flaps, and the different

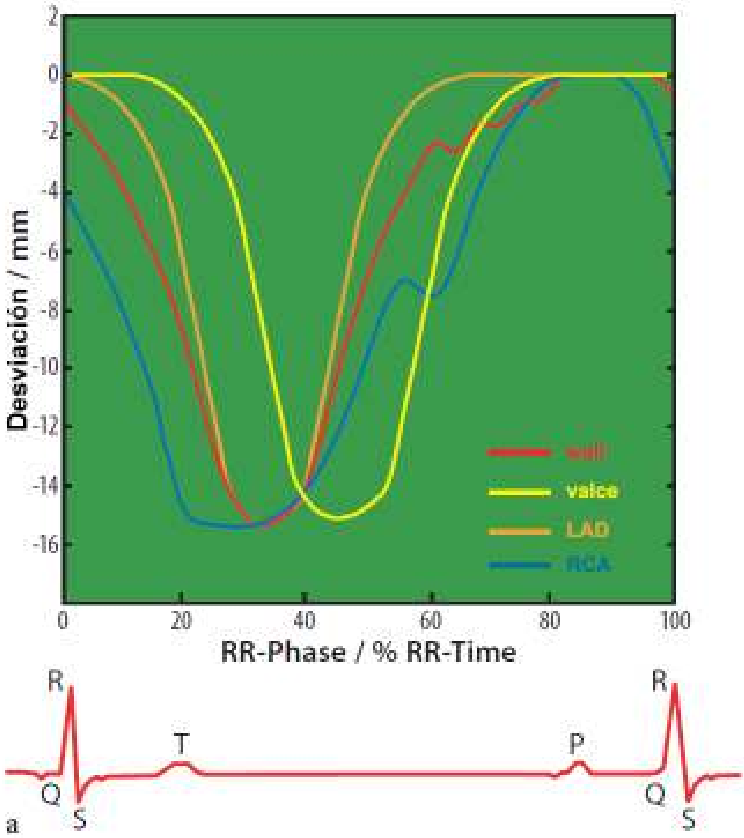


Figure 5.7: Cardiac motion during the different phases of the cardiac cycle [62]. Left ventricular wall (Wall). Aortic valve flaps (Valce). Left anterior descending coronary artery (LAD). Right coronary artery (RCA).

segments of the coronary arteries in representative transaxial planes (Fig. 5.7). The least amount of movement of the major cardiac anatomy and the coronary arteries, and thus the least amount of dislocation over time, is observed in end-systole and mid- to end–diastole of the cardiac cycle (Fig. 5.7) [62].

Due to the very complex 3D motion pattern of the heart, the intensity of movement varies for different cardiac anatomies and different coronary vessels, and within the cardiac cycle. Therefore, we evaluated the optical flow at different times of the cardiac cycle.

In order to detect the optical flow in cardiac CT images taken at different times of the cardiac cycle, we consider that the movement is due to the change of position of different cardiac structures and not due to the relative motion between the observer and the scene or to the variations in scene illumination as is the case of synthetic images. The images correspond to the axial plane (see Appendix B) of cardiac CT.

In Fig. 5.8 we show the cartesian Hermite coefficients of the slice 52 of the CT sequence at 20% of the cardiac cycle.

In Fig. 5.9 we steered the cartesian Hermite coefficients of Fig. 5.8 according to maximum energy direction.

To verify the constancy assumptions between two images in different times of the cardiac cycle we compared their Hermite coefficients. Fig. 5.10 shows the coefficients $L_0$ of slice 52 of the CT sequence at 20% and 30% of the cardiac cycle, where there is a displacement $(u, v)$ of the pixels in the position $(x, y)$ from a time $t$ to a time $(t + 1)$. Fig. 5.11 shows the steered coefficients $l_{n,\theta}$ for both periods of the cardiac cycle.

In order to show the different movements of cardiac structures, we compute the optical flow at different times of the cardiac cycle. In the following paragraphs we show the optical flow obtained using our implementation and compared with the implementation of [63] in 2D.

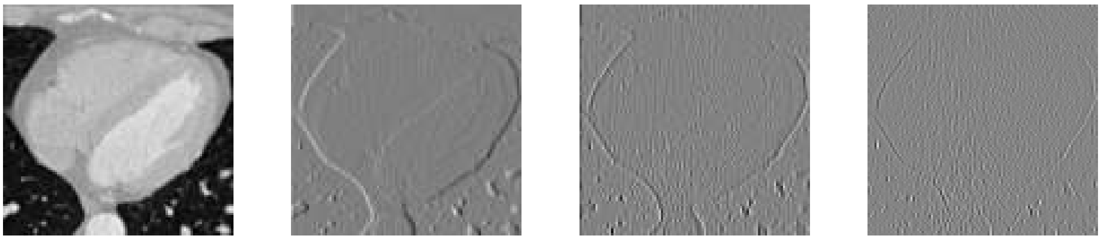
Due to restrictions used we can determine the optical flow regularized in other areas that not dependent on the textures or artifacts of the images, such as, homogeneous areas that are also of interest for medical diagnosis, despite the noise and the irregular motion of the images used.

## 0% of the cardiac cycle

Figs. 5.12(a) and 5.12(b) show slice 52 of the CT sequence at 0% and 10% of the cardiac cycle (systole) respectively.

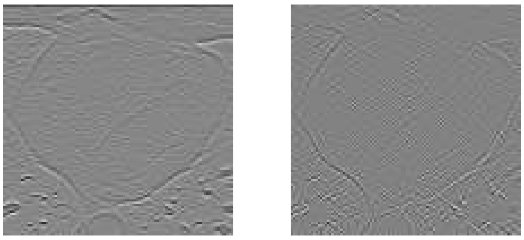
The resulting optical flow applying the approach of [63] and the steered Hermite transform approach are shown in Figs. 5.13(a) and 5.13(b) respectively.

From a visual evaluation we notice that our algorithm shows clearer and better defined flows especially in areas close to the cavities, such as the left ventricle, being the study of movement of this structure of major importance to physicians.



(a) $L_{0,0}(x, y)$ (b) $L_{1,0}(x, y)$ (c) $L_{2,0}(x, y)$ (d) $L_{3,0}(x, y)$

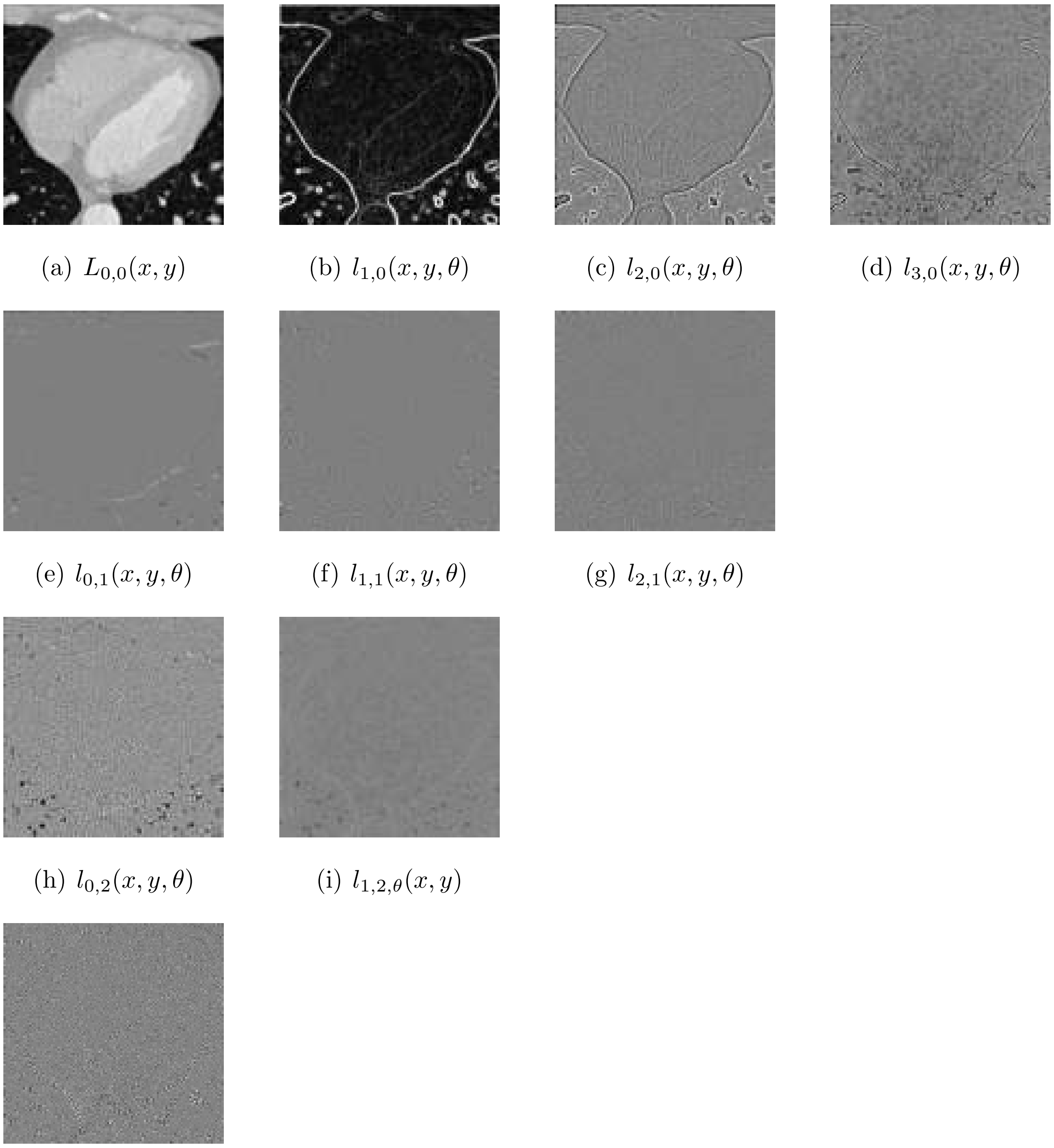
(e) $L_{0,1}(x, y)$ (f) $L_{1,1}(x, y)$ (g) $L_{2,1}(x, y)$



(h) $L_{0,2}(x, y)$ (i) $L_{1,2}(x, y)$

(j) $L_{0,3}(x, y)$

Figure 5.8: Cartesian Hermite coefficients of the slice 52 of the CT sequence at 20% of the cardiac cycle for $N = 3$ ($n = 0, 1, 2, 3$). (a) $L_{0,0}(x, y)$ (represents the DC Hermite coefficient). (b) $L_{1,0}(x, y)$. (c) $L_{2,0}(x, y)$. (d) $L_{3,0}(x, y)$. (e) $L_{0,1}(x, y)$. (f) $L_{1,1}(x, y)$. (g) $L_{2,1}(x, y)$. (h) $L_{0,2}(x, y)$. (i) $L_{1,2}(x, y)$. (j) $L_{0,3}(x, y)$.



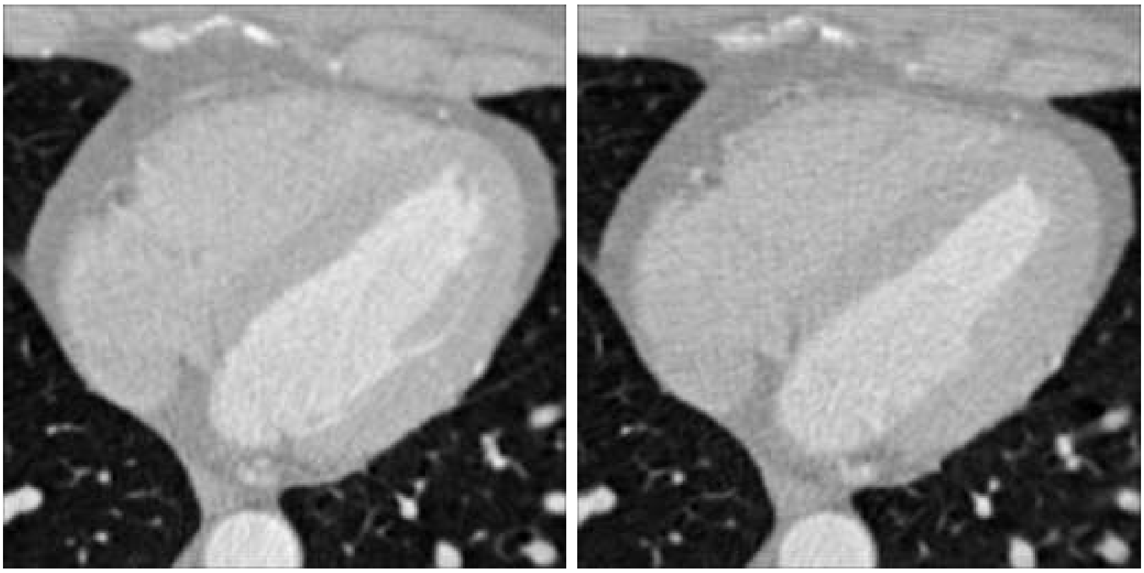
(a) $L_{0,0}(x, y)$ (b) $l_{1,0}(x, y, \theta)$ (c) $l_{2,0}(x, y, \theta)$ (d) $l_{3,0}(x, y, \theta)$

(e) $l_{0,1}(x, y, \theta)$ (f) $l_{1,1}(x, y, \theta)$ (g) $l_{2,1}(x, y, \theta)$

(h) $l_{0,2}(x, y, \theta)$ (i) $l_{1,2,\theta}(x, y)$

(j) $l_{0,3,\theta}(x, y)$

Figure 5.9: Steered Hermite coefficients of the slice 52 of the CT sequence at 20% of the cardiac cycle for $N = 3$ ($n = 0, 1, 2, 3$). (a) $l_{0,0}(x, y)$ (represents the DC Hermite coefficient). (b) $l_{1,0,\theta}(x, y)$. (c) $l_{2,0,\theta}(x, y)$. (d) $l_{3,0,\theta}(x, y)$. (e) $l_{0,1,\theta}(x, y)$. (f) $l_{1,1,\theta}(x, y)$. (g) $l_{2,1,\theta}(x, y)$. (h) $l_{0,2,\theta}(x, y)$. (i) $l_{1,2,\theta}(x, y)$. (j) $l_{0,3,\theta}(x, y)$.



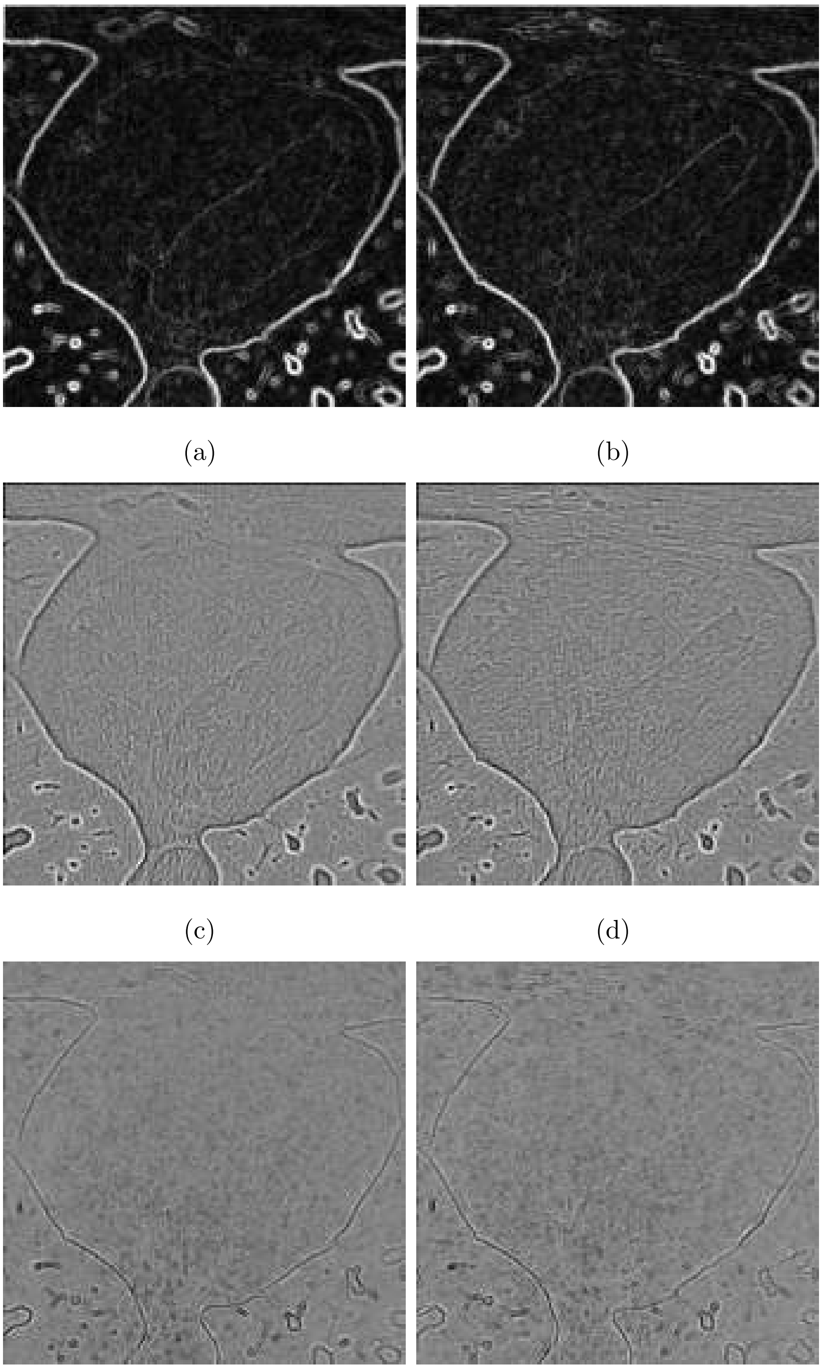
(a) (b)

Figure 5.10: Cartesian Hermite coefficients $L_{00}$ of the slice 52 of the CT sequence at 20% and 30% of the cardiac cycle. (a) $L_{0,0}(x, y, t)$. (b) $L_{0,0}(x + u, y + v, t + 1)$.

Fig. 5.14 shows the isolated displacements of the left ventricle. For this purpose we use a binarization of the image and apply algorithms of mathematical morphology for a simple segmentation of the left ventricle. Fig. 5.14(a) shows the optical flow resulting from the algorithm of [63] and Fig. 5.14(b) the results of the steered Hermite transform algorithm.

In the optical flow field the size of the arrows is proportional to the magnitude of the displacement, which results in an overlap between the vectors. To avoid this we show the velocity field using a color code (Fig. 5.15). Figs. 5.15(a) and 5.15(b) show the color code results corresponding to the displacements of Fig. 5.14.

The left ventricular wall contraction in systole can be seen in Figs. 5.13, 5.14 and 5.15.



(a) (b)

(c) (d)

(e) (f)

Figure 5.11: Steered Hermite coefficients up to order $N = 3$ of the slice 52 of the CT sequence at 20% and 30% of the cardiac cycle. (a) $l_{1,\theta}(x, y, t)$. (b) $l_{1,\theta}(x{+}u, y{+}v, t{+}1)$. (c) $l_{2,\theta}(x, y, t)$. (d) $l_{2,\theta}(x + u, y + v, t + 1)$. (e) $l_{3,\theta}(x, y, t)$. (f) $l_{3,\theta}(x + u, y + v, t + 1)$.

### 20% of the cardiac cycle

Figs. 5.17, 5.18 and 5.19 show the estimated optical flow between 20% and 30% corresponding to cardiac cycle of slice 52 shown on Fig. 5.16, also showing the left ventricular wall contraction.

### 50% of the cardiac cycle

The movement of relaxation during diastole can be seen between 40% to 60% of the cardiac cycle [62]. Figs. 5.21, 5.22 and 5.23 show the optical flow between 50% and 60% of cardiac cycle of slice 52 shown on Fig. 5.20 using both the approach of [63] and the steered Hermite transform. According to [62] the least amount of movement is observed in end-systole and mid- to end-diastole of the cardiac cycle.

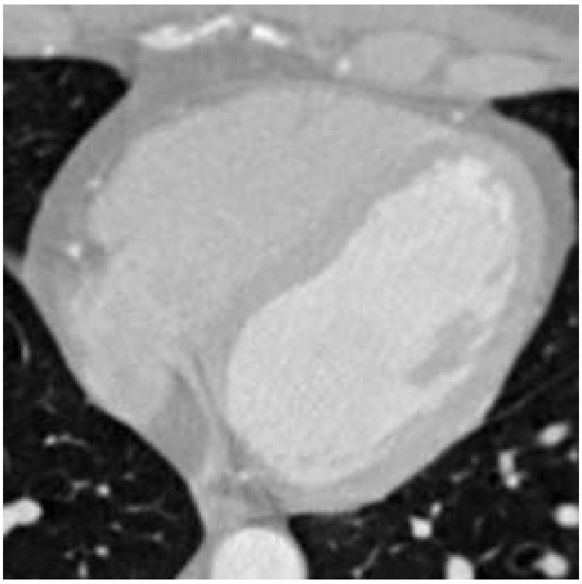
## 5.4 Quantitative evaluation of optical flow in cardiac CT images

In order to provide a quantitative measure of our proposal we calculate the Root Mean Squared *RMS* of the *Backward reconstruction* using the optical flow algorithm of [63] and the steered Hermite coefficients. The computed optical flow for a particular image in a sequence is used to estimate the next image in that sequence. Then the *RMS* between this estimated image and the actual next image is calculated.

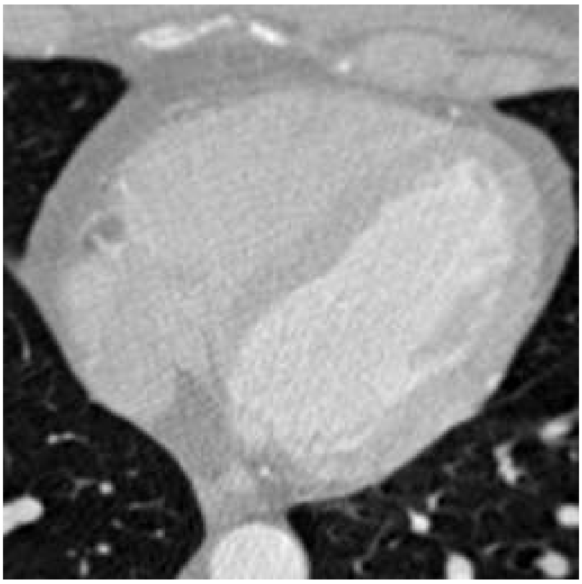
For the reconstruction algorithm we used the *Backward reconstruction* [46]. Let us consider two adjacent images in the CT sequence $L(x, y, t)$ and $L(x, y, t+1)$, and the optical flow $(u, v)$ between them. Given the values pixels of second image, we can calculate the gray value at $(x, y)$ in the first image using spline interpolation.

$$RMS\ error = \sqrt{\frac{\sum_x \sum_y \left(L(x,y,t) - \hat{L}(x,y,t)\right)^2}{\mathcal{M} \times \mathcal{N}}} \tag{5.3}$$

where $L(x, y, t)$ and $\hat{L}(x, y, t)$ are the true and reconstructed images of size $\mathcal{M} \times \mathcal{N}$ at time $t$.

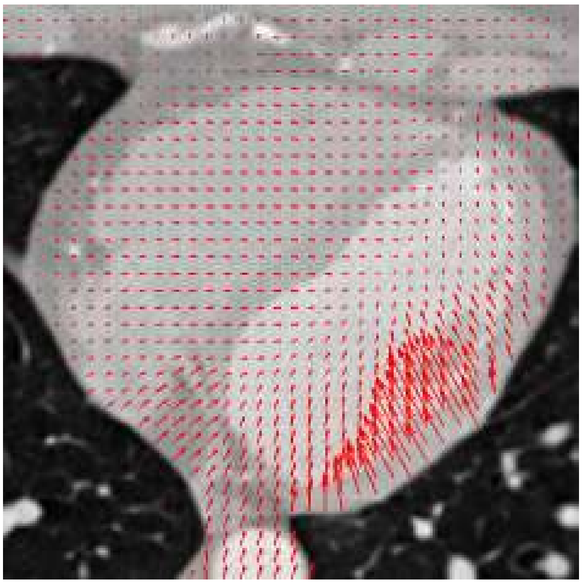


(a)

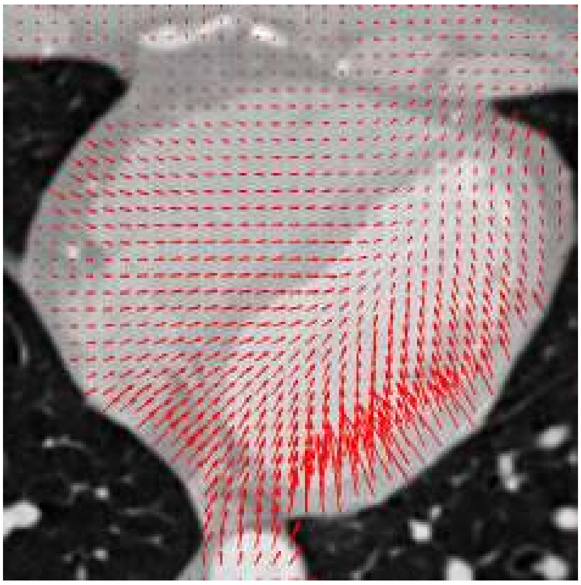


(b)

Figure 5.12: Slice 52 of the CT sequence. (a) 0% of the cardiac cycle. (b) 10% of the cardiac cycle.

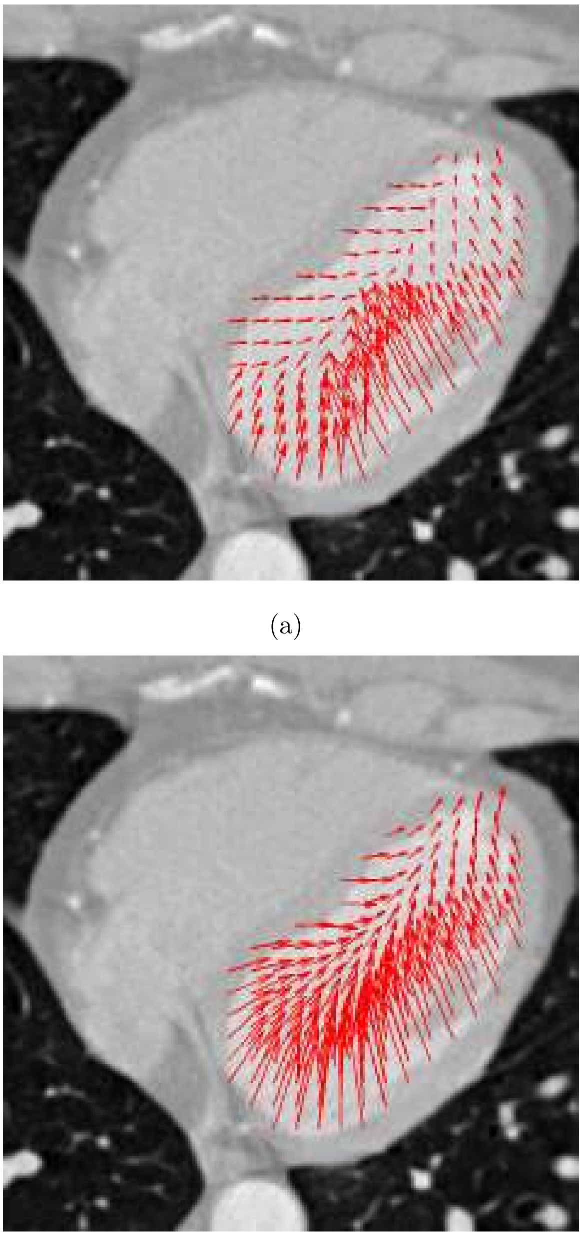


(a)



(b)

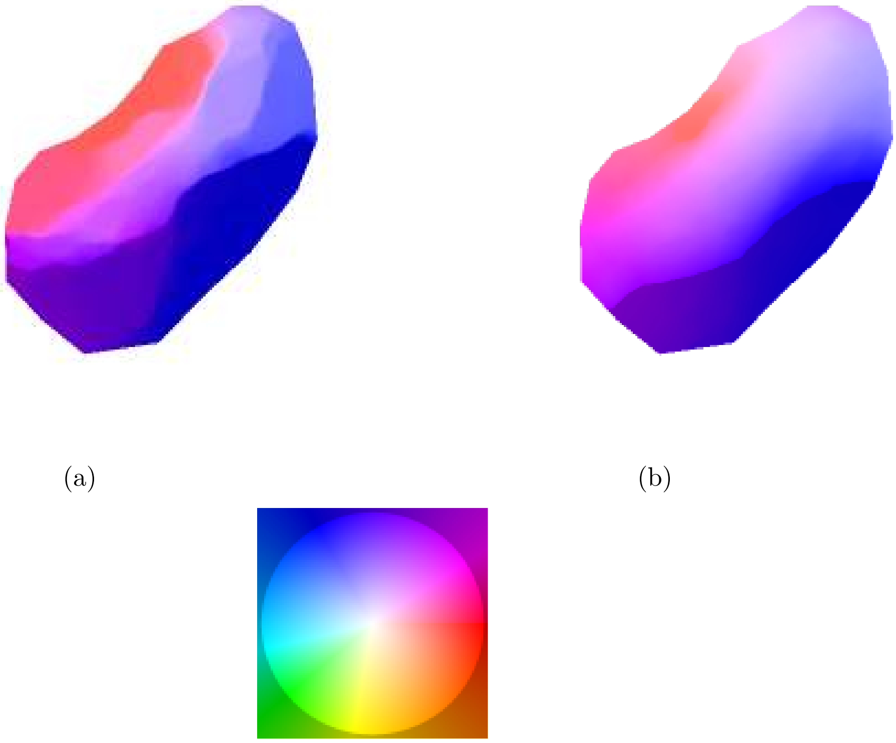
Figure 5.13: Optical flow of slice 52 of the CT sequence between 0% and 10% of the cardiac cycle. (a) Optical flow algorithm of [63]. (b) Optical flow using steered Hermite transform approach.



(a)

(b)

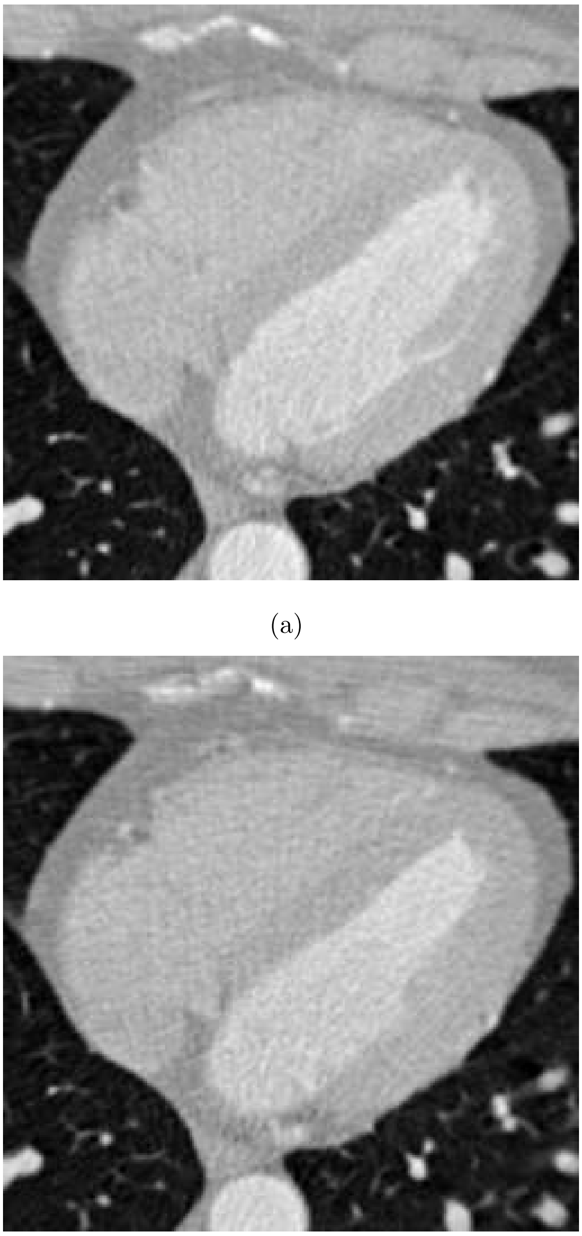
Figure 5.14: Optical flow left ventricle of slice 52 of the CT sequence between 0% and 10% of the cardiac cycle. (a) Optical flow left ventricle algorithm of [63]. (b) Optical flow left ventricle using steered Hermite coefficients.



(a) (b)

(c)

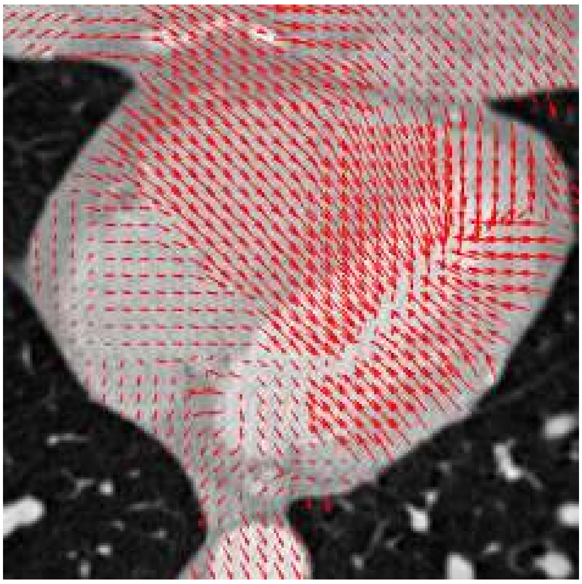
Figure 5.15: Optical flow left ventricle of slice 52 of the CT sequence between 0% and 10% of the cardiac cycle (color coding). (a) Optical flow left ventricle algorithm of [63] (color coding). (b) Optical flow left ventricle using steered Hermite coefficients (color coding). (c) Reference color wheel.



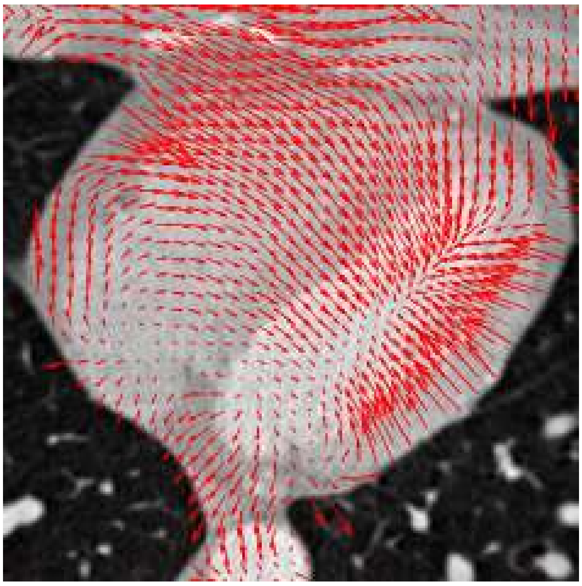
(a)

(b)

Figure 5.16: Slice 52 of the CT sequence. (a) 20% of the cardiac cycle. (b) 30% of the cardiac cycle.

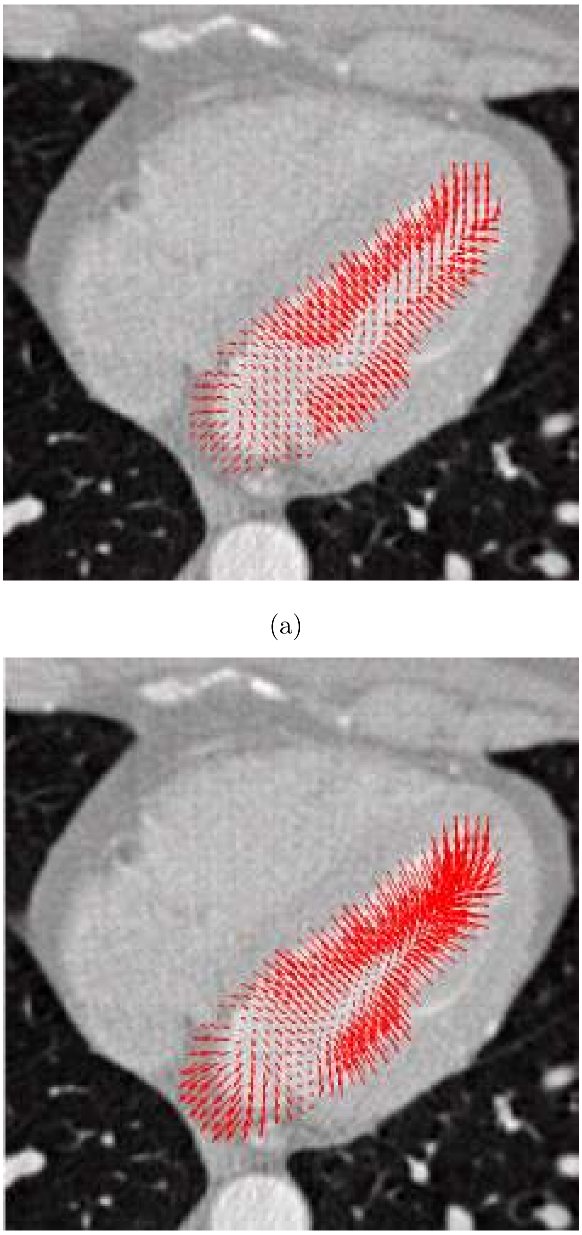


(a)



(b)

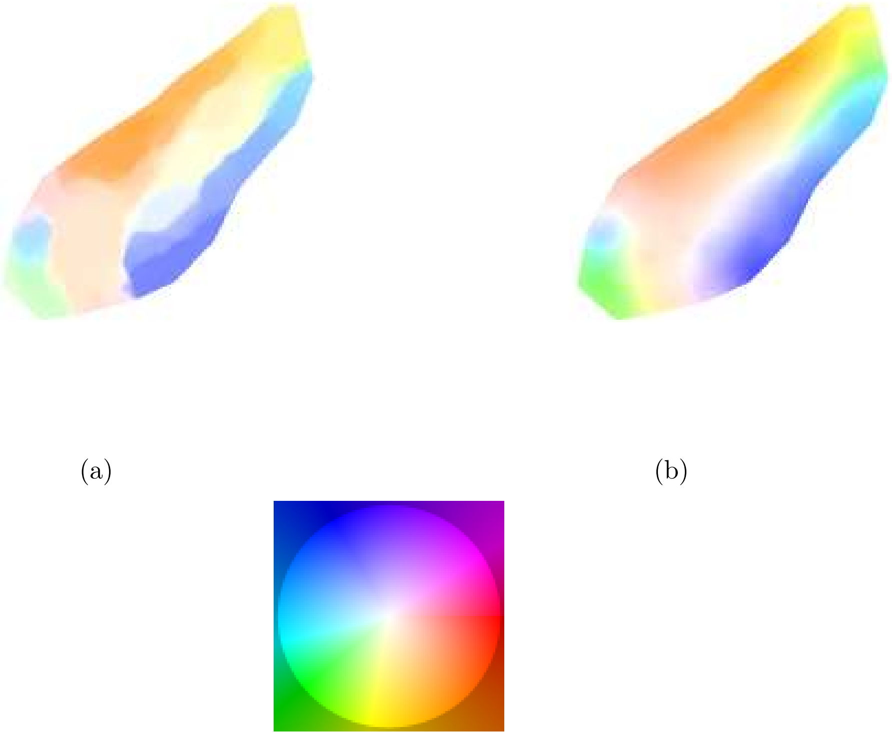
Figure 5.17: Optical flow of slice 52 of the CT sequence between 20% and 30% of the cardiac cycle. (a) Optical flow algorithm of [63]. (b) Optical flow using steered Hermite transform approach.



(a)

(b)

Figure 5.18: Optical flow left ventricle of slice 52 of the CT sequence between 20% and 30% of the cardiac cycle. (a) Optical flow left ventricle algorithm of [63]. (b) Optical flow left ventricle using steered Hermite coefficients.

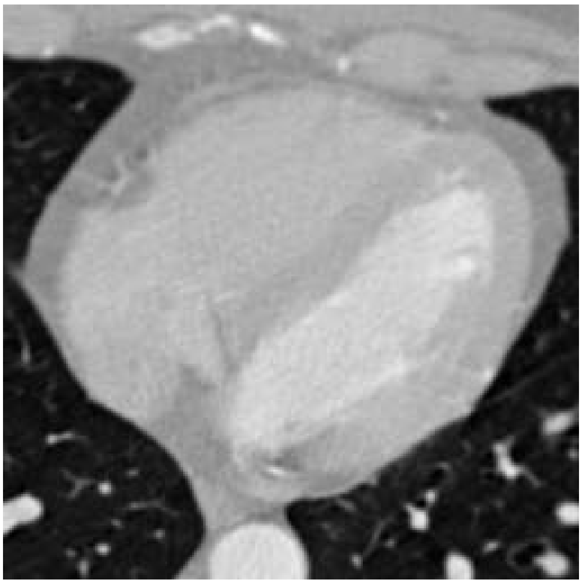


(a)

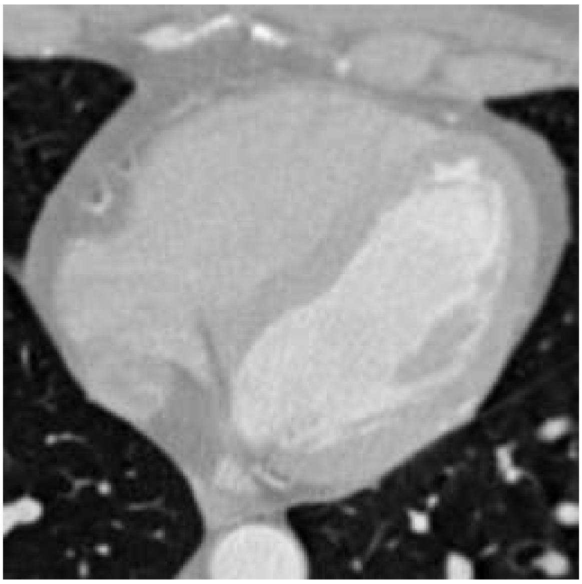
(b)

(c)

Figure 5.19: Optical flow left ventricle of slice 52 of the CT sequence between 20% and 30% of the cardiac cycle (color coding). (a) Optical flow left ventricle algorithm of [63] (color coding). (b) Optical flow left ventricle using steered Hermite coefficients (color coding). (c) Reference color wheel.

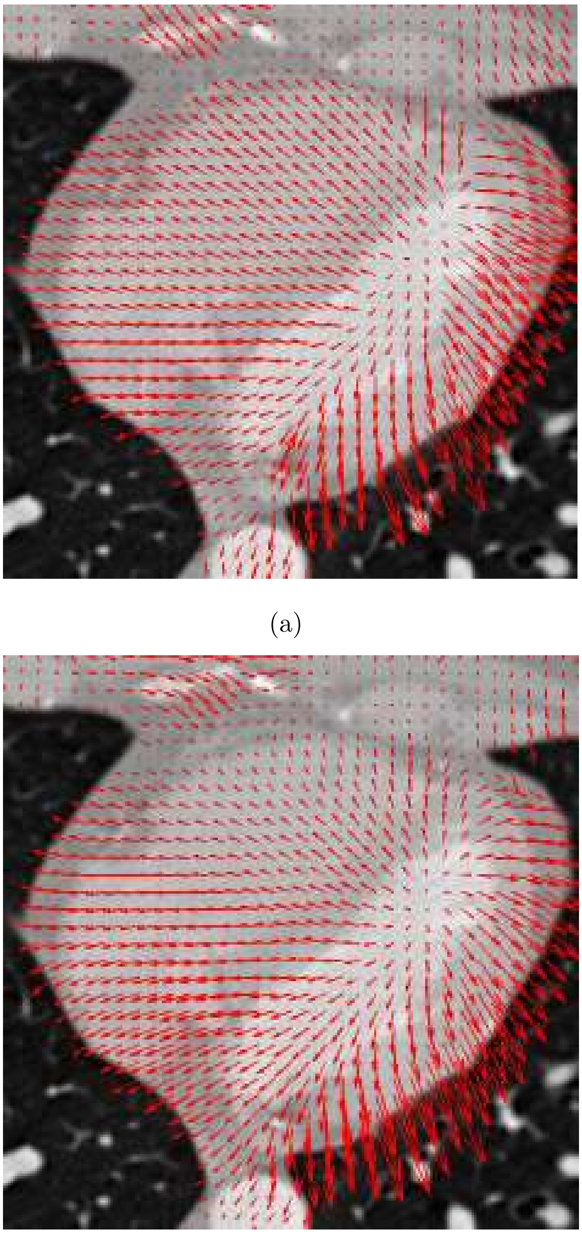


(a)



(b)

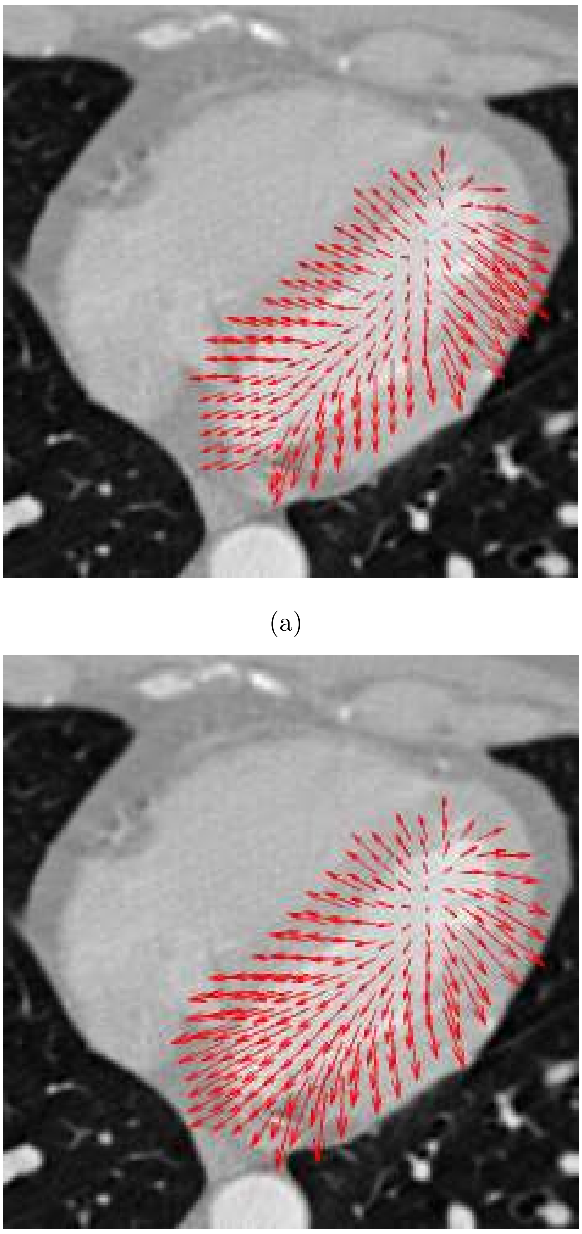
Figure 5.20: Slice 52 of the CT sequence. (a) 50% of the cardiac cycle. (b) 60% of the cardiac cycle.



(a)

(b)

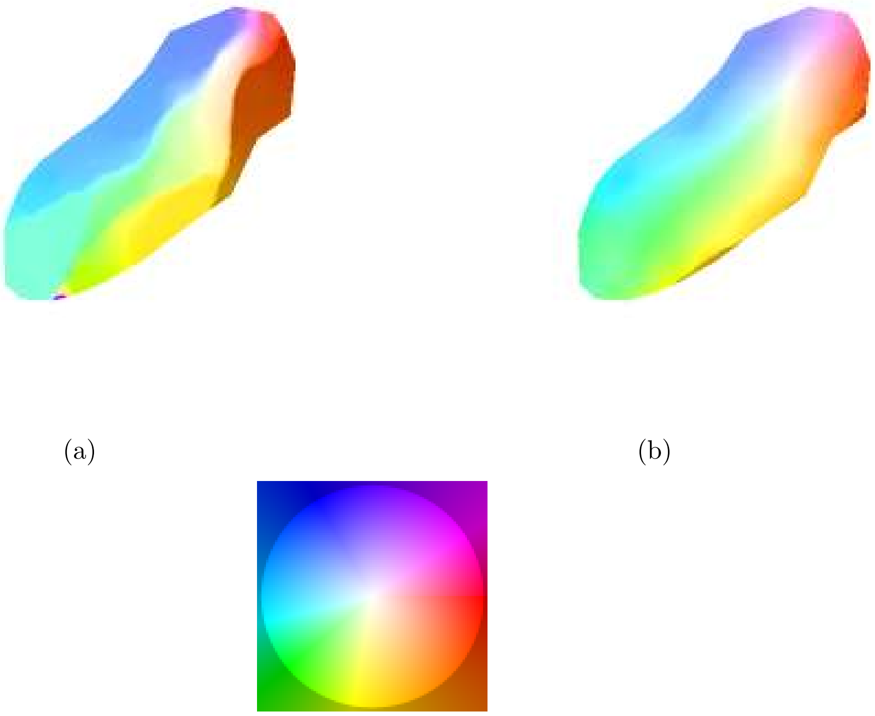
Figure 5.21: Optical flow of slice 52 of the CT sequence between 50% and 60% of the cardiac cycle. (a) Optical flow algorithm of [63]. (b) Optical flow using steered Hermite transform approach.



(a)

(b)

Figure 5.22: Optical flow left ventricle of slice 52 of the CT sequence between 50% and 60% of the cardiac cycle. (a) Optical flow left ventricle algorithm of [63]. (b) Optical flow left ventricle using steered Hermite coefficients.



(a) (b)

(c)

Figure 5.23: Optical flow left ventricle of slice 52 of the CT sequence between 50% and 60% of the cardiac cycle (color coding). (a) Optical flow left ventricle algorithm of [63] (color coding). (b) Optical flow left ventricle using steered Hermite coefficients (color coding). (c) Reference color wheel.

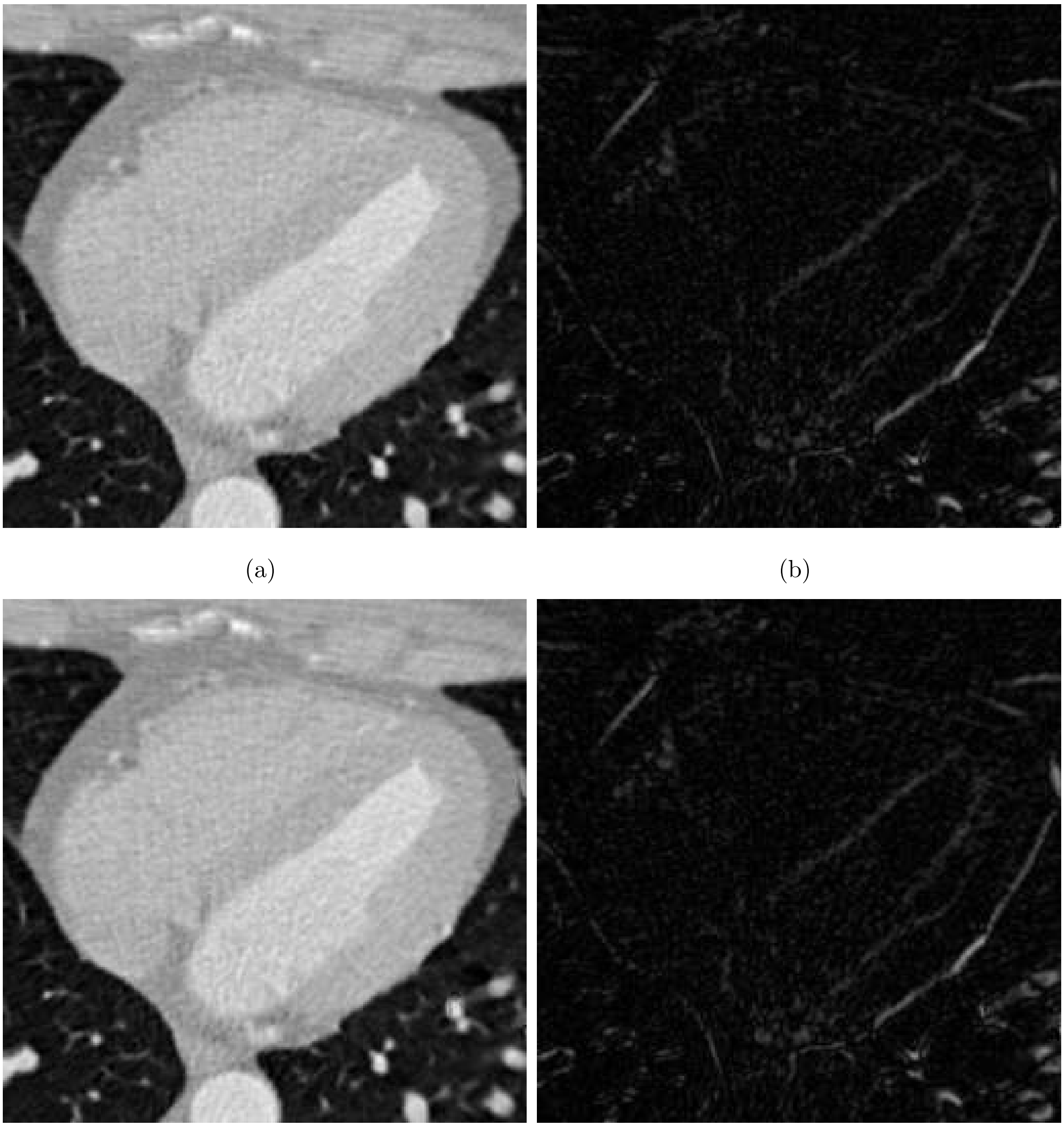
| **% cardiac cycle** | **RMS error [63]** | **RMS error Hermite** |
|---|---|---|
| 0%-10% | 28.6214 | 26.0571 |
| 20%-30% | 17.5157 | 16.9614 |
| 50%-60% | 19.5726 | 19.2855 |

Table 5.3: Reconstruction RMS error.

Fig. 5.24 shows the reconstruction error for the optical flow at 20% of the cardiac cycle of the slice 52 for the approach of [63] and using the steered Hermite coefficients.

Table 5.3 shows the reconstruction error for the optical flow from 0%-10%, 20%-30% and 50%-60% of the cardiac cycle images of slice 52, first applying the algorithm of [63] and then using the steered Hermite transform.

In Table 5.3 we note that the RMS error is smaller in our approach, this is because when using the coefficients of order larger than 2 the estimate is more robust to noise. It is important to stress that in spite of the noisy nature of CT images our algorithm has shown accuracy.



(a) (b)

(c) (d)

Figure 5.24: Reconstruction error slice 52 of the CT sequence 20% of the cardiac cycle. (a) *Backward reconstruction* using the optical flow of [63]. (b) Absolute error between the reconstructed first image and the true first image (20% of the cardiac cycle) using approach of [63]. (c) *Backward reconstruction* using the steered Hermite transform. (d) Absolute error between the reconstructed first image and the true first image (20% of the cardiac cycle).

# Chapter 6

# Conclusions

> Your beliefs become your thoughts,
> Your thoughts become your words,
> Your words become your actions,
> Your actions become your habits,
> Your habits become your values,
> Your values become your destiny.
>
> Mahatma Gandhi (1869–1948).

This thesis proposed a method of optical flow estimation using the steered Hermite transform. The Hermite transform (Martens [52, 51]) is a model that incorporates some of the more important properties of the first stages of the human visual system, such as the overlapping Gaussian receptive fields, the Gaussian derivative model of early vision (Young [85, 86, 87]), and the multiresolution analysis (Escalante-Ramírez [24], Silván-Cárdenas [73]).

We have followed a continuous, rotationally invariant differential approach for motion estimation which incorporates certain restrictions such as constraints of intensity and gradient constant, and piecewise smooth optical flow field, in order to handle the opening problem. In order to improve the robustness of classical methods, our approach use additional nonlinear restrictions suggested by novel differential approaches. They include higher order derivatives on local image constraints to achieve intensity change invariance, the use of isotropic flow–driven regularization to reduce smoothing at motion boundaries and robust optimization functions to be more lenient with respect to outliers in the solution. These constancy assumptions are not linearized thus allowing implementing a multiresolution strategy to handle large displacements.

Image primitive structures extracted from the steered Hermite coefficients have been used as local image constraints, in order to make the algorithm more robust to noise. This has improved the estimation performance, allowing the algorithm to analyze perceptually relevant oriented image structures. Varying the spatial scale of the Gaussian window allows analyzing objects of different spatial dimensions. We use this feature to build the estimation method on a multiresolution approach, starting from the lowest resolution we define a linear system and the solution is interpolated and propagated to the next finer level.

Estimation performance has been tested in several ways. Angular error was calculated from several sequences with ground truth. Smaller values were found with our method than with one of the best state-of-the-art competitive approaches in the majority of cases.

For highly textured images our method allows obtaining an angular error smaller than that obtained with the algorithm of [63]. This is due to the fact that the steered Hermite transform obtains local orientations of structures in noise or high variability of the intensities. We ran objective and subjective tests on CT images. From a visual evaluation we notice that our algorithm shows clearer and better defined flows especially in areas close to the cavities, such as the left ventricle, being the study of movement of this structure of major importance to physicians. In order to perform an objective test on these images we reconstructed sequential images from their adjacent images and their motion vectors. We then measured the backward reconstruction error and found smaller errors in our proposal than in competitive approaches. It is important to stress that in spite of the noisy nature of CT images our algorithm has shown accuracy.

# Appendix A

# Computed tomography

> Our reward is in the effort, not outcome. A total effort is full victory.
>
> Mahatma Gandhi (1869–1948).

One of the most important advances in diagnostic radiology in recent decades has been the cross-sectional imaging of the human body. After the introduction of computed tomography (CT), ultrasound (US) and magnetic resonance imaging (MRI) diagnostic proposals, treatment and monitoring of disease have changed completely. In particular, spiral computed tomography has become a useful tool for cross-sectional imaging, this due to its robustness, which is moderately invasive and relatively inexpensive. In particular, CT angiography was significantly enhanced by the rotational speed of the order of sub-seconds and the fineness of the collimation [19].

In 1895, William Roentgen produced and detected electromagnetic radiation in a wavelength range today known as x rays or Roentgen rays. While investigating materials that could stop these rays he passed his hand in front of the x-ray source and produced the first x-ray image. He was awarded the Nobel Prize for physics in 1901. Clinical CT itself was invented by Godfrey Hounsfield in 1972. He also received the Nobel Prize (for medicine) in 1979 along with Allan Cormack (Fig. A.1) [8].

In a CT, an X-ray beam is rotated rapidly around the patient's body, and the data

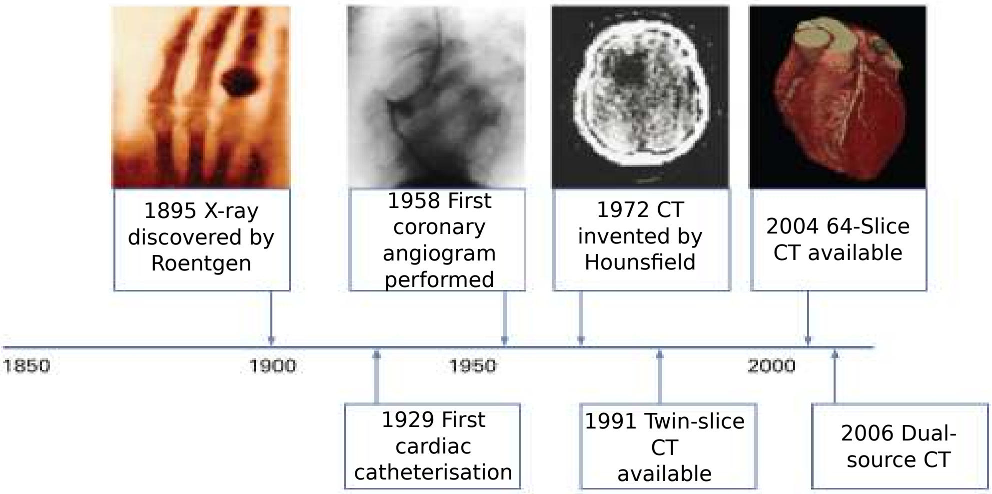


Figure A.1: CT timeline [8].

are transmitted to a computer, where an algorithm assigns to each point (voxel) of a cross section of the image plane a value in a gray scale (Hounsfield unit), indicating the attenuation of the X-rays on tissue (Fig. A.2) [70].

Initial CT scanners (Fig. A.3) were limited to tomographic sections of the brain that took several minutes to acquire. Slip–ring technology replaced the cables of older scanners and permitted fast spiral scanning whereby the x–ray source and detector continuously rotate around the patient while the patient table moves through the scanner [8].

Spiral CT generates cross–sectional images of a volume by obtaining multiple measurements of the x-rays attenuation from several projections. The requirement for image reconstruction is that all the above-mentioned measurements should lie in the same plane, which is not the case in spiral CT since scanning is performed with simultaneous patient translation. The complete in–plane data set of measurements is obtained through interpolation from the measurements that precede and follow that plane. Given the need for interpolation, we should consider that the more distant the measurements, the less accurate the interpolated values. The distance between the measurements depends on the table feed and on the gantry rotation period [19].

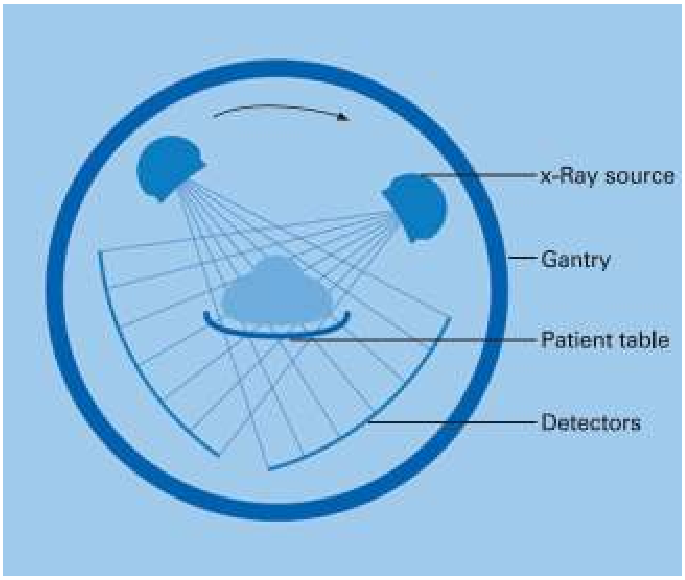


Figure A.2: CT schematic [8].

In spiral CT, images can be reconstructed at any z–axis position within the scanned volume since missing information is obtained through interpolation. In MSCT, image reconstruction is performed using a filtered backprojection kernel[1]. A convolution filtering before the backprojection is needed to overcome the additive nature of the reconstruction technique. Image filtering is generally used to enhance specific features such as edges. However, image filtering affects image pixel noise and, as a consequence, image contrast [19].

The scan speed was further improved by changing the shape of the radiation beam (to a fan shape) and increasing the number and quality of x-ray detectors in what are called multidetector (or multislice) CT scanners (Fig. A.4), where the rotation speed has dropped to less than half a second and the resolution in the sub-millimeter order, allowing scanned large areas in great detail. Developed in 1991, the Elscint CT Twin was the first multislice (MSCT) scanner and had two parallel banks of x-ray detectors to acquire two slices for each gantry rotation. Quickly the technology developed with scanners with four slices, then 16 slices, and current scanners having

---

[1] Backprojection is an approximated algorithm that assigns to a defined pixel the values that are collected along projection lines passing through the same pixel.

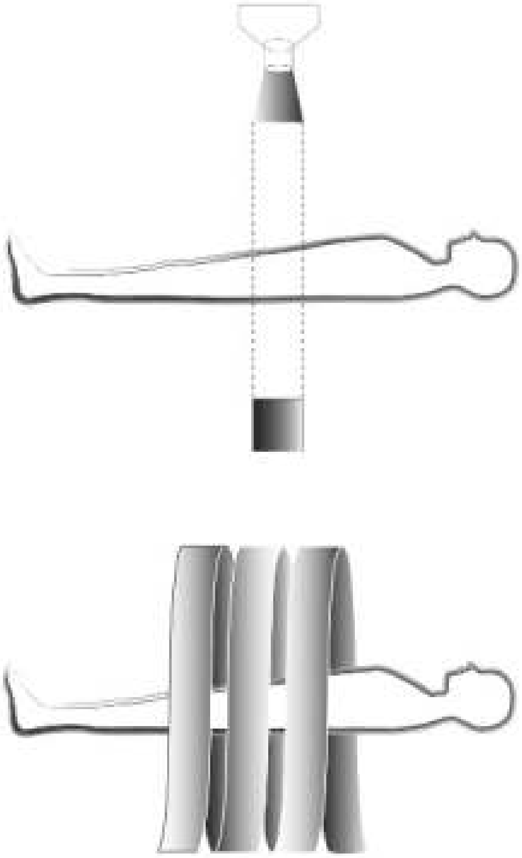


Figure A.3: Single CT scanner [8].

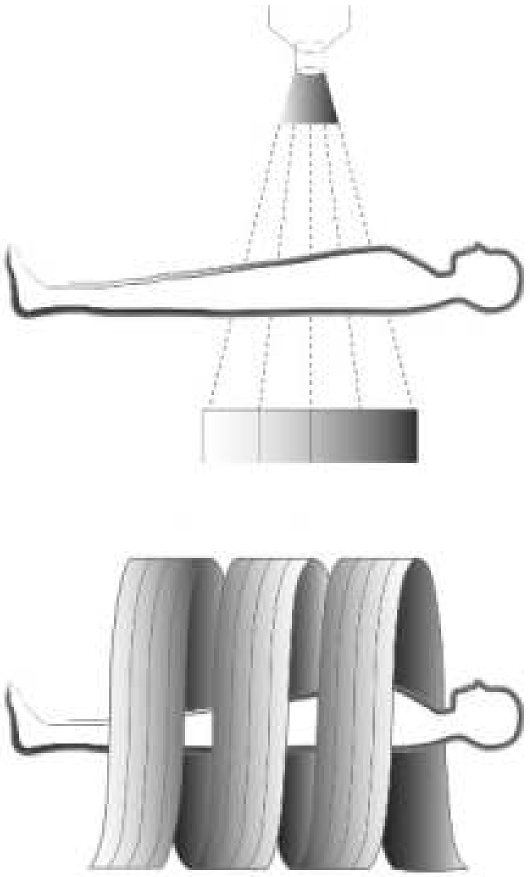


Figure A.4: Multislice CT scanner [8].

up to 64 slices and prototype 256-slice detectors in development. Recently, the first dual-source MSCT was launched. This has two pairs of x-ray sources and multislice detectors mounted at 90 degrees to each other [8].

The limitations of spiral CT have been significantly reduced with MSCT. Increased number of detectors and faster gantry rotation speed permit an increase in the overall scan speed ranging from 4 to more than 25 times, compared with single detector array equipment.

# Appendix B

# Cardiac CT images

> There are only two mistakes one can make along the road to truth; not going all the way, and not starting.
>
> Sidhartha Gautama (563–483 b.c).
> Dhammapada.

The heart has been extensively evaluated radiologically, but mostly in standard two-dimensional images (e.g., X-rays and angiograms of the chest). Standard radiographs of the chest (anterior and posterior) may reveal the silhouette of the heart, as well as the great arteries and pulmonary vasculature, but cannot show small structures and easily determine overlapping structures (Fig. B.1). CT images can be enhanced and manipulated in various ways. Generally an ionized contrast agent is injected intravenously during scanning, allowing smaller structures become visible. Moreover, the data can be reconstructed on the computer to provide images through different planes of the body or 3-dimensional images (Fig. B.2) [70].

MSCT for the evaluation of the heart and coronary arteries was introduced with the generation of 4-row detector instruments in 1999 [19].

The human body can be viewed in three standard anatomic planes, which are oriented perpendicular to each other: sagittal, coronal, and transverse (Fig. B.3). These planes are aligned with the thoracic midline structures, the aorta and esoph-

Figure B.1: Planar projection of standard chest radiograph [70].

Figure B.2: Three-dimensional computed tomographic image revealing complex anatomy and spatial relationships [70].

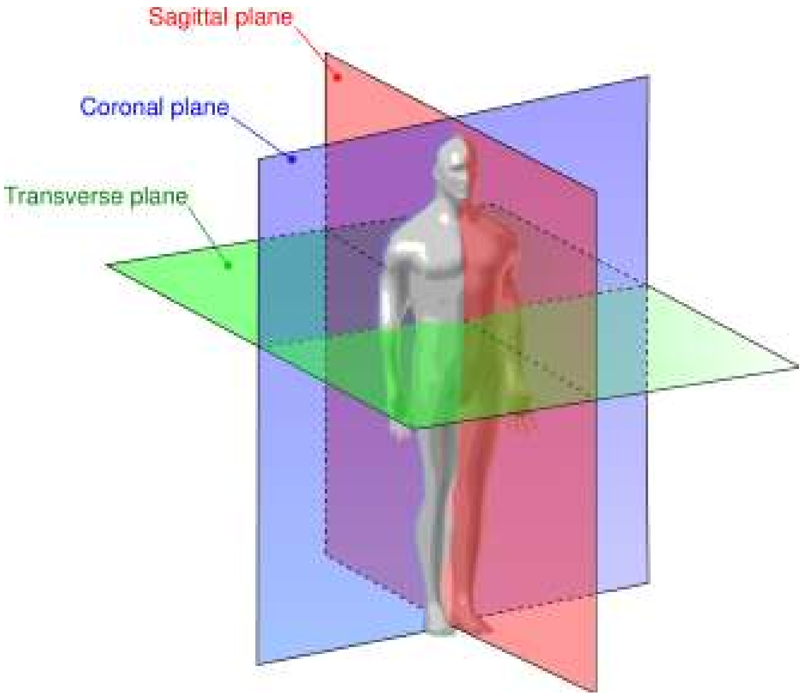


Figure B.3: Human standard anatomic planes [55].

agus. In contrast, the heart is oriented obliquely in the chest and therefore imaging in standard anatomic planes is suboptimal to visualize cardiac anatomy and pathology. The heart's three standard planes are its vertical and horizontal long axis and its so-called short axis. These cardiac axes are tilted against the standard anatomic planes, as shown in Fig. B.4 [28].

The axial plane is the first image plane in CT and usually gives a good overview of cardiac and coronary anatomy (Fig. B.5(a) and B.5(b)). The long vertical axis (Fig. B.6) or two-chamber view is easily produced from the axial plane, this corresponds to a vertical plane through the cardiac apex and the center plane of the mitral valve into the left atrium. This view is adequate to delineate the configuration of the left ventricle and to evaluate contraction of the anterior and lower segment of ventricular left myocardium [28].

The short axis view is oriented perpendicular to the vertical long axis and is parallel with the mitral valve plane and the cardiac base (Fig. B.7). The short axis therefore has a double-oblique angulation to account for the dorsoventral and

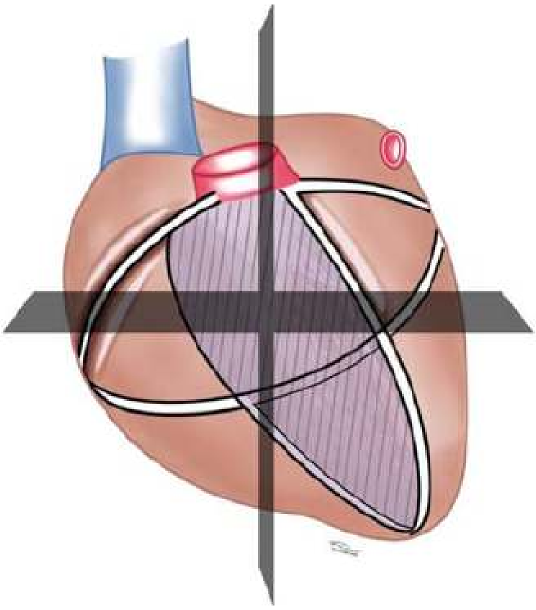
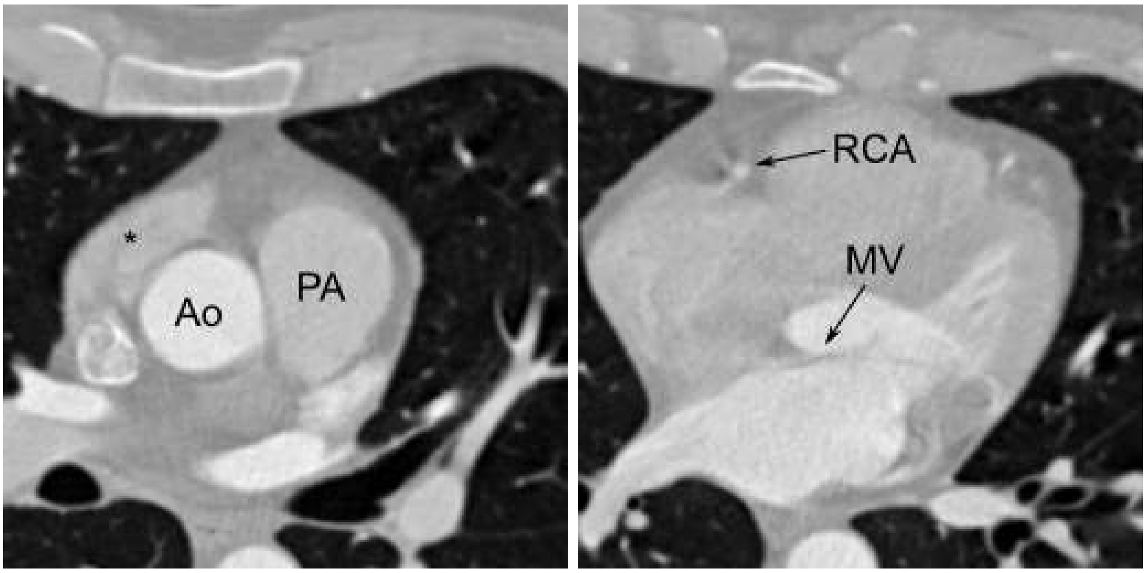


Figure B.4: Orientation of the vertical long axis and the short axis of the heart in relation to the standard anatomic planes [28].



(a)

(b)

Figure B.5: Axial anatomy of the heart. (a) Sample of the right atrial appendage, the ascending aorta and pulmonary artery (*, Ao and PA, respectively). (b) Sample of the right coronary artery and the mitral valve (RCA and MV).

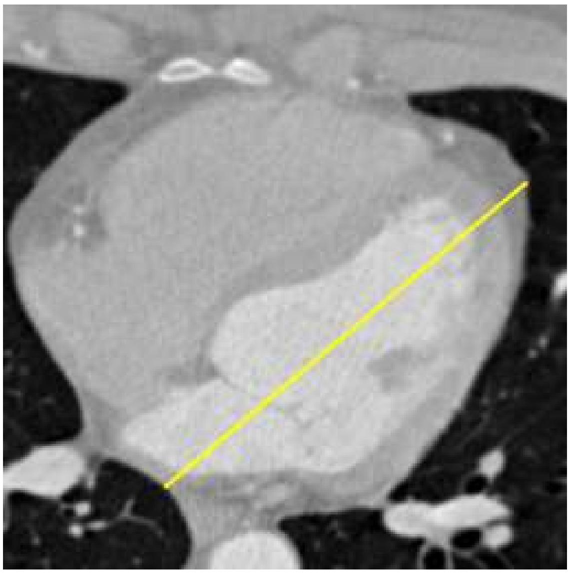


Figure B.6: Vertical long axis orientation. It is shown as a line from the cardiac apex through the middle of the mitral valve plane.

medioleftlateral tilt of the heart. Due to the alignment of the short axis view with the atrioventricular grooves, it can be used to display the right coronary artery down to the cardiac crux, the posterolateral branches of the distal right coronary artery, and the left circumflex coronary artery in its course in the left atrioventricular groove. The inferior facet of the right ventricle is parallel with the diaphragm and thus the transverse plane has a sharp angle (also called the acute margin), with the anterior free wall and the outflow tract of the right ventricle giving the right ventricle an almost triangular shape in this view. The left ventricle has a circular aspect on short axis view. Right and left ventricular motion can also be visualized with the short axis view and it is the basis for volumetric measurements used in global ventricular function evaluation [28].

The horizontal long axis corresponds to a tilted plane from cardiac apex through the middle of the mitral valve plane to the cardiac base. It is perpendicular to the vertical long axis and displays both ventricles and atria in their largest diameters. This view is a frontal view onto the heart from the anterior aspect directed to the inferior wall and thus gives a good overview of the size and configuration of both ventricles. It can be used to delineate the mitral and tricuspid valves. As shown

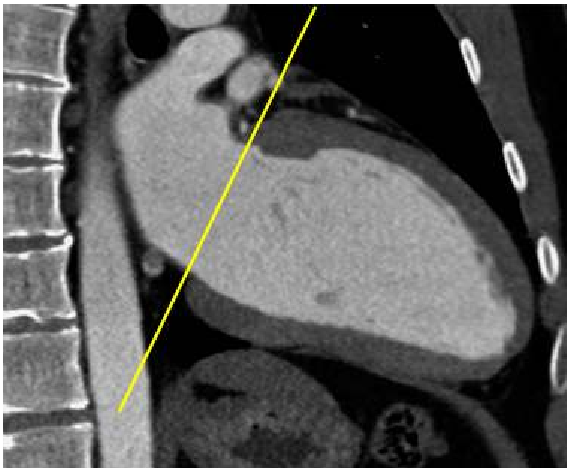


Figure B.7: Short axis orientation as planned from a vertical long axis view [28].

in Fig. B.8, the horizontal long axis imaging plane displays the left main coronary artery and the left anterior descending with its diagonal branches. The right coronary artery, like the proximal left circumflex coronary artery, is depicted in an orthogonal cut plane [28].

Figure B.8: Horizontal long axis orientation as planned from a vertical long axis view. Note the S-shape of the interventricular septum and the thin membranous part of the septum close to the cardiac base. [28].

# Appendix C

# Relation between Gaussian derivatives and the Hermite coefficients

> You have to do the right thing. It may not be in your power, may not be in your time, that there'll be any fruit. But that doesn't mean you stop doing the right thing. You may never know what results come from your action. But if you do nothing, there will be no result.
>
> Mahatma Gandhi (1869–1948).

The 1D Hermite coefficients are achieved by the inner product between the signal located by the Gaussian window and the Hermite polynomials

$$L_k = \left\langle L(x), H_k\left(\frac{x}{\sigma}\right)\right\rangle = \int_{-\infty}^{\infty} \mathcal{G}(x) L(x) H_k\left(\frac{x}{\sigma}\right) dx \tag{C.1}$$

It can be shown that

$$\begin{aligned}
L_k &= \left\langle L(x), H_k\left(\frac{x}{\sigma}\right)\right\rangle = \left\langle L^{(1)}(x), H_{k-1}\left(\frac{x}{\sigma}\right)\right\rangle = \cdots \\
\cdots &= \left\langle L^{(2)}(x), H_{k-2}\left(\frac{x}{\sigma}\right)\right\rangle = \cdots \\
\cdots &= \left\langle L^{(k)}(x), H_0\left(\frac{x}{\sigma}\right)\right\rangle
\end{aligned} \tag{C.2}$$

where

$$\begin{aligned}
L^{(k)}(x) &= \frac{\partial^k L(x)}{\partial^k} \\
H_0\left(\frac{x}{\sigma}\right) &= 1
\end{aligned} \tag{C.3}$$

Therefore:

$$L_k = \left\langle L^{(k)}(x), 1\right\rangle = L^{(k)}(x) \tag{C.4}$$

The Eq. C.4 states that the *kth* order Gaussian derivative of the image is the inner product of the image and the *kth* order Hermite polynomial [48].

# Appendix D

# Simplification of the Euler-Lagrange equations

> If you wish to know the past, then look at the present which is the result of it. If you wish to know the future, then look at the present which is the cause of it.
>
> Albert Einstein (1879–1955).

The Euler-Lagrange equations for the Eq. (4.8) are:

$$\begin{aligned}
&\Psi'\Big(|L_0(X+W)-L_0(X)|^2 \\
&+\gamma|l_{n,\theta}(X+W)-l_{n,\theta}(X)|^2\Big)\cdot \\
&\Bigg[|L_0(X+W)-L_0(X)|\frac{\partial L_0(X+W)}{\partial u(x)} \\
&+\gamma\left(|l_{n,\theta}(X+W)-l_{n,\theta}(X)|\frac{\partial l_{n,\theta}(X+W)}{\partial u(x)}\right)\Bigg]- \\
&\alpha\ div\left(\Psi'\Big(|\nabla u|^2+|\nabla v|^2\Big)\nabla u\right)=0
\end{aligned} \tag{D.1}$$

$$\begin{aligned}
&\Psi'\Big(|L_0(X+W)-L_0(X)|^2\\
&+\gamma|l_{n,\theta}(X+W)-l_{n,\theta}(X)|^2\Big)\cdot\\
&\left[|L_0(X+W)-L_0(X)|\frac{\partial L_0(X+W)}{\partial v(y)}\right.\\
&\left.+\gamma\left(|l_{n,\theta}(X+W)-l_{n,\theta}(X)|\frac{\partial l_{n,\theta}(X+W)}{\partial v(y)}\right)\right]-\\
&\alpha\ div\left(\Psi'\Big(|\nabla u|^2+|\nabla v|^2\Big)\nabla v\right)=0
\end{aligned}\tag{D.2}$$

Applying the chain rule for partial derivatives to equation Eq. (D.1) we get:

$$\begin{aligned}
\frac{\partial L_0(*)}{\partial u(x)}&=\frac{\partial L_0(*)}{\partial x}\\
\frac{\partial L_0(*)}{\partial v(y)}&=\frac{\partial L_0(*)}{\partial y}\\
\frac{\partial l_{n,\theta}(*)}{\partial u(x)}&=\frac{\partial l_{n,\theta}(*)}{\partial x}\\
\frac{\partial l_{n,\theta}(*)}{\partial v(y)}&=\frac{\partial l_{n,\theta}(*)}{\partial y}
\end{aligned}$$

from Eq. (C.4) of Appendix C we have that:

$$\frac{\partial L_0(X+W)}{\partial x}=L_{1,0}(X+W)\tag{D.3}$$

and

$$\begin{aligned}
\frac{\partial l_{n,\theta}(x)}{\partial x}&=\frac{\partial}{\partial x}\sum L_{m,n-m}\cdot g_{m,n-m}(\theta)\\
\frac{\partial l_{n,\theta}(x)}{\partial x}&=\sum L_{(m)+1,n-m}\cdot g_{m,n-m}(\theta)\\
\frac{\partial l_{n,\theta}(x)}{\partial x}&=l_{n,\theta_{(m)+1}}(x)
\end{aligned}\tag{D.4}$$

And in a similar way for partial derivatives to Eq. (D.2).

Eq. (D.3) and Eq. (D.4) simplifies Eq. (D.1) and Eq. (D.2):

$$
\begin{aligned}
&\Psi'\Big(|L_0(X+W)-L_0(X)|^2\\
&+\gamma|l_{n,\theta}(X+W)-l_{n,\theta}(X)|^2\Big)\cdot\\
&\Bigg[|L_0(X+W)-L_0(X)|L_{10}(X+W)\\
&+\gamma\left(|l_{n,\theta}(X+W)-l_{n,\theta}(X)|l_{n,\theta_{(m)+1}}(X+W)\right)\Bigg]-\\
&\alpha\ div\left(\Psi'\Big(|\nabla u|^2+|\nabla v|^2\Big)\nabla u\right)=0
\end{aligned}
\tag{D.5}
$$

$$
\begin{aligned}
&\Psi'\Big(|L_0(X+W)-L_0(X)|^2\\
&+\gamma|l_{n,\theta}(X+W)-l_{n,\theta}(X)|^2\Big)\cdot\\
&\Bigg[|L_0(X+W)-L_0(X)|L_{01}(X+W)\\
&+\gamma\left(|l_{n,\theta}(X+W)-l_{n,\theta}(X)|l_{n,\theta_{(n)+1}}(X+W)\right)\Bigg]-\\
&\alpha\ div\left(\Psi'\Big(|\nabla u|^2+|\nabla v|^2\Big)\nabla v\right)=0
\end{aligned}
\tag{D.6}
$$

# Appendix E

# Linearization via Taylor expansion

> Live as if you were to die tomorrow.
> Learn as if you were to live forever.
>
> Mahatma Gandhi (1869–1948).

There are nonlinear terms in Eqs. (4.11) and (4.12) as:

$$\begin{aligned} &L_0(X + W^{k+1}) - L_0(X) \\ &l_{n,\theta}(X + W^{k+1}) - l_{n,\theta}(X) \end{aligned}$$

Applying a first order Taylor expansion we get:

$$\begin{aligned} &L_0(X + W^{k+1}) - L_0(X) \approx \\ &\Big(L_0(X + W^k) + du^k \frac{\partial L_0(X)}{\partial x} + dv^k \frac{\partial L_0(X)}{\partial y}\Big) - L_0(X) \approx \\ &\Big(L_0(X + W^k) + du^k L_{01}(X) + dv^k L_{10}(X)\Big) - L_0(X) \approx \\ &\Big(L_0(X + W^k) - L_0(X)\Big) + du^k L_{01}(X) + dv^k L_{10}(X) \end{aligned} \tag{E.1}$$

and

$$\begin{aligned}
&l_{n,\theta}(X+W^{k+1}) - l_{n,\theta}(X) \approx \\
&\Big(l_{n,\theta}(X+W^{k}) + du^{k}\frac{\partial l_{n,\theta}(X)}{\partial x} + dv^{k}\frac{\partial l_{n,\theta}}{\partial y}\Big) - l_{n,\theta} \approx \\
&\Big(l_{n,\theta}(X+W^{k}) + du^{k} l_{n,\theta_{(m)+1}}(X) + dv^{k} l_{n,\theta_{(n)+1}}(X)\Big) - l_{n,\theta} \approx \\
&\Big(l_{n,\theta}(X+W^{k}) - l_{n,\theta}(X)\Big) + du^{k} l_{n,\theta_{(m)+1}}(X) + dv^{k} l_{n,\theta_{(n)+1}}(X) \qquad \text{(E.2)}
\end{aligned}$$

# References

[1] M. Abramowitz, I. A. Stegun, Handbook of Mathematical Functions, Dover, New York, 1965.

[2] E. H. Adelson, J. R. Bergen, Spatiotemporal Energy Models for the Perception of Motion, Journal of the Optical Society of America A: Optics, Image Science, and Vision 2 (2).

[3] L. Álvarez León, J. Esclarín Monreal, M. Lefébure, J. Sánchez Pérez, A PDE model for computing the Optical Flow, in: Sixteen Congreso de Ecuaciones Diferenciales y Aplicaciones, 1999, pp. 1349–1356.

[4] American Health Assistance Foundation, Anatomy of the Eye, `http://www.ahaf.org/glaucoma/about/understanding/anatomy-of-the-eye.html`, [accessed 01.07.2013] (Apr. 2012).

[5] P. Anandan, A Computational Framework and an Algorithm for the Measurement of Visual Motion, International Journal of Computer Vision 2 (3) (1989) 283–310.

[6] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, R. Szeliski, A Database and Evaluation Methodology for Optical Flow, in: IEEE International Conference on Computer Vision 0 (2007) 1–8.

[7] J. L. Barron, D. J. Fleet, S. S. Beauchemin, Performance of Optical Flow Techniques, International Journal of Computer Vision 12 (1) (1994) 43–77.
URL `http://dx.doi.org/10.1007/BF01420984`

[8] W. T. Bax, J. J. Davies, L. C. Roberts, Cardiac CT and CT coronary angiography: technology and application, Heart 94 (6) (2008) 781–792.

[9] S. S. Beauchemin, J. L. Barron, The Computation of Optical Flow, ACM Computing Surveys 27 (3) (1996) 433–467.

[10] J. Bigün, G. H. Granlund, J. Wilklund, Multidimensional Orientation Estimation with Applications to Texture Analysis and Optical Flow, IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (8) (1991) 775–790.

[11] M. J. Black, P. Anandan, The Robust Estimation of Multiple Motions: Parametric and Piecewise-Smooth Flow Fields, Computer Vision and Image Understanding 63 (1) (1996) 75–104.

[12] J. A. Bloom, T. R. Reed, A Gaussian Derivative-Based Transform, IEEE Transactions on Image Processing 5 (3) (1996) 551–553.

[13] T. Brox, Classical Variational Optical Flow, `http://lmb.informatik.uni-freiburg.de/resources/binaries`, [accessed 10.12.2011] (2011).

[14] A. Bruhn, J. Weickert, Towards Ultimate Motion Estimation: Combining Highest Accuracy with Real-Time Performance, in: Proceedings of Tenth IEEE International Conference on Computer Vision, vol. 1, IEEE Computer Society Press, 2005, pp. 749–755.

[15] A. Bruhn, J. Weickert, C. Feddern, T. Kohlberger, C. Schnörr, Variational Optical Flow Computation in Real Time, IEEE Transactions on Image Processing 14 (5) (2005) 608–615.

[16] A. Bruhn, J. Weickert, C. Schnörr, Combining the Advantages of Local and Global Optic Flow Methods, in: Proceedings of the Twenty-fourth DAGM Symposium on Pattern Recognition, Springer-Verlag, London, UK, 2002, pp. 454–462.

[17] A. Bruhn, J. Weickert, C. Schnörr, Lucas/Kanade Meets Horn/Schunck: Combining Local and Global Optic Flow Methods, International Journal of Computer Vision 61 (3) (2005) 211–231.

[18] G. Buckberg, J. I. E. Hoffman, A. Mahajan, S. Saleh, C. Coghlan, Cardiac Mechanics Revisited: The Relationship of Cardiac Architecture to Ventricular Function, Circulation 118 (24) (2008) 2571–2587.

[19] F. Cademartiri, G. Luccichentia, A. van Der Lugt, P. Pavone, P. M. Pattynama, P. J. de Feyter, G. P. Krestin, Sixteen-Row multislice computed tomography: basic concepts, protocols, and enhanced clinical applications, Seminars in Ultrasound, CT and MRI 25 (1) (2004) 2–16.

[20] P. Charbonnier, L. Blanc-Féraud, G. Aubert, M. Barlaud, Two deterministic half-quadratic regularization algorithms for computed imaging, in: IEEE International Conference on Image Processing, vol. 2, IEEE Computer Society Press, 1994, pp. 168–172.

[21] R. Courant, D. Hilbert, Methods of Mathematical Physics, vol. I, Interscience Publishers, New York, 1953.

[22] J. Crook, Why is the retina back-to-front?, `http://www.catalase.com/retina.htm`, [accessed 03.14.2013] (Apr. 1998).

[23] M. Douma, Retinal Ganglion Cells Calculate Color, `http://www.webexhibits.org/colorart/ganglion.html`, [accessed 01.07.2013] (2006).

[24] B. Escalante-Ramírez, J. L. Silván-Cárdenas, Advanced modeling of visual information processing: A multiresolution directional-oriented image transform based on Gaussian derivatives, Signal Processing: Image Communication 20 (9-10) (2005) 801–812.

[25] B. Escalante-Ramírez, J. L. Silván-Cárdenas, H. Yuen-Zhou, Optic Flow Estimation using the Hermite Transform, in: Proceedings of the SPIE in Applications of Digital Image Processing XXVII, vol. 5558, 2004, pp. 632–643.

[26] A. Estudillo-Romero, B. Escalante-Ramírez, Rotation-invariant texture features from the steered Hermite transform, Pattern Recognition Letters 32 (16) (2011) 2150–2162.

[27] C. L. Fennema, W. B. Thompson, Velocity Determination in Scenes Containing Several Moving Objects, Computer Graphics and Image Processing 9 (4) (1979) 301–315.

[28] R. Fischbach, Cardiac and Cardiothoracic Anatomy in CT, in: Multi-slice and Dual-source CT in Cardiac Imaging, 2nd ed., Springer Berlin Heidelberg, 2007, pp. 23–39.

[29] D. J. Fleet, A. D. Jepson, Hierarchical construction of orientation and velocity selective filters, IEEE Transactions on Pattern Analysis and Machine Intelligence, 11 (3) (1989) 315–325.

[30] D. J. Fleet, A. D. Jepson, Computation of component image velocity from local phase information, International Journal of Computer Vision 5 (1990) 77–104.

[31] W. T. Freeman, E. H. Adelson, The Design and Use of Steerable Filters, IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (1991) 891–906.

[32] B. Galvin, B. Mccane, K. Novins, D. Mason, S. Mills, Recovering Motion Fields: An Evaluation of Eight Optical Flow Algorithms, in: British Machine Vision Conference, 1998, pp. 195–204.

[33] J. J. Gibson, The perception of the visual world, The American Journal of Psychology, 64 (1951) 622–625.

[34] J. J. Gibson, The Senses Considered as Perceptual Systems, Houghton Mifflin Company, Boston, 1966.

[35] R. L. Gregory, Eye and Brain: The Psychology of Seeing, 2nd ed., McGraw Hill, NY, USA, 1976.

[36] D. Heeger, Model for the extraction of image flow, Journal of Optical Society of America, A: Optics, Image Science, and Vision 4 (8).

[37] Henry Guennadi Levkin, Imageprocessing/videocodecs/programming, `http://www.hlevkin.com/TestImages/additional.htm`, [accessed 01.20.2013] (Jan. 2004).

[38] B. K. P. Horn, B. G. Schunck, Determining Optical Flow, Artificial Intelligence 17 (1-3) (1981) 185–203.

[39] IEEE, Local correlation measures for motion analysis: a comparative study (1982).

[40] R. Jain, W. N. Martin, J. K. Aggarwal, Segmentation through the detection of changes due to motion, Computer Graphics and Image Processing 11 (1) (1979) 13–34.

[41] R. Jain, D. Militzer, H.-H. Nagel, Separating non-stationary from stationary scene components in a sequence of real world tv-images, in: Proceedings of the 5th International Joint Conference on Artificial Intelligence, vol. 2, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1977, pp. 612–618.

[42] R. Jain, H.-H. Nagel, On the Analysis of Accumulative Difference Pictures from Image Sequences of Real World Scenes, IEEE Transactions on Pattern Analysis and Machine Intelligence 1 (2) (1979) 206–214.

[43] J. J. Koenderink, A. J. van Doorn, Generic Neighborhood Operators, IEEE Transactions on Pattern Analysis and Machine Intelligence 14 (6) (1992) 597–605.

[44] M. D. Levine, Vision in Man and Machine, McGraw-Hill, New York, 1985.

[45] J. Limb, J. Murphy, Estimating the velocity of moving images in television signals, Computer Graphics and Image Processing 4 (4) (1975) 311–327.

[46] T. Lin, J. L. Barron, Image Reconstruction Error for Optical Flow, in: Vision Interface, 1994, pp. 73–80.

[47] C. Liu, W. T. Freeman, E. H. Adelson, Y. Weiss, Human-Assisted Motion Annotation, `http://people.csail.mit.edu/celiu/motionAnnotation/database/cameramotion.zip`, [accessed 10.12.2011] (2008).

[48] H. Liu, T.-H. Hong, M. Herman, R. Chellappa, A General Motion Model and Spatio-Temporal Filters for Computing Optical Flow, International Journal of Computer Vision 22 (1997) 141–172.

[49] B. D. Lucas, T. Kanade, An Iterative Image Registration Technique with an Application to Stereo Vision, in: Proceedings of the Seventh International Joint Conference on Artificial Intelligence (IJCAI '81), 1981, pp. 674–679.

[50] D. C. Marr, E. C. Hildreth, Theory of Edge Detection, in: Proceedings of the Royal Society of London, vol. 207 of B, Biological Sciences, 1980, pp. 187–217.

[51] J.-B. Martens, The Hermite Transform–Applications, IEEE Transactions on Acoustics, Speech and Signal Processing 38 (9) (1990) 1607–1618.

[52] J.-B. Martens, The Hermite Transform–Theory, IEEE Transactions on Acoustics, Speech and Signal Processing 38 (9) (1990) 1595–1606.

[53] E. Mémin, P. Pérez, A Multigrid Approach for Hierarchical Motion Estimation, in: Proceedings of the Sixth International Conference on Computer Vision (ICCV '98), IEEE Computer Society, Washington, DC, USA, 1998, pp. 933–938.

[54] E. Mémin, P. Pérez, Hierarchical Estimation and Segmentation of Dense Motion Fields, International Journal of Computer Vision 46 (2) (2002) 129–155.

[55] Y. Mrabet, Human anatomy planes, `http://commons.wikimedia.org/wiki/File:Human_anatomy_planes.svg`, [accessed 01.20.2013] (Aug. 2012).

[56] H.-H. Nagel, Analysing sequences of TV-frames, in: Proceedings of the fifteen International Joint Conference on Artificial Intelligence, vol. 2, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1977, pp. 626–626.

[57] H.-H. Nagel, Constraints for the Estimation of Displacement Vector Fields from Image Sequences, in: Proceedings of the Eighth International Joint Conference on Artificial Intelligence, vol. 2, 1983, pp. 945–951.

[58] H.-H. Nagel, On the estimation of optical flow: Relations between different approaches and some new results, Artifical Intelligence 33 (1987) 299–324.

[59] H.-H. Nagel, W. Enkelmann, An Investigation of Smoothness Constraints for the Estimation of Displacement Vector Fields from Image Sequences, IEEE Transactions on Pattern Analysis and Machine Intelligence 8 (5) (1986) 565–593.

[60] K. Nakayama, J. M. Loomis, Optical velocity patterns, velocity-sensitive neurons, and space perception: a hypothesis, Perception 3 (1974) 63–80.

[61] M. Ogata, T. Sato, Motion–detection model with two stages: Spatiotemporal filtering and feature matching, Journal of The Optical Society of America A-optics Image Science and Vision 9 (1992) 377–387.

[62] B. Ohnesorge, T. Flohr, Principles of Multi-slice Cardiac CT Imaging, in: Multi-slice and Dual-source CT in Cardiac Imaging, 2nd ed., Springer Berlin Heidelberg, 2007, pp. 71–126.

[63] N. Papenberg, A. Bruhn, T. Brox, S. Didas, J. Weickert, Highly Accurate Optic Flow Computation with Theoretically Justified Warping, International Journal of Computer Vision 67 (2) (2006) 141–158.

[64] P. Perona, J. Malik, Scale-Space and Edge Detection Using Anisotropic Diffusion, IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (7) (1990) 629–639.

[65] M. Petrou, A new imaging architecture and an alternative interpretation of the structure of the human retina, in: International Symposium on Information Technology. ITSim 2008, vol. 1, 2008, pp. 1–9.

[66] R. W. Rodieck, Quantitative analysis of cat retinal ganglion cell response to visual stimuli, Vision Research 5 (12) (1965) 583–601.

[67] B. Sakitt, H. Barlow, A Model for the Economical Encoding of the Visual Image in Cerebral Cortex, Biological Cybernetics 43 (2) (1982) 97–108.

[68] C. Schnörr, Segmentation of visual motion by minimizing convex non-quadratic functionals, in: Proceedings of the 12th International Conference on Pattern Recognition, vol. A, 1994, pp. 661–663.

[69] C. Schnörr, R. Sprengel, A nonlinear regularization approach to early vision, Biological Cybernetics 72 (2) (1994) 141–149.

[70] P. Schoenhagen, A. E. Stillman, S. S. Halliburton, R. D. White, CT of the heart: Principles, advances, clinical uses, Cleveland Clinic Journal of Medicine 72 (2) (2005) 127–140.

[71] U. J. Schoepf, C. R. Becket, L. K. Hofmann, E. Yucel, Multidetector-row CT of the heart, Radiologic Clinics of North America 41 (3) (2003) 491–505.

[72] J. L. Silván-Cárdenas, B. Escalante-Ramírez, Optic–flow Information Extraction with Directional Gaussian–Derivatives, in: Proceedings of the Fifteenth International Conference on Pattern Recognition, vol. 3, IEEE Computer Society, Washington, DC, USA, 2000, pp. 190–193.

[73] J. L. Silván-Cárdenas, B. Escalante-Ramírez, The Multiscale Hermite Transform for Local Orientation Analysis, IEEE Transactions on Image Processing 15 (5) (2006) 1236–1253.

[74] G. Szegö, Orthogonal Polynomials, vol. 53 of Colloquium Publications, American Mathematical Society, New York, 1959.

[75] C. Tomasi, R. Manduchi, Bilateral Filtering for Gray and Color Images, in: Proceedings of the Sixth International Conference on Computer Vision, ICCV '98, IEEE Computer Society, Washington, DC, USA, 1998, pp. 839–846.

[76] S. Uras, F. Girosi, A. Verri, V. Torre, A Computational Approach to Motion Perception, Biological Cybernetics 60 (1988) 79–87.

[77] A. M. van Dijk, J.-B. Martens, Image representation and compression with steered Hermite transforms, Signal Processing 56 (1) (1997) 1–16.

[78] R. van Rullen, S. J. Thorpe, Rate Coding Versus Temporal Order Coding: What the Retinal Ganglion Cells Tell the Visual Cortex, Neural Computation 13 (6) (2001) 1255–1283.

[79] A. B. Watson, J. Albert J. Ahumada, Model of human visual–motion sensing, Journal of the Optical Society of America A: Optics, Image Science, and Vision 2 (2) (1985) 322–341.

[80] J. Weickert, On Discontinuity-Preserving Optic Flow, in: Computer Vision and Mobile Robotics Workshop, 1998, pp. 115–122.

[81] J. Weickert, Coherence-enhancing diffusion of colour images, Image and Vision Computing 17 (3–4) (1999) 201–212.

[82] J. Weickert, C. Schnörr, A Theoretical Framework for Convex Regularizers in PDE-Based Computation of Image Motion, International Journal of Compututer Vision 45 (3) (2001) 245–264.

[83] R. Wilson, G. H. Granlund, The Uncertainty Principle in Image Processing, IEEE Transactions on Pattern Analysis and Machine Intelligence 6 (6) (1984) 758–767.

[84] J. Xiao, H. Cheng, H. Sawhney, C. Rao, M. Isnardi, S. Corporation, Bilateral Filtering–based Optical Flow Estimation with Occlusion Detection, in: European Conference on Computer Vision (ECCV), ECCV'06, 2006, pp. 211–224.

[85] R. A. Young, The Gaussian derivative theory of spatial vision: Analysis of cortical cell receptive field line–weighting profiles, Tech. Rep. GMR-4920, General Motors Research Laboratories, Detroit, Mich, USA (1985).

[86] R. A. Young, Simulation of human retinal function with the Gaussian derivative model, in: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 1986, pp. 564–569.

[87] R. A. Young, The Gaussian derivative model for spatial vision: I. Retinal mechanisms., Spatial Vision 2 (4) (1987) 273–293.